

**DRAFT**  
**Report to the ad hoc committee of the**  
**National Academies of Sciences, Engineering, and Medicine**

**Evidence to Advance Reform in the Global Security and Justice Sectors:**  
**A Workshop-based Consensus Study Series**

*Evidence and Police Use of Force: Theories without Data*

By

Geoffrey P. Alpert  
University of South Carolina

Kyle McLean  
Clemson University

Seth Stoughton  
University of South Carolina

August 2021

Two ongoing crises have dominated recent national and international news—the SARS-CoV-2 pandemic and the use of excessive force by police. These crises intersect. The reaction to death of George Floyd at the hands of Minneapolis police officers gave rise to demonstrations not just across the United States, but around the world, even in the midst of a global pandemic. But even when in-person gatherings were otherwise discouraged, many public health officials suggested that the use of excessive force by police was such a serious and pervasive social problem that gathering *en masse* to protest was an acceptable risk (Resnick, 2020).

Both the SARS-CoV-2 pandemic and the issue of excessive force have raised important questions about the type, amount, and degree of scientific evidence that serves as a sufficient basis for governmental policymaking. Society is troubled by a pandemic that has taken hundreds of thousands of lives in the United States and abroad and resulted in an untold number of serious short- and long-term health consequences. Yet, there is a public debate about whether the scientific research supporting the effectiveness of masks and the safety of mRNA vaccines is strong enough to support public health recommendations or mandates. In many cases, these debates are fueled by ignorance and a deep misunderstanding of the research that underlies public health policymaking.

In a similar manner, society is troubled by police uses of force that occur too frequently, are avoidable, excessive, or disproportionately targeted at specific populations. Addressing these concerns necessarily requires evidence of the nature and extent of the problem. Yet, as with masks and vaccines, ignorance and misunderstandings about the available evidence abound. Research demonstrates, for example, that African-Americans killed by police are more likely to be unarmed than white Americans killed by the police (Nix et al., 2017). What often goes unappreciated, however, is the scope of the problem. A recent study demonstrated a substantial

portion of individuals believe that more than 1,000 unarmed black Americans are killed every year by the police (McCaffree & Saide, 2021), when in fact the number of unarmed individuals, of any race, killed by the police is relatively small (n=93) (Nix et al., 2017). The disconnect between the *perceived* scope of the issue and the actual scope of the issue has important implications for how optimally to calibrate a public policy response. And a distinct lack of accessible data further complicates the picture; only one state, New Jersey, has anything approaching standardized, robust data on statewide uses of force, and only because of a result of a prolonged effort by a private media organization to obtain and code all uses of force in the state over a four-year period (NJ Advance Media, 2018). Addressing appropriately important crises—whether the COVID-19 pandemic or the use of excessive force by police—requires scientific evidence that is sufficient, robust, granular, and representative enough to understand the issues and develop successful responses while also being easily communicated and understood. The bases for public policymaking on technical matters must withstand scientific and social scrutiny and stand above political whims and ideological beliefs.

There is one substantial difference in the responses to the COVID crisis and the use of force by police: to alleviate the COVID pandemic, billions of dollars have been spent to develop the scientific evidence necessary to inform the policy recommendations surrounding masking, physical distancing, and vaccinations. To date, no similar investment has been made in understanding or resolving the issue of excessive force. Indeed, Operation Warp Speed—the effort to quickly develop a COVID vaccine—had a reported budget of \$12.4 *billion* (Barone, 2020), while the entirety of the research and evaluation budget for the Office of Justice Programs in 2020 was only \$79 *million* (Office of Justice Programs, 2021).

Despite the regrettable lack of investment, social scientists across a range of disciplines have dedicated their careers to understanding and reducing the use of excessive force by the police. Scholars in fields ranging from sociology (e.g., Legewie, 2016) to economics (e.g., Fryer, 2018) to political science (e.g., Mullinix et al., 2021) and, of course, criminology (e.g., Nix et al., 2017), have attempted to understand the scope, causes, and possible solutions to the issue of excessive force. Despite this considerable attention, only recently has the use of force been studied with sufficiently improved methods and rigorous research, and only recently has that research started to have a systemic impact on public policy.<sup>1</sup>

This modern effort to develop sound evidence about policing has been linked to a growing social concern and an increasing acceptance of evidence-based methods and practices from police command staff, policy makers, and trainers. Research findings have formed the basis for many decisions which were previously left to dubious theories, unsupported assumptions, and political whims. Unfortunately, and as we will show, this evidence base remains far too sparse and full of gaps to meet the needs and requests of the public and police practitioners. If we are truly to consider the issue of police use of force as being as urgent and pressing as the COVID pandemic, then we must dedicate greater resources to the scientific study of policing as a distinct set of social practices rather than as a topic suited principally for philosophical discussion and theoretical musings.

This is not to say that no policy should be enacted and no practices undertaken without a body of indisputable scientific evidence as support. That would be unrealistic. It is to say,

---

<sup>1</sup> For our purposes, excessive force refers to force that exceeds what the courts and/or the public deem reasonable in a given society. When we refer to the study of the use of force, this includes excessive force, as well as reasonable uses of force. Academic studies in this area tend to focus on the use of force rather than excessive force as excessive force requires a determination of unreasonableness that is unavailable in most datasets. Accordingly, we focus on the study of the use of force with the intent of reducing the problem of excessive force.

however, that public policy should be predicated on robust evidence when it is available, on reliably developed findings when possible, and on sound theoretical assumptions when needed. To the detriment of police agencies, officers, and communities, it is far too common to see use-of-force policymaking based in nothing more than conventional wisdom (read: the way things have always been done), optimism, untested assumption. Even recent reforms reflect a number of good ideas, but few empirical studies of their effectiveness.

Understanding, explaining, and controlling police uses of force and excessive force requires a multifaceted approach that incorporates a thorough understanding of policies, training, supervision, selection, and accountability. There are, of course, a range of additional considerations; unfortunately, the dictates of time and space limit the scope of our discussion. In this paper, we will review the research findings in each of the identified areas to show the evidence that drives our knowledge and decisions concerning uses of force remains in its infancy. In each of these substantive areas, scholars, policymakers, and philosophers have generated a lot of good ideas, tested only a few of them, and replicated an even smaller number of them. This leaves us with an abundance of testable ideas, but sizeable gaps in our knowledge of their efficacy. We must increase what we know in a rigorous manner to have faith in our research findings so they can be used to develop rational and reasonable responses to this critical issue.

## **Policies**

Policies regarding the use of force represent perhaps the clearest area in which decisions are dominated by the conventional wisdom about what seems like a good idea instead of data about what works. This is dictated, in no small part, by the need for agencies to adopt defensible policies that comply with state and federal law (Stoughton et al., 2020), reflect existing

jurisprudence on the use of force (Stoughton, 2021), and garner public acceptance.

Unfortunately, jurisprudence and public sentiment do not always make for an effective means of reducing the use of force.

