

Exploring a Craft Learning Model for Reviewing Patrol Officer Decision-Making in Encounters with the Public

James J. Willis and Heather Toronjo

Since the 1960s, police departments have turned to rules and procedures to help control how patrol officers, as legal decision-makers, exercise their discretionary authority. The logic of the administrative rulemaking model depends on the development and enforcement of bureaucratic rules and regulations. The public outcry over high-profile incidents of police abuses of authority has renewed interest in this approach. This article conceptualizes a complementary craft learning model to supplement rulemaking. This model harnesses patrol officers' knowledge and skills, learned through experience, to the development of criteria for assessing and guiding how they use their discretion in less dramatic encounters with the public. Using in-depth interviews with thirty-eight patrol officers reacting to a video clip of a fairly routine and low-key neighbor dispute, we derive seven evaluative standards (accountability, lawfulness, problem diagnosis, repair of harm, economy, fairness, and safety and order). We then explore how these standards could be used by first-line supervisors to structure reviews of patrol officer decision-making through a process of reflection-in-action. Our purpose is to imagine a reform strategy that tries to account for the complex technical and normative dimensions of everyday police work to facilitate more deliberate, transparent, and principled decisions.

In a modern state the actual ruler is necessarily and unavoidably the bureaucracy, since power is exercised neither through parliamentary speeches nor monarchical enunciations but through the routines of administration.

Max Weber

It is quite clear that criteria of workmanship and procedures for examining it must be developed from within policing. The purely technical aspects of workmanship in policing can only acquire formulation and development from the practitioners of the craft themselves.

Egon Bittner

James J. Willis is Professor of Criminology, Law and Society at George Mason University, Email: jwillis4@gmu.edu.

Heather Toronjo is Research Faculty in the Schar School of Policy and Government at George Mason University, Email: htoronjo@gmu.edu.

The authors thank the anonymous reviewers for their very thoughtful comments, and Drs. Richard Lippke, Allison Redlich, and David Weisburd for their helpful feedback on earlier drafts. The lead author would also like to acknowledge his intellectual debt to Dr. Stephen Mastrofski, whose ideas and insights over many years have contributed greatly to his understanding of the police craft.

INTRODUCTION

Due to the breadth of the police mandate as “a mechanism for the distribution of non-negotiably coercive force” (Bittner 1970, 46) and to the ambiguity of statutes defining the limits to police powers, police officers have enormous discretionary authority (Davis 1969; 1975; Walker 1993; Thacher 2016). In choosing how to exercise this authority on the front lines, patrol officers rely heavily on the experiences they have accumulated “handling a multitude of what seem like unique situations over and over again” (Bayley and Bittner 1984, 35). They consider the craft knowledge and skills they have learned in response to the context-dependent “exigencies” of different types of encounters with the public as essential to making good judgments (Bittner 1970, 46). The questionable bases of these judgments have been revealed over the past years, as Americans have taken to the streets to protest problems of racial discrimination, arbitrary enforcement, and excessive force. In doing so they have highlighted one of the most significant challenges facing American policing today: the need for more effective strategies for channeling discretion that account for the knowledge and skills that officers value and for the realities of police practice (Bittner 1970; Thacher 2008; Worden and Dole 2019).

To many it appears that police officers are “out of control,” neither protecting nor serving, but using their authority in ways that are pernicious, arbitrary, and intrusive (Friedman 2017; Bazelon 2020). High-profile cases and dramatic outcomes capture national attention, but most police-civilian encounters do not involve fighting crime, making arrests, or using force (Bittner 1990, 240–43; Bayley 1994). Instead, they comprise numerous low-key but challenging “human troubles” (Black 1971, 1090), such as arbitrating or mediating between disputants, managing recalcitrant youths, or responding to public disorder complaints. Due to a lack of administrative guidance (Tift 1975), patrol officers are often left to rely on their own individual experiences and department norms to decide how best to intervene (Brown 1981). And yet research suggests that people care very much about the quality of a police officer’s choices (Mastrofski 1996; Tyler 2004), and they are quite comfortable observing and judging how rightfully or reasonably they feel treated (Meares et al. 2015). These peacekeeping activities rarely receive “equal billing” to reform movements focused on much larger and more ambitious goals for transforming, or even abolishing, police organizations (Mastrofski 1999, 1; Searcey 2020). Thus, a core question for advocates of police reform, one rendered more urgent by recent events, is what might be done to improve the quality and defensibility of the choices patrol officers make when exercising their authority in everyday encounters with the public.