What have come to be called “best practices” for use of force policies are guided by a number of key organizations, each of which has released so-called “model policies.” The International Association of Chiefs of Police (IACP) has released their National Consensus Policy, the Police Executive Research Foundation (PERF) has released its Guiding Principles on the Use of Force, the College of Policing in the UK has a list of general principles and three core questions for every officer to evaluate during a potential use of force encounter, and the Australia and New Zealand Policing Advisory Agency (ANZPAA) has a list of recommended use of force principles. These documents overlap in several areas, such as a focus on ensuring that officers use force only when it is reasonable, necessary, and proportionate. Also, they each make a number of distinct recommendations that range from including a statement on the sanctity of life in agency policy, adopting a range of components of “de-escalation,” articulating that officers have a duty to intervene, to, in the case of PERF’s Guiding Principles, advocating for agencies to provide all patrol officers with shields. Regardless of their content, each of these documents aims to build upon clear legal, moral, and philosophical foundations. Notably absent from that foundation, though, is empirical evidence of effectiveness.

Empirical examinations of use-of-force policies are few and far between. Early examinations of force policies used the use of force continuum to create a force factor score for examining the prevalence of potentially problematic use of force incidents (Alpert & Dunham, 2004; Terrill, 2005). This work, while informative, provides little information on whether the continuum, or certain characteristics of a continuum, reduced uses of force. Similarly, Terrill and

Paoline (2013) conducted a national survey of police departments to assess the prevalence of use of force continua in the early 2010s. Their study demonstrated that while use of force continua are adopted in most police departments, there is substantial variation in the structure and specification of the continua. Again, this provides little indication of the substantive impact, if any, of these continua on police use of force.

A small number of studies have determined that more restrictive force policies lead to reductions in the use of force, likely beginning with the work by James Fyfe (1979). White (2001) used interrupted time series analyses to examine the impact of a series of policy changes on officer-involved shootings in the Philadelphia Police Department, concluding that restrictive policies can reduce the number of shootings. Similarly, Terrill and Paoline (2017) studied less-lethal force policies in eight departments across the country, concluding that more restrictive policies were associated with fewer use of force incidents.

Still, analyses in this area struggle to identify specific causes. Methodologically, studies of agency policies are plagued by endogeneity problems. For legal reasons, different policies cannot be randomly assigned to individual officers. Policies also tend to be revised substantially after major incidents or leadership changes, introducing variables beyond the reasons for the policy revision. And policy changes are often wholesale, with the rewriting of an entire policy or set of policies, rather than with isolated changes to specific policy components. Additionally, policy revisions—of whatever variety—cannot easily be disentangled from modifications to training or supervisory expectations that accompany those revisions. Accordingly, analyses that examine alterations to use of force policies are challenged to determine whether any impact results from not only the policy change, but a host of external factors, and if so what specific policy change drives which impact. Similarly, analyses that compare use of force policies across

multiple jurisdictions must contend with the issue that more restrictive force policies may reflect different agency cultures or that training within these departments that may be contributing to any detected effects. To be clear, these observations are not a criticism of the scholars who have worked in this area, but rather a statement on how difficult it is to design and conduct this type of research.

In fact, the only recent peer-reviewed study that identified a specific change in use of force policy and measured the impact on reported uses of force in our review was conducted by Shjarback and colleagues (2020). These authors assessed the impact of a change in policy at the Dallas Police Department that required officers to report when they pointed a firearm at a subject which resulted in a slow but lasting reduction officer-involved shootings.

Finally, it should be noted that several journalistic studies of the use of force (e.g., the Washington Post's Fatal Force Project, Campaign Zero) have provided some recommendations on the contents of use of force policies based on data. These studies are not peer-reviewed and come with many of the same caveats noted above regarding separating the independent impact of individual policies. Overall, their recommendations are largely consistent with the findings that more restrictive policies result in fewer uses of force. Still, it is striking that these groups have created more evidence-based suggestions than many of the professional organizations who spend significant effort promulgating and promoting their model force policies for nationwide adoption, and the government.

## **Training**

Compared to agency policies, police training has been the focus of considerably more research. Over the past decade, in-service training programs have been studied for their potential to reduce uses of force. One study examined the impact of a procedural justice training program



in Chicago Police Department on trainees' attitudes compared to a control group of untrained officers (Skogan, Van Craen, & Hennessy, 2015). This study moved the literature forward as one of the few evaluations of police training at the time but was disadvantaged by the lack of experimental design and its failure to specifically measure use of force. Procedural justice—the outcome of interest in the study—has been linked to citizens' compliance (Tyler, 2006; Tyler & Huo, 2002), such that there is a theoretical connection between procedural justice and the use of force, but the study did not look at uses of force as a distinct outcome.

Giacomantonio and colleagues (2020) later pushed this type of evaluation further by examining the impact of a verbal judo training program in the Halifax Regional Police (CAN) on officers' performance in simulated scenarios. On the one hand, their study used advanced technology—an immersive virtual reality simulator—to expose treatment and control officers to the same stimuli to test for differences in behavior. On the other hand, the study lacked an experimental design and showed no differences in the use of force between trained and untrained officers. Additionally, there is no robust evidence as to the relationship between changes in performance in simulated scenarios, especially using virtual reality simulators, and changes to performance in the field.

Following these initial studies, a number of experimental evaluations have been conducted to test for differences in administrative reports of uses of force. These include an evaluation of Polis Solutions' T3 training program by McLean and colleagues (2021), as well as an evaluation of PERF's ICAT program by Engel and colleagues (*Unpublished*). McLean and colleagues' study was strengthened by an experimental design and repetitive training exposure. The results showed no significant differences in the number of force incidents involving trained officers compared to the number involving non-trained officers. McLean and colleagues' study

randomized individual officers to treatment and control conditions but was limited by potential contamination issues and agency measures of use of force (see McLean et al., 2021 for a discussion). Engel and colleagues' use of a stepped-wedge design limits contamination issues compared to the McLean and colleagues' study but left unaddressed the possibility that the observed reductions could be attributed to exogenous changes that coincided with the introduction of training, such as distinct policy changes or the occurrence of critical incidents in the jurisdiction (see also Worden and colleagues, 2020, for a stepped-wedge study impacted by external factors).

Wood and colleagues (2020) also studied a procedural justice training program in the Chicago Police Department using a stepped-wedge design. The study had similar issues with the potential for external factors influencing the use of force (see Correction to Wood et al., 2020), and found a small reduction in the use of force in the Chicago Police Department. Finally, White and colleagues (2021) are currently evaluating a de-escalation training program in the Tempe Police Department but have yet to release results from the analysis of use of force incidents.

These studies represent the first rigorous reviews of training programs designed to reduce the number and level of uses of force. Several commonalities dominate the studies and create serious issues for providing scientific recommendations about police training. First, each study evaluated a *different* training program, and no known studies attempt to replicate existing evaluations of police training. Accordingly, the evidence base for any training program is limited to a single study. Even in the most promising of studies—perhaps *especially* with the most promising of studies—replication is important to improve confidence in the generalizability of the findings.

Second, each study evaluated an *in-service* training program. In-service training programs are undoubtedly important, but the focus on in-service training in experimental evaluations has resulted in a complete lack of evidence on what programs and techniques are effective in pre-service (academy) training of new officers.

Third, each study except for the one that evaluated Polis Solutions' T3 program evaluated a program that involved one or two days of training. Best practices for adult learning in training from other fields suggest that high-repetition training over more prolonged periods is important for both retention and understanding (Ericsson, 2004; Ericsson & Ward, 2007).

To be clear, a number of promising training programs exist that may reduce officers' uses of force or excessive force. Programs like Georgetown University's Active Bystandership for Law Enforcement promote peer intervention when officers perceive a colleague to be using or about to use excessive force (Active Bystandership for Law Enforcement Project, 2021). The Marietta, GA, Police Department has adopted a Brazilian Jiu-Jitsu program that they claim will make officers more willing to communicate with individuals before turning to physical control techniques *and* more willing to use empty hand controls rather than resorting to intermediate weapons such as TASERS (Huddleston, 2021). Additionally, the LaGrange, GA, Police Department has developed "Shoot to Incapacitate," a controversial program that trains officers to shoot at subjects' pelvic girdles or upper legs in certain deadly force situations to reduce lethality *vis-à-vis* center mass (Schrade, 2021). All three programs reflect innovative ideas and are based on sound theoretical assumptions, but none of the three is backed by rigorous scientific research into their effectiveness in achieving their stated goals.

In sum, our review of studies measuring the impact of police training on the use of force is consistent with our review in each of the other areas. Good ideas on how to improve training

abound, but empirical research continues to lag the desire to implement evidence-based training programs. Dedicating more money, time, and resources to this critical topic is vital first step in reducing the amount of and improving the quality of police use of force.

## **Officer Selection**

Research has also sought to determine whether certain officers are more likely to use force than other officers with the end goal of identifying personal characteristic—e.g., age, race, gender, education, etc.—that are positively correlated with a reduced reliance on force.

Although many studies have attempted to tackle this issue, it is important to consider that this path may be one of the least promising areas for reducing the use of force because it depends on two distinct causal predicates. The evidence must first establish that officers with certain characteristics are more or less likely to use force. The evidence must then identify the recruiting or selection strategies that are most effective at identifying any relevant latent characteristics (e.g., personality) *and* increasing the pool of or appeal to officer-applicants with desirable characteristics while simultaneously deterring or screening out officer-applicants with undesirable characteristics. Accordingly, the challenge for conducting or operationalizing research in this area is much greater than for studying or implementing a policy revision, training, supervision, or accountability.

To be clear, this does not mean that the study of recruiting and selection in policing is without merit or promise (see Lough and Ryan, 2010). Indeed, having a police department that truly reflects the community it serves is vital morally, ethically, and scientifically as a manner of improving police-community relations. Improving racial and gender representation in policing is valuable for a number of reasons, but there is little evidence to support the assertion that it will reduce uses of force.

The study of officer demographic and uses of force has generated a much larger body of research than those explored in other areas of this report. The reason for this likely has to do with the issue of funding that is raised in our introduction. Experimental evaluations of policies, training, supervision, and accountability require considerable resources in the form of grant funding, which we have already noted is remarkably rare in the study of policing. Obtaining use-of-force reports and even videos from a police department and examining the characteristics of these incidents requires considerable researcher effort and statistical rigor, but it does not demand anywhere near the level of agency cooperation or monetary resources as implementing a randomized field trial or conducting systematic observational research.

A range of hypotheses have been offered for why various personal characteristics might affect officers' use-of-force decisions, but despite (or perhaps because of) the size of this body of work, the findings are mixed and contradictory.

**Gender.** Studies of officer gender have found both that gender has no impact on the use of force (Kaminski, Digiovanni, & Downs, 2004; McCluskey & Terrill, 2005; Paoline & Terrill, 2007) and that male officers are more likely to use force than female officers (Garner et al., 2002).

**Officer race.** Scholars have determined that officer race has no impact on force (Friedrich, 1977; McCluskey & Terrill, 2005; Worden, 1995), that minority officers use more force (Cohen & Chaiken, 1972, Sun & Payne, 2004), and that minority officers use less force (Ba et al., 2021). In examining the totality of this body of research, then, it is challenging to be confident of any conclusion.

**Age at time of hiring.** Building on brain development neuroscience, scholars have argued that officers who are older at the time they are hired may be less inclined to use force

(Wasilczuk, forthcoming). A 1972 study concluded that “men who were oldest at time of appointment . . . were substantially less likely than average to have civilian complaints,” but age at time of hiring was one of the weakest among half-dozen factors that predicted a reduction in misconduct (Cohen & Chaiken, 1972). A 2009 study of career-ending misconduct at the New York Police Department concluded that older age at appointment had some protective effect, but it was very limited (-0.06) (Kane & White). Even with more robust evidence about the effects of officer age at the time of hiring, no known study has attempted to quantify the relative difficulty of recruiting and hiring older candidates.

**Education.** The idea that education may impact the use of force has been studied for at least 40 years (see e.g., Sherman & Blumberg, 1981), and continues to be updated. More recent studies suggest that officers with a college education are less likely to use force (Chapman, 2012; Paoline & Terrill, 2007; Rydberg & Terrill, 2010). While this provides some evidence about the impact of a college education on use of force (the first of the causal pathways discussed above), it does not address the recruitment of college educated individuals into police departments (the second necessary causal pathway).

Beyond the limited research and contradictory findings, it is worth considering the practical barriers that limit the potential of hiring initiatives to reduce meaningfully the use of force. First, and perhaps most obviously, agencies have limited funds, which limits their ability to recruit and select preferred candidates. Second, agencies often hire from a limited pool. Many police departments are already struggling to hire enough officers without putting additional requirements in place; both PERF and the IACP have released reports in recent years on the struggles of recruiting police officers. Third, there can be legal constraints that limit an agency’s ability to hire, including a constitutional prohibition on using racial or gender quotas in

public-sector hiring (Stoughton, 2014). Fourth, putting some hiring preferences into place may hurt *other* hiring efforts. For example, there is evidence that adopting higher education requirements may hurt efforts to improve the racial diversity of police departments (Carter & Sapp, 1991)—something we have already identified as a moral and ethical imperative. Some of these obstacles are surmountable; an agency cannot change an officer’s race, obviously but they may be able to require or incentivize ongoing education with tuition coverage or making degrees a necessary component of promotion. Such programs can be expensive, however, which implicates the first obstacle. And while there is evidence that more highly-educated officers use force less often, it is not clear whether that effect attaches only to officers who initially enter policing with a degree or also to officers who earn a degree while engaged in policing. Again, that observation brings us back to the primary theme of good ideas that have little evidence to support their efficacy.

## **Supervision**

Supervision also represents one of the more promising areas of research when considering the weight of evidence on its impact on uses of force. Findings that effective supervision impacts an officer’s use of force are hardly surprising given the role that bosses have in encouraging or even requiring particular types of behavior on their shifts. At the same time, the practical implementations of these findings are unclear given the limited research on how identify, promote, and train good supervisors. Practically speaking, we might know that it matters without knowing how to do it.