Efforts to control discretion take many guises, including implementing new forms of citizen oversight, changing recruitment and training, and tightening disciplinary mechanisms (Mastrofski 2004; President’s Taskforce 2015; Walker 2016). However, just as scholars in the 1960s and 1970s advocated for administrative rulemaking by local police departments (Goldstein 1967; Davis 1969, 1975), there is a resurgence of interest in this approach (Friedman 2017). Supporters suggest that departmental guidelines can channel decision-making in ways that are practicable and help meet the law’s purpose. From this perspective, police agencies should develop policies that help define, clarify, and restrict police discretion, and implement rigorous accountability mechanisms for

their enforcement. Well-known police and legal scholars, including Samuel Walker (Walker and Archbold 2014), Christopher Slobogin (2016), and Barry Friedman (2017) espouse the virtues of the administrative rulemaking model. More stringent guidelines and accountability structures were also key features of the George Floyd Justice in Policing Act of 2021 reform bill.¹ In light of this, one scholar of police governance recently concluded, “we are in the midst of a rule-making renaissance” (Ponomarenko 2019: 4).

Administrative rulemaking is an important means of regulation and compliance, but it also has limitations (Krantz 1979; Walker 1993; Ponomarenko 2019). Patrol officers make decisions in varied situations that are unique and uncertain. In organizational terms, police work occurs in a task environment that is heterogeneous, and where the “technology” for getting things done is poorly developed (Mastrofski and Ritti 2000). Put differently, the idiosyncrasies of street-level encounters are not well understood, nor is the technical capacity of the police for producing desirable outcomes (Engel and Worden 2003).

These features present significant challenges to an administrative rulemaking model that uses a top-down approach to controlling discretion through the enforcement of general bureaucratic rules (Bittner 1983). It is difficult developing regulations sufficiently plentiful for the extraordinarily diverse situations that officers are called upon to handle, and that are capable of covering the many factors relevant to an officer’s decision (Bittner 1990; Mastrofski 2000). Policies that try to create rules around all this complexity can become inflexible and cumbersome and, even when detailed, “may remain obtuse in the face of unpredictable circumstances that continue to rise” (Thacher 2020, 756). This helps explain why administrative reviews are evoked selectively, often as a formal disciplinary process focused on rule compliance in relation to relatively rare critical incidents, particularly deadly force (and more recently the use of nonlethal force), domestic violence, and high-speed pursuits (Walker 1993).

Moreover, the rulemaking model tends not to afford much importance to the valuable “resources of knowledge, skill, and judgment” that experience teaches patrol officers constitute the craft of police work (Goldstein 1967, 1123; Bittner 1983; Sklansky and Marks 2008; Willis 2013). And yet research suggests that based on their hands-on experience, some patrol officers have developed a level of artistry in their work that is hard to ignore (Willis and Mastrofski 2017). Muir’s (1977) professional police officer is capable of handling complex situations with uncanny skill, and Bittner (1967) describes a set of sophisticated techniques that allow patrol officers to make prudent judgments when keeping the peace on skid row. At its best, the domain of an officer’s professional competence includes, “making good arrests, deescalating crises, investigating crimes, using coercion and language effectively, abiding by the law and protecting individual rights, developing knowledge of the community, and imparting a sense of fairness to one’s actions” (Mastrofski 1988, 65). This is not to romanticize police expertise as a “privileged, divine-like attribute” (Lvovsky 2021, 4), but to simply recognize the potential benefits of systematically accessing the kind of craft knowledge that is unavailable to outsiders, who lack the same level of mastery or experience (Moore 1995; Thacher 2008). According to Egon Bittner (1990: 262), one of the most meticulous observers

1. For details on the bill, see <https://judiciary.house.gov/issues/issue/?IssueID=14924>.

of the police, we should look to the individual officers of a highly complex vocation “to make police work what it should be.”

Despite the potential of using the police craft as a means to channel police discretion (Bayley and Bittner 1984; Mastrofski 1996), this perspective remains underexplored. Kenneth Davis, one of rulemaking’s strongest advocates, explicitly *excluded* the knowledge of patrol officers when outlining the major elements of regulatory supervision. According to Davis (1975, 11), those aloft the police organization should be solely responsible for controlling discretion by writing detailed rules and procedures: as “top officers” they “obviously [had] the skills and broad understanding that patrolmen typically lack.”