Early studies on the role of supervision in police use of force used surveys to correlate officer attitudes and behaviors to supervisor’s attitudes and behaviors (Engel, 2000; 2001; 2002; Engel & Worden, 2003). Using these techniques, Engel (2000) identified that officers with

supervisors she classified as “active” were more likely to use force. Active supervisors were characterized by increased activity in the field, the tendency to take over situations, and an emphasis on serving a dual role as both a street officer and a supervisor. Modeling this active behavior then, unsurprisingly encourages street officers to be more active and engage in more uses of force.

A more recent study examined the impact of procedurally fair supervision on officers’ uses of force (Owens et al., 2018). In this study, officers identified as working in high-risk-circumstances were randomly assigned to either a treatment or control condition. In the treatment condition, officers had a non-disciplinary meeting with their supervisor in which the supervisor modeled procedurally fair behavior, reviewed a recent incident the officer engaged in, and pointed out areas of success and potential areas for improvement. This cognitive debrief was juxtaposed to the control group who were debriefed in a “matter-of-fact” manner, without any emphasis on procedural justice. As a result of the intervention, treated officers were less likely than control officers to engage in uses of force in the six weeks following their supervisory meeting. In fact, the officers in the experimental group passed forward the procedurally just treatment they learned from their supervisors to the public in their subsequent interactions (Owens et al., 2017).

This study then is particularly promising as it not only reinforces that supervisors have a meaningful impact on officers’ uses of force, but also identifies a specific intervention that can promote this type of supervision. While this finding is promising, it suffers from many of the same caveats as the studies of police training. The intervention as studied in a single police department, has never been subject to replication, and only demonstrated reduced uses of force for six weeks. There is no evidence on whether the effects of supervisory meetings are purely



transient, on the optimal number, frequency, and duration of such supervisory meetings, or on the opportunity costs of any long-term, repetitive intervention (for example, might this commitment of a supervisor's time reduce their attention to other supervisory responsibilities?).

Despite these unanswered questions, the clear message from this literature is that supervision is a promising avenue for reducing the use of force and merits dedicating considerably more resources to evaluating and replicating promising programs.

### **Accountability: Holding the Police Responsible**

Holding the police accountable has only been taken seriously for a relatively short time. The 1981 U.S. Commission on Civil Rights *Who's Guarding the Guardians* was one of the first examinations of police accountability, resulting in a Report that is a "must read" for students of police reform. This report and previous reports that looked at police misconduct or other aspects of it (including the 1973 National Advisory Commission: Task Force on Police, see Walker, 2012) had one thing in common: a lack of meaningful data and evidence-based recommendations. Another theme among the early reports on reform was the division of internal and external control mechanisms (see Noble and Alpert, 2009 and Kane and White, 2009). Rather than review the litany of accountability systems that have been described over the years, we will mention only a few topical areas including the problem officer, review boards, qualified immunity, and consent decrees.

#### *The Problem Officer*

Over the years, police leaders and academics have been concerned with the small number of officers to whom the majority problems have been attributed. Data from New Jersey, for example, shows that ten percent of officers in the state accounted for 38 percent of all uses of

force, and a small group of 252 officers used force at a rate five times higher than the state average (NJ Advance Media, 2018).

One internal accountability tool noted in *Who's Guarding the Guardians* is the Early Warning System, later re-named to the Early Identification System (EIS). An EIS uses internally-generated data on various aspects of officers' performance—e.g., ranging from uses of force to use of sick leave—to alert supervisors to officers who are potentially problematic before they become a serious concern. While there is evidence that EIS can be effective if managed properly with moderate goals (Walker et al., 2001), a recent systematic review of EIS shows that there were only eight studies that matched selection criteria for proper methods, and six of those were single agency studies. The results of the reviewed studies are encouraging, but measures of success/effectiveness are inconsistent (Gullion and King, 2020). Here, as elsewhere, a consistent theme in research on police use of force is data drawn from single-agency studies that gives rise to weak results.

Chalfin and Kaplan (2021) recently published a study examining the potential outcomes if EIS were properly implemented such that agencies committed their limited resources to identifying and extricating problem officers. Concluding that removing “high-complaint ‘bad apples’” from police departments would result in only small reductions in civilian complaints, Chalfin and Kaplan contend that agencies should focus their resources on broader reforms rather than committing substantial resources to EIS. Sierra-Arévalo and Papachristos (2021) counter by challenging the analytical approach that Chalfin and Kaplan relied upon, concluding that the Chalfin and Kaplan estimates are comparable to the modest results seen in other programs. Specifically, Sierra-Arévalo and Papachristos make an argument similar to ours: the existing literature is all but devoid of empirically based recommendations that result in substantial

reductions in the use of force or citizen complaints. Rather than committing resources to policies and programs with indeterminate outcomes, Sierra-Arévalo and Papachristos suggest, a modest (4-6%) but proven reduction may be preferable.

This argument is worth discussing in greater length because it reflects the conundrum replete in our findings. A policymaker or police chief who wishes to commit resources to reducing the uses of force in their agency and wants to make an evidence-based policy determination must choose between reforms associated with modest reductions, such as the development of EIS or the implementation of the procedural justice training program evaluated by Wood and colleagues (2020), or programs that are theoretically sound but which have yet to be evaluated, such as Active Bystandership for Law Enforcement. Even assuming the policymaker is scientifically literate and well informed, their position is not an enviable one (see Fleming and Rhodes, 2018 and Aronie and Alpert, 2020).

In the event that a policymaker agrees with Sierra-Arévalo and Papachristos' (2021) and these “rotten apples” are fired from an individual police department, it is not clear that policing as an industry is rid of the problematic officer. Robust decertification programs, where fired officers are no longer certified to be a police officer in a given state, are sparse (Goldman, 2013; Goldman, 2016) and many fired officers find work with other agencies, especially during difficult recruiting periods (Grunwald & Rappaport, 2020). Hiring an experienced officer is an immediate cost savings to an agency and provides the agency with the ability to put the person to work without the lag time of getting training and experience.

That said, hiring an officer who was fired, asked to leave, resigned in lieu of termination can create significant liability; a study of such “wandering officers” in Florida found that they were significantly more likely to receive complaints for violent or sexual conduct even when

controlled for age, gender, and education (Grunwald & Rappaport, 2020). Plaintiff's attorneys highly value cases involving an officer who repeated the bad behavior that caused them to leave a prior agency; indeed, such situations can open the door to municipal liability claims that are otherwise difficult to establish because they require evidence that the agency adopted an unconstitutional policy or custom. While the total cost of lawsuits against the police is unknown, news reports and academic assessments suggest they are quite significant (see Schwartz, 2014). A 2021 report by the Marshall Project shows that New York and Chicago have paid out more than two billion dollars in police misconduct claims over the past 10 years (<https://www.themarshallproject.org/2021/02/22/police-misconduct-costs-cities-millions-every-year-but-that-s-where-the-accountability-ends>). The Report concludes, "If not for inconsistent and shoddy record-keeping, we might know if settlements make a difference in police misconduct."

The potential for liability for on-going or repeat misconduct is complicated by the observation that it can be difficult to fire an officer because of the substantive and procedural protections provided by collective bargaining agreements (Rushin, 2017; Rushin, 2019) or state Law Enforcement Officers Bill of Rights, not all states decertify officers, and departmental discipline is often overturned in arbitration (Rushin, forthcoming) or by civil service boards (see Grunwald and Rappaport, 2020).

### *Civilian Review Boards*

One of the external reform mechanisms that is often touted as a necessary tool for police reform is the Civilian Review Board. The National Association for Civilian Oversight of Law Enforcement (<https://www.nacole.org>) provides an impressive array of civilian oversight models. It does not, however, offer any empirical evidence demonstrating the effectiveness of civilian

oversight. When the Council on Criminal Justice: Task Force on Policing published a Policy Assessment on Civilian Oversight in April 2021 ([https://assets.foleon.com/eu-west-2/uploads-7e3kk3/41697/civilian\\_oversight.2690411fd370.pdf](https://assets.foleon.com/eu-west-2/uploads-7e3kk3/41697/civilian_oversight.2690411fd370.pdf)), it reviewed the variety of models and the potential merits of civilian oversight and concluded (P. 1, 3):

Rigorous empirical research on the impact of civilian oversight and the relative merits of different models does not exist. Other research has yielded mixed findings about the ability of civilian oversight to reduce excessive use of force and other forms of police misconduct. ... Research to date has been largely descriptive of existing civilian oversight entities, either taking a deep look at the operations of a single entity or comparing and contrasting across the functions and structures of several entities.

Specifically, the report reviews the most influential studies and notes many methodological problems, including the lack of control over policy changes during the research period. The result is a characteristically mixed message: civilian oversight has promising aspects, it is helpful as a matter of public relations and transparency, but there is no empirical evidence that it works to reduce police misconduct or the use of force.

### *Consent Decrees*

Consent Decree, Settlement Agreements, and Collaborative Reforms offer another model of external reform. Each category of reform differs from the others in terms of degree and oversight requirements. While consent decrees have been around for a few decades, they and their data have been shrouded in secrecy until recently (Jiao, 2021 and Alpert et al., 2017). Unfortunately, the secrecy under which the consent decrees have operated has kept valuable evaluation data out of the hands of researchers or restricted the publication of studies; monitors' reports and agency data may be public, but the research behind them is unlikely to be made available. While there are a number of issues to consider, such as federal versus state consent decrees and lawsuits, the lack of the use of these data are lost opportunities.

Fortunately, there are exceptions. Jeff Fagan’s work in New York provides us with the best example of quality research and usable policy recommendations (see Fagan and Geller, 2020, Fagan, *forthcoming*). It is important to acknowledge the earlier work of Jim Fyfe who used data from litigation in which he was involved to develop important information on police use of force and deadly force (see Fyfe, 1988).

The research conducted in several agencies involved in consent decrees could assist reform efforts across the country. On the one hand, a comprehensive evaluation of the paragraphs of the Findings Letters and/or decrees would provide a real-time roadmap of where agencies are failing and where they need assistance. These data are available and could be used by other agencies to review their policies, training, supervision, and accountability systems. On the other hand, the data collected by the agencies to demonstrate their improvement and ability to fulfill the requirements of the consent decree or other reform agreements with the government could be analyzed to develop a roadmap or toolkit for other agencies to follow. The successes of one agency could be used by others. In other words, understanding the problems that got agencies under a consent decree and the ways in which they improved to fulfill the requirements—a process that requires input and oversight of agency command staff, external monitors, and a battery of government attorneys—would be of enormous assistance to other agencies. For the most part, though, these data are not easily available and have not been used to assist other agencies.

Moving to the impact or effectiveness of the consent decree, it is difficult to determine whether the evidence justifies the significant amount of time, money, and effort that they inevitable require. In some respect, consent decrees are inherently political; President Biden re-instated consent decrees after President Trump sharply curtailed their use, but both decisions

appear to have been made as a matter of political values rather than evidence of success or failure.

The relative merits of a consent decree are difficult to assess with any precision. There is some evidence that consent decrees and other forms of structural reform litigation has been effective in reducing misconduct (Rushin, 2015), but we do not know if consent decrees were an efficient mechanism for reform. Perhaps the agencies would have benefitted equally or even more from locally-driven reform supported by comparable funding. Several of the best descriptive studies of consent decrees are by Goh (2020), Rushin, (2014), Chanin (2015, 2017) and Harmon (2009). Chillar conducted an extensive and thorough review of the evaluation research conducted on consent decrees and published an important article that concluded (2021: 7):

While some consent decree evaluations find compelling evidence regarding the outcomes used to track compliance and the positive impacts police and citizens experienced, it is important to note that the studies were conducted during the period in which the department was under consent decree or was recently released from federal oversight. This ensured that the departments had no time to revert to the problematic policies and practices that warranted the implementation of the consent decree. Following the exit of the monitoring team in the cities of Pittsburgh, Washington D.C., Cincinnati, Detroit, and Prince George's County in Maryland, Chanin (2014, 2015) used focus groups and administrative data to investigate the effect of consent decrees. While results indicate that consent decrees offered departments the opportunity to bring about organizational change, such change faltered and reverted to business as usual upon the absence of DOJ oversight.

Chillar's observation reflects an issue raised in several of the articles; whether reforms being measured in the short-term term persist. His own research examined the short-term impacts on police stops of community members by the Newark, NJ Police Department. Overall, Chillar found that the consent decree reforms likely explained improvements in data quality collected in field interrogations, but that there were unexplained decreases in reported field interrogations of blacks and Latinos. Significantly, he concluded that regardless of specific reforms and oversight

efforts, cultural change in the agency was necessary to improve police behavior and actions toward black and Latino community members.

The evidence on consent decrees shows improved departments in the short term, but there is also evidence that in the absence of on-going oversight, agencies will revert to many of the problematic behaviors that resulted in the consent decree in the first place.

### *Qualified Immunity*

Perhaps no aspect of the legal regulation of policing has been subject to more contentious public debate than qualified immunity. Qualified immunity is a legal doctrine that protects officers who violate constitutional rights from civil liability unless it was “clearly established” at the time that their action was a violation. In other words, so long as a “reasonable officer” could have thought that a particular action did not violate the Constitution, an officer will not be liable for taking that action even if it *did* violate the Constitution. As applied by the courts, this has become a significant bar; as the leading qualified immunity scholar described, “[c]urrent Supreme Court doctrine suggests that an officer violates clearly established law only if there is a prior court of appeals or Supreme Court decision holding virtually identical facts to be unconstitutional” (Schwartz, 2021). The constitutional standards that regulate policing, however—such as Fourth Amendment standards of proof like “probable cause” and “reasonable suspicion” are themselves often indistinct (Stoughton et al., *forthcoming*). The result is a series of questions about both the substantive standard (i.e., whether an officer violated the constitution) and qualified immunity (i.e., whether the violation was clearly established) that are both framed in ways that often favor and defer to officers.