Last, bureaucratic rules generally establish minimum requirements for performance (Bittner 1983; Ericson 2007), fostering judgments of adequacy or “failures to do anything wrong.” In this respect, they function differently to operative standards common to a wide range of other professional practices, such as teaching, nursing, and social work. The latter raise performance expectations by trying to articulate fine-grained distinctions between “good and bad work practices” (Bittner 1990, 146). At its best, craft or “workmanship” can establish expectations for skilled performance that far surpass those of general laws or policies (Bittner 1983; Klein 2011).

Some evidence suggests that newer versions of the administrative rulemaking model are opening up opportunities for police officers to share their street-level experiences during the review process to promote learning and improve decision-making (Thacher 2020). Still, attempts to combine the dual goals of experiential learning and rule-based compliance raise questions about their compatibility. Police agencies are punishment-centered bureaucracies (Gouldner 1954), whose existing review processes for strengthening internal accountability are better equipped at identifying and sanctioning misconduct than promoting skilled performance (Mastrofski 2000). As a result, efforts to encourage the kind of low-risk and open-ended review process for facilitating learning about discretionary choices are easily undermined by features of more traditional compliance-monitoring systems (Schön 1983). For example, strict accountability can discourage efforts to search for and experiment with innovative, and potentially more effective, alternatives to standard police responses (Schön 1987, 4). Compstat’s crime control elements were designed to strengthen both creative problem-solving and internal accountability, but middle managers’ fear of adverse consequences for failing to reduce crime quickly led them to eschew innovative solutions in favor of standard short-term crime responses (i.e., increasing arrests or saturation patrols) (Willis, Mastrofski, and Weisburd 2007).

THE LOGIC OF A CRAFT LEARNING MODEL

Developing Craft-Based Standards

Given these concerns, we start to develop an alternative model for helping broaden reform efforts to combat questionable and troubling uses of police discretion. We conceptualize a craft learning model for harnessing practical knowledge generated through officers’ hard-won experiences to specific performance standards. Despite

numerous reform efforts over the last fifty years, there has been “little effort to capitalize upon police experience” to influence officers’ discretion (Goldstein 1967, 1132). Nor is this part of contemporary debates about police reform. Police agencies continue to rely heavily on traditional measures of patrol officer performance (e.g., tallies of arrests, incidents, and citations), and to do little to learn about and document how police exercise their authority in daily police-public encounters. Even when there is an official record, “there is only the vaguest indication of police action (“advised,” “referred,” “warned”)” (Mastrofski 1996, 216). Most encounters are resolved by informal means, but little is known about what means were used, and to what extent they were improper or unwise. In the next section, we outline how performance standards could be used by police supervisors, as part of an in-depth and continuous review process for scrutinizing and helping patrol officers reflect on the judgments that skilled policing demands.

Standards for evaluating practices, although they might only be tacitly understood, are central to being a member of a profession. As Yanow and Tsoukas write (2009, 1344, emphasis in original):

To be a member of a practice is to be someone for whom what is going on in the practice *matters*. Practices are constituted by certain collective self-understanding that situate practitioners relative to particular standards of excellence and to obligations, held both collectively within the practitioner community which individuals aspire to join or to which they belong. These self-understandings cannot be qualitatively neutral: they are articulated through contrasts (e.g. of right and wrong uses of concepts) and, hence, entail an evaluative component. Both self-understandings and evaluative components are learned through engaging in and with the practice, not through thinking about them.

Unlike the more constraining properties of “if-then provisions in rules and procedures” (Engel and Worden 2003, 132), craft-based standards allow for greater flexibility in helping reviewers make judgments about the goals that practitioners should *strive* to meet when making decisions. Professionals set standards to evaluate their own performance in specific contexts because they want their “behavior to measure up to the standards as much as possible” (Cohen and Feldberg 1991, 40). Similarly, standards can help distinguish high quality police work from work that is merely acceptable or unacceptable. Thus, there is an aspirational quality to craft-based standards grounded in the realities of actual practice that is not easily captured by abstract and general rules. In the same way that craftspeople are distinguished by their virtuoso skills, creativity, and desire to “do a job well for its own sake,” craft standards can help distinguish superior performers from those who merely meet their minimal obligations (Sennett 2008, 9).

A current barrier to the development of craft-based standards in policing is the deeply entrenched view among police themselves that it is impossible to articulate performance criteria *ex ante* for a given situation (Bayley and Bittner 1984). From this perspective, police work is simply too unpredictable, varied, and complex for the creation of decision-making frameworks that offer useful guidance for the front-lines. The barrier this presents is rendered more challenging by the assumption that police work depends

on an “inchoate lore” rather than a body of “technical knowledge or elaborate schemes of norms” for offering guidance (Bittner 1990, 262).