In part because of these ambiguities and the barriers to accountability that they create, there are arguments to maintain and to abolish qualified immunity. On the one hand, it is argued



that the doctrine helps weed out weak cases, develop acceptable practices, develop law, and protect the pocketbooks of municipalities. On the other hand, as Professor Joanne Schwartz has noted (2017: 1) “The United States Supreme Court appears to be on a mission to curb civil rights lawsuits against law enforcement officers, and appears to believe qualified immunity is the means of achieving its goal. The Supreme Court has long described qualified immunity doctrine as robust—protecting ‘all but the plainly incompetent or those who knowingly violate the law.’” (citing *Mallory v Briggs*, 475 U.S. 335, 341 (1986)). Her article and *Kinports* (2016) cover the arguments and opinions where the use of qualified immunity has been used to create a shield for the police and has protected them from abuses and maltreatment of citizens. Schwartz presents an analysis of the role of qualified immunity in constitutional litigation that leads her to conclude that qualified immunity has been seen as unclear and unreasonably protective of government actors.

The evidence on qualified immunity suggests that the law, as it now stands, not only shields officers from liability for abuses, it also shields their employers. Private employers are liable for the actions for their employees under a legal theory known as “vicarious liability” or “*respondeat superior*,” but public employers are not. Nevertheless, police agencies virtually always cover the costs of defending their officers in litigation as well as the settlement or judgment costs (Schwartz, 2014). When individual officers are insulated from liability, it is almost always their employers, rather than the officers themselves, that benefit.

What is far less clear is whether eliminating qualified immunity—or revising municipal liability doctrines or tort reform, for that matter—would have any significant effect on officer action. Schwartz has predicted that there would be more civil rights litigation but that the average time, cost, and complexity would decrease and that the success rates for plaintiffs and

defendants would remain roughly stable (Schwartz, 2020). While certainly well informed, Schwartz's predictions are based on doctrinal analysis and interviews with attorneys rather than empirical modeling.

## **Discussion**

Generating an evidence-base for dealing with the use of force in policing is exceedingly difficult. Reducing unnecessary and excessive force is an objective of vital importance to the public. National attention to this issue has generated the immediate need for departments to "do something." Unfortunately, from a scientific perspective, doing something quickly is often not supported by or conducive to generating solid research findings. Just as we did not immediately roll out untested vaccines nationwide, we must not rush to roll out untested programs to reduce the use of force nationwide just because there is an urgent need for improvement. Indeed, the importance of the issue is precisely why we must take a deliberative approach; getting it right slowly comes with real costs, but those costs are far less than continuing to get it wrong. This is not to suggest that no one take action, but to make sure those actions are evaluated to determine their efficacy. A society that considers this issue as urgent as the coronavirus pandemic should make funding more readily available and demand a commitment to not only relying on theoretically sound and evidence-based reforms, but also taking part in research from their policing agencies.

Our review of the literature suggests that quality studies of policies, training, selection, supervision, and accountability exist in small numbers, with few of the many policies and programs implemented by police agencies having been evaluated properly. As noted multiple times, the current state of inquiry on police use of force can best be summarized as a lot of good ideas with very few well-supported by research. More detailed policies, different equipment,

better communication skills, the extended use of mental health professionals, improved supervision, changes to accountability strategies, and other reforms could all work—or, more likely, could all marginally advance the ultimate goal as components of comprehensive reform—but it will take time and diligent research to determine how to best make them work and whether each particular effort is individually or collectively worth the time, effort and cost.

Perhaps as important, replications of existing studies are non-existent. Studies of policy changes, training, selection or supervision programs, or accountability measures were frequently limited to single departments and never contained more than a few. Creating a robust evidence base *requires* replication. Replications are uncommon for a number of reasons. In an analysis of articles in 100 psychological journals with the highest impact scores, Makel et al. (2012) report that only 1% of published studies are replications. It appears that editors, who place a higher prize on originality and positive results, are reluctant to publish replications. In fact, a study of journal editors by Madden, Easley and Dunn (1995, p. 85) concluded: “A paradox of replication in the social sciences is that a researcher who operationally replicates and finds nonsupport for previous work may be accused of not being true to the original method, but if the researcher finds support for the previous work, then the argument will be made that nothing new has been learned.” As Smart (1964, p. 232) noted, “withholding negative results from publication has a repressive effect on scientific development.” Tautologically, the difficulty of getting replications published reduces the desirability of conducting them, although it is worth noting that the more controversial the topic the more likely a replication study will be published (Madden, Easley and Dunn (1995).

The second major reason for the lack of replication studies is the difficulty of obtaining funding for major applied research projects in policing. As noted in the introduction, the research

and evaluation budget for the Office of Justice Programs—the major national agency that is responsible for funding these types of evaluations—is relatively small, especially considering that it covers not only policing, but other areas of criminal justice (e.g., courts, corrections, and forensic sciences). The annual Research and Evaluation on Policing solicitation funds evaluations of policing programs, but usually less than 5 per year (it funded 4 in 2020). In a competitive environment where only a select number of evaluations of policing programs are funded every year, it is nearly impossible to obtain funding to replicate an evaluation of a program in a different context. Even CrimeSolutions.gov, a program sponsored by the National Institute of Justice (NIJ) that rates research to determine certainty of the outcomes and effectiveness, does not use replications as an evaluation criterion. The National Institute for Justice (or another federal agency) should create a solicitation dedicated to replications of studies that show promise by prior evidence. Such a model exists in the private sector (Arnold Ventures).<sup>2</sup>

It is well known that replications are critical to building and accumulating knowledge and that the bias against publishing and funding them has an impact on conducting them. Programs and policies with promising results should continue to be studied in multiple departments and the results replicated to ensure that the effects are generalizable. Programs and policies without promising results may need to be discarded but may also be worth salvaging and modifying if researchers and practitioners can identify the areas where programs went wrong or if the measured outcomes were inappropriate.

In the end, the strongest recommendation we can provide is that more resources—in the form of time, money, effort, and commitment from police researchers and agencies alike—must

---

<sup>2</sup> <https://craftmediabucket.s3.amazonaws.com/uploads/RFP-for-RCTs-in-Criminal-Justice-CJ-EBP.pdf>

be dedicated to study police behavior and uses of force. The implementation of new programs and policies *must* be accompanied by evaluations and resources should be expended to replicate those evaluations and ensure the accuracy of the findings. In this area, the United States, in particular, lags considerably behind other countries, yet it never seems to learn its lesson. Each new use of force controversy is accompanied by a plethora of new ideas and little support for evaluating them.

### References

- Alpert, G.P. & Dunham, R.G. (2004). *Understanding police use of force: Officers, suspects, and reciprocity*. New York: Cambridge University Press.
- Alpert, G., McLean, K., & Wolfe, S. (2017). Consent decrees: An approach to police accountability and reform. *Police Quarterly*, 20, 239-249.
- Aronie, J, and Alpert, G. (2020). 16 to a dealer's 10: Could blackjack odds help inform police pursuit policies? *PoliceOne*. March 2
- Ba, B.A., Knox, D., Mummolo, J., & Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in Chicago. *Science*, 371, 696-702.
- Barone, E. (2020). The Trump Administration's 'Operation Warp Speed' has spent \$12.4 billion on vaccines. How much is that, really? *Time*. Retrieved from: <https://time.com/5921360/operation-warp-speed-vaccine-spending/>.
- Carter, D.L., & Sapp, A.D. (1991). *Police education and minority recruitment: The impact of a college requirement*. Police Executive Research Forum.
- Chalfin, A., & Kaplan, J. (2021). How many complaints against police officers can be abated by incapacitating a few "bad apples?" *Criminology & Public Policy*, 20, 351-370.
- Chanin, J. (2014). On the implementation of pattern or practice police reform. *Criminology, Criminal Justice, Law & Society*, 15, 38-56.
- Chanin, J. (2015). Examining the sustainability of pattern or practice police misconduct reform. *Police Quarterly*, 18, 163-192.

- Chanin, J. (2017). Early stages of pattern or practice police misconduct reform: An examination of the Department of Justice's investigation and negotiation processes. *Police Quarterly*, 20, 250-274.
- Chapman, C. (2012). Use of force in minority communities is related to police education, age, experience, and ethnicity. *Police Practice and Research*, 13, 421-436.
- Chillar, V.F. (2021). The racial divide at micro places: A pre/post analysis of the effects of the Newark Consent Decree on field inquiries (2015-2017). *Journal of Research in Crime and Delinquency*. DOI: 10.1177/00224278211030964.
- Cohen, B. & Chaiken, J.M. (1972). *Police background characteristics and performance*. Rand Institute.
- Engel, R.S. (2000). The effects of supervisory styles on patrol officer behavior. *Police Quarterly*, 3, 262-293.
- Engel, R.S. (2001). Supervisory styles of patrol sergeants and lieutenants. *Journal of Criminal Justice*, 29, 341-355.
- Engel., R.S. (2002). Patrol officer supervision in the community policing era. *Journal of Criminal Justice*, 30, 51-64.
- Engel., R.S., Corsaro, N., Isaza, G.T., & McManus, H.D. (2020). *Examining the impact of Integrating Communications, Assessment, and Tactics (ICAT) de-escalation training for the Louisville Metro Police Department: Initial findings*. IACP/UC Center for Police Research and Policy. Retrieved from: [https://www.theiacp.org/sites/default/files/Research%20Center/LMPD\\_ICAT%20Evaluation%20Initial%20Findings%20Report\\_FINAL%2009212020.pdf](https://www.theiacp.org/sites/default/files/Research%20Center/LMPD_ICAT%20Evaluation%20Initial%20Findings%20Report_FINAL%2009212020.pdf).
- Engel, R.S., & Worden, R.E. (2006). Police officers' attitudes, behavior, and supervisory influences: An analysis of problem solving. *Criminology*, 41, 131-166.
- Ericsson, K.A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, 79, 570-581.
- Ericsson, K.A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science*, 16, 346-350.
- Fagan, J. (Forthcoming). No runs, few hits and many errors: Street stops, bias and proactive policing. *UCLA Law Review*.
- Fagan, J., & Geller, A. (2020). *Profiling and consent: Stops, searches, and seizures after Soto*. *Virginia Journal of Social Policy & the Law*, 27, 16-58.

- Fleming, J. & Rhodes, R. (2018). Can experience be evidence? Craft knowledge and evidence-based policing. *Policy and Politics*, 46, 3-26.
- Friedrich, R.J. (1977). *The impact of organizational, individual, and situational factors on police behavior*. University of Michigan [Thesis].
- Fyrer, R.G. (Forthcoming). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*.
- Fyfe, J. (1979). Administrative interventions on police shooting discretion: An empirical examination. *Journal of Criminal Justice*, 7, 309-323.
- Fyfe, J. (1988). Police use of deadly force: Research and reform. *Justice Quarterly*, 9, 165-205.
- Garner, J.H., Maxwell, C.D., & Heraux, C.G. (2002). Characteristics associated with the prevalence and severity of force used by the police. *Justice Quarterly*, 19, 706-746.
- Giacomantonio, C., Goodwin, S., & Carmichael, G. (2018). Learning to de-escalate: Evaluating the behavioural impact of Verbal Judo training on police constables. *Police Practice and Research*, 21, 401-417.
- Goldman, R.L. (2016). The importance of state law in police reform. *Saint Louis University Law Journal*, 60, 363-390.
- Goldman, R.L. (2013). A model decertification law. *Saint Louis University Public Law Review*, 32, 147-156.
- Goh, L.S. (2020). Going local: Do consent decrees and other forms of federal intervention in municipal police departments reduce police killings? *Justice Quarterly*, 37, 900-929.
- Grunwald, B. & Rappaport, J. (2020). The wandering officer. *Yale Law Journal*, 129, 1676-1782.
- Gullion, C. & King, W. (2020). Early intervention systems for police: a state-of-the-art review. *Policing: An International Journal*, 43, 643-658.
- Harmon, R. (2009). Promoting civil rights through proactive policing reform. *Stanford Law Review*, 62, 1-68.
- Huddleston, D. (2021). Marietta police officers learning Brazilian Jiu Jitsu to help keep the streets safe. *WSB-TV 2*. Retrieved from: <https://www.wsbtv.com/news/local/cobb-county/marietta-police-officers-learning-brazilian-jiu-jitsu-help-keep-streets-safe/CCYMTFR47FFD7B4QU5WOUQGC6Q/>.

- Jiao, A. (2021). Federal consent decrees: A review of policies, processes, and outcomes. *Police Practice and Research*, 22, 793-804.
- Kaminski, R.J., DiGiovanni, C., & Downs, R. (2004). The use of force between the police and persons with impaired judgment. *Police Quarterly*, 7, 311-338.
- Kane, R., & White, M. (2009). *Bad cops, police misconduct and the New York City Police Department*. New York: New York University Press.
- Kinports, K. (2016). The Supreme Court's quiet expansion of qualified immunity. *Minnesota Law Review*, 100, 62-78.
- Legewie, J. (2016). Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology*, 122, 379-424.
- Lough, J. and Ryan, M. (2010). Psychological profiling of Australian police officers: a three-year examination of post-selection performance. *International Journal of Police Science & Management*, 12, 480-486.
- Madden, C., Easley, R., Dunn, M. (1995). How journal editors view replication research. *Journal of Advertising*, 24, 77-87.
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7, 537-542.
- McCaffree, K. & Saide, A. (2021). *How informed are Americans about race and policing?* Skeptic Research Center, CUPES-007.
- McCluskey, J.D., & Terrill, W. (2005). Departmental and citizen complaints as predictors of police coercion. *Policing: An International Journal*, 28, 513-529.
- McLean, K., Wolfe, S.E., Rojek, J., Alpert, G.P., & Smith, M.R. (2020). Randomized controlled trial of social interaction police training. *Criminology & Public Policy*, 19, 805-832.
- Mullinix, K.J., Bolsen, T., & Norris, R.J. (2021). The feedback effects of controversial police use of force. *Political Behavior*, 43, 881-898.
- NJ Advance Media (2018). The force report. *NJ.com*. Retrieved from <https://force.nj.com/>.
- Noble, J. & Alpert, G. (2009). *Managing accountability systems for police conduct: Internal affairs and external oversight*. Prospect Heights, IL: Waveland Press.
- Office of Justice Programs. (2021). *FY 2021 performance budget*. US DOJ. Retrieved from: <https://www.justice.gov/doj/page/file/1246736/download>.