Because any evaluation of police discretion (and a potential framework for reviewing and scrutinizing its use) first requires the possibility of identifying defensible criteria for its assessment (Mastrofski 2004), we test this view directly. Of course, given discretion’s broad domain, we would expect officers to think differently from one another about what comprises skilled police work in any specific situation. At the same time, we wondered if there was a sufficient level of consensus among patrol officers on what constitutes skilled performance to act as a useful tool for learning, and for channeling discretion along more “principled routes” (Thacher 2016, 533). With this in mind, we conducted in-depth interviews with thirty-eight patrol officers in one agency reacting to the performance of two officers in a video-taped incident involving a dispute between two neighbors. If there were patterns of agreement around the strengths and weaknesses of how the officers depicted chose to respond to the dispute, and around suggestions for improvement, these insights could be a basis for envisioning a craft learning model.

Applying Craft Standards to a Supervisory Review Process Based on Reflection-in-Action

While our main focus is to explore empirically the feasibility of deriving craft-based performance standards, we go one step further by envisioning a model for how these standards might be integrated into a more systematic learning process for improving decision-making. We focus on a neighbor dispute, but similar to Herman Goldstein’s problem-oriented policing model (1979), this approach could be applied to different types of encounters that share common features (e.g., traffic stops, truant teenagers) (Mastrofski 1996, 229). Local agencies would be responsible for prioritizing those encounters that they believe warrant most attention and developing and codifying different performance criteria for them. We cannot address all the concerns our model raises, and so we limit ourselves to conceptualizing a review process centered on “reflection-in-action,” and to describing some of its potential benefits. Such a model, designed around meaningful reflection on one’s actions, is consistent with an emerging interest in making policing a more critically reflective practice and in taking advantage of body-worn camera (BWC) technology for this purpose (Charles 2000; Christopher 2015; Phelps et al. 2018). In addition to building on the small body of existing police scholarship to conceptualize how reflective practice might be a key feature of a craft learning model, we use our interview data to illustrate its specific practical implications for learning about and channeling discretion.

The rapid spread of BWCs across U.S. police agencies over the last few years means that, unlike police work in the past, many police supervisors now have access to high quality audio and video recordings for helping appraise how officers behave during their encounters with the public (Nix et al. 2020). This footage can capture a tremendous amount of detail, including variations in body language and tone of voice, and it can be played slowly and stopped to help officers recall what they were thinking in the actual moment and context of deciding. Much attention has been paid to how BWCs can strengthen compliance with bureaucratic rules and help deter use-of-force and police

misconduct (Lum et al. 2020). In comparison, there has been much less exploration of the potential of this technology as a more constructive learning tool for capturing and improving craft knowledge and skills (Willis and Mastroski 2017). In fact, current research suggests that body-worn cameras are being harnessed much more tightly to police agencies' existing accountability structures for investigating complaints and identifying and disciplining officers for wrongdoing. In comparison, it appears they are not being used as a more innovative tool for helping officers deliberate on, and learn from, more routine exercises of their discretion when they engage with the public (Koen, Willis, and Mastroski 2018; Koen and Mathna 2019; White and Malm 2020).

First articulated by John Dewey (1933) and later developed by Donald Schön (1983), reflection-in-action is part of an epistemology of practice whose point of departure is "the competence and artistry already embedded in skillful practice" (Schön 1987, xi). Key to this artistry is the ability of some practitioners, especially experts, to "reflect in the midst of action without interrupting it," particularly when dealing with situations characterized by "uncertainty, disorder, and indeterminacy" (Dewey 1933; Schön 1983, 16; 1987, 26). Since routine responses are less likely to apply to these problematic situations, reflection-in-action may lead to a practitioner changing the initial framing of a problem, reconsidering underlying assumptions, rethinking a particular strategy, and conducting "on-the-spot" experiments, in order to adapt to changing circumstances, and to pursue an approach likely to be more successful (Schön 1983, 62–63). Given this background, using video footage to enable practitioners to think actively and carefully on how they reflected-in-action during an encounter with the public, particularly around how they adjusted in those moments which were unexpected or particularly challenging, could play an important role in the development of practical skills and new knowledge, which might then inform future actions (Schön 1987, 31).

Because first-line supervisors work closely with individual officers and instruct them on how to perform their craft (Muir 1977; Engel and Worden 2003), they are well situated to periodically review BWC footage of what initially might have appeared to seem a fairly routine encounter, but that an officer then experienced as problematic. It is common for BWC policies to grant supervisors permission to access BWC footage to conduct individual performance appraisals, but studies suggest this is rarely done (White, Flippin, and Malm 2019; Willis 2022). Since video footage can provide a close approximation to the "surrounding world of actual life-related practices," it can serve as part of a process of open-ended inquiry designed to help expose and interrogate the implicit decision-making processes essential to professional practice (Yanow and Tsoukas 2009, 1342). This focus on thoughtfully reflective deliberation and learning in the context of actual incidents as they evolve distinguishes craft-based learning reviews from an administrative rulemaking model. A craft learning model might bring attention to what policy requires, but its assessment of an officer's choices goes well beyond this. Furthermore, it does not insist that officers justify their responses retroactively in an environment carrying a high risk of censure or punishment.

Evaluations of what constitutes more or less desirable exercises of police authority can obviously be complicated and controversial, particularly when law and policy offer little guidance. How then would a supervisor help an officer understand whether their performance was substandard, merely adequate, or truly exemplary? As David Thacher observes (2020, 762), there is no simple answer to this question "for Dewey's alternative

model of rationality is fundamentally open-ended.” It is worth noting that this challenge is not unique to policing. Professionals in other fields, such as medicine, engineering, and law, may struggle to reach a “firm and uncontroversial evaluation of a particular decision” (Thacher 2020, 762). From the perspective of the craft learning model, this observation opens up rather than closes down discussions during the review process. With this in mind, the key task of supervisory review is to help patrol officers reflect on their choices, articulate how they tried to meet or exceed key performance standards, examine how and why they might have fallen short, and consider areas for improvement.

Other concerns revolve around individual police officer bias and existing norms in policing that might contribute to troubling police practices, such as favoring wealthy over poor suspects, or treating civilians differently due to their race (Lum 2009). Establishing performance criteria for supervisory review is clearly not sufficient for preventing questionable or morally indefensible decision-making (Thacher 2016). To mitigate these concerns, departments could involve community members in a similar but separate approach for reviewing body-worn camera footage. The purpose would be to solicit input on officer performance (Waddington et al. 2015). This could then be part of the process for formulating standards and identifying responses that are consistent with community values and expectations. While examples of troubling applications of the police craft and disagreements about police performance will undoubtedly remain, creating a system for critically reviewing officer performance to help increase transparency, foster debate, solicit community input, and improve external accountability, would be a large step forward in monitoring and managing everyday uses of police authority (Mastrofski 1996).

Although it cannot offer a definitive guide to action, reflection-in-action structured around performance standards does offer a number of benefits. In narrowing the scope for evaluating what should matter in handling an encounter to a practicable set of relevant criteria, standards help identify to what extent officers took some of the most relevant factors into consideration when making their decisions. Making these priorities explicit also helps reduce the “vulnerability” of patrol officers to “arbitrary performance evaluations after the fact” (Mastrofski 1996, 223). Moreover, conversations around specific standards can help with identifying important gaps in an officer’s knowledge and skills that might require additional training, and with clarifying to what extent an officer’s choices and goals were matched with the police department’s own priorities. Should these be misaligned, attempts could be made to understand where major differences lie and to address them. Giving patrol officers’ ownership in defining and developing performance criteria, rather than imposing regulations from above, may also increase acceptance of this new model for reviewing and guiding discretion.

Second, learning about officers’ attempts at spontaneous improvisation and the understandings behind them could help supervisors understand how an officer framed an initial problem, and how this might have shaped subsequent choices, perhaps even limiting their scope. It could also lead to the identification of innovative solutions, which could be disseminated to other officers to become part of a broader repertoire of “examples, images, understandings, and actions” for handling similar encounters in the future (Schön 1987, 66). In doing so, this kind of review process could help collect knowledge about surprising situations, where more routine responses fell short of

expectations. This knowledge could then accumulate and become a valuable resource for organizational learning (Schön 1987).

Finally, a nonpunitive reflective review process could contribute to patrol officers' normative understanding of important public values in policing, as they are encouraged to consider how the meaning and application of these values influenced their choices in relation to the practical realities of street-level policing (Thacher 2008). Albert Reiss (1971, 21) once noted that, as a profession, policing requires officers to make decisions affecting the "fate of people" that were not just technical but involved making "moral judgments." These judgments, powerfully shaped by situational contexts, tend to be implicit and rarely exposed to careful