- Owens, E., Weisburd, D., Amendola, K.L., & Alpert, G.P. (2018). Can you build a better cop? Experimental evidence on supervision, training, and policing in the community. *Criminology & Public Policy*, 17, 41-87.
- Owens, E., Weisburd, D., Amendola, K.L., & Alpert, G.P. (2017). Research in Brief: Passing it Forward: Modeling Procedural Justice in the Workplace to Improve Public-Police Interactions. *Police Chief* 79: 14-15.
- Paoline, E.A. & Terrill, W. (2007). Police education, experience, and the use of force. *Criminal Justice and Behavior*, 34, 179-196.
- Peeples, L. (2020). What the data say about police brutality and racial bias – and which reforms might work. *Nature*, 583, 22-24.
- Resnick, B. (2020). Police brutality is a public health crisis. *Vox*. Retrieved from: <https://www.vox.com/science-and-health/2020/6/1/21276828/pandemic-protests-police-public-health-black-lives-matter>.
- Rushin, S. (2015). Structural reform litigation in American police departments. *Minnesota Law Review*, 99, 1343-1422.
- Rushin, S. (2017). Police union contracts. *Duke Law Journal*, 66, 1191-1266.
- Rushin, S. (2019). Police disciplinary appeals. *University of Pennsylvania Law Review*, 167, 545-610.
- Rushin, S. (Forthcoming). Police arbitration. *Vanderbilt Law Review*.
- Rydberg, J., & Terrill, W. (2010). The effect of higher education on police behavior. *Police Quarterly*, 13, 92-120.
- Schwartz, J. (2014). Police indemnification. *New York University Law Review*, 89, 885-1006.
- Schwartz, J. (2017). How qualified immunity fails. *The Yale Law Journal*, 127, 2-76.
- Schwartz, J. (2020). After qualified immunity. *Columbia Law Review*, 120, 309-388.
- Schwartz, J. (2021). Qualified immunity's boldest lie. *The University of Chicago Law Review*, 88, 605-684.
- Sierra-Arévalo, M., & Papachristos, A. (2021). Bad apples and incredible certitude. *Criminology & Public Policy*, 20, 371-381.
- Sherman, L.W., & Blumberg, M. (1981). Higher education and police use of deadly force. *Journal of Criminal Justice*, 9, 317-331.

- Shjarback, J.A., White, M.D., & Bishopp, S.A. (2021). Can police shootings be reduced by requiring officers to document when they point firearms at citizens? *Injury Prevention*. DOI: 10.1136/injuryprev-2020-043932.
- Skogan, W.G., Van Craen, M., & Hennessy, C. (2015). Training police for procedural justice. *Journal of Experimental Criminology*, 11, 319-334.
- Smart, R. G. (1964). The importance of negative results in psychological research. *Canadian Psychologist*, 5, 225–232.
- Stoughton, S.W. (2021). How the Fourth Amendment frustrates the regulation of police violence. *Emory Law Journal*, 70, 521-585.
- Stoughton, S.W. (2014). The incidental regulation of policing. *Minnesota Law Review*, 98, 2179-2235.
- Stoughton, S.W., Noble, J.J., & Alpert, G.P. (2020). *Evaluating police uses of force*. New York: NYU Press.
- Stoughton, S., McLean, K., Nix, J., & Alpert, G. (Forthcoming). Policing suspicion: Qualified immunity and “clearly established” standards of proof. *Journal of Criminal Law & Criminology*.
- Sun, I.Y., & Payne, B.K. (2004). Racial differences in resolving conflicts: A comparison between black and white police officers. *Crime & Delinquency*, 50, 516-541.
- Terrill, W. (2005). Police use of force: A transactional approach. *Justice Quarterly*, 22, 107-138.
- Terrill, W. & Paoline, E.A. (2013). Examining less lethal force policy and the force continuum: Results from a national use-of-force study. *Police Quarterly*, 16, 38-65.
- Terrill, W. & Paoline, E.A. (2017). Police use of less lethal force: Does administrative policy matter? *Justice Quarterly*, 34, 193-216.
- Tyler, T.R. (2006). *Why people obey the law*. Princeton, NJ: Princeton University Press.
- Tyler, T.R. & Huo, Y.J. (2002). *Trust in the law: Encouraging public cooperation with the police and courts*. New York: Russell Sage.
- Walker, S. (2012). Institutionalizing police accountability reforms: The problem of making police reforms endure. *Saint Louis University Public Law Review*, 32, 57-92.
- Walker, S., Alpert, G.P., & Kenney, D.J. (2001). *Early warning systems: Responding to the problem police officer*. National Institute of Justice: Research in Brief.

- Wasilczuk, M.K. (*Forthcoming*). How police hiring policies fail to account for emerging adulthood. *Buffalo Law Review*.
- White, M.D. (2001). Controlling police decisions to use deadly force: Reexamining the importance of administrative policy. *Crime & Delinquency*, 47, 131-151.
- White, M.D., Mora, V.J., Orosco, C., & Hedberg, E.C. (2021). Moving the needle: Can training alter officer perceptions and use of de-escalation? *Policing: An International Journal*, 44, 418-436.
- Wood, G., Tyler, T.R., & Papachristos, A.V. (2020). Procedural justice training reduces police use of force and complaints against officers. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 9815-9821.
- Worden, R. (1995). The causes of police brutality: Theory and evidence on police use of force. In W. Geller & H. Toch (Eds.), *And justice for all: Understanding and controlling police abuse of force* (pp. 31-60). Washington, DC: Police Executive Research Forum.
- Worden, R., McLean, S.J., Engel, R.S., Cochran, H., Corsaro, N., Reynolds, D., Najdowski, C.J., & Isaza, G.T. (2020). *The impacts of implicit bias awareness training in the NYPD*. IACP/UC Center for Police Research and Policy & the John F. Finn Institute for Public Safety. Retrieved from:  
[https://www1.nyc.gov/assets/nypd/downloads/pdf/analysis\\_and\\_planning/impacts-of-implicit-bias-awareness-training-in-%20the-nypd.pdf](https://www1.nyc.gov/assets/nypd/downloads/pdf/analysis_and_planning/impacts-of-implicit-bias-awareness-training-in-%20the-nypd.pdf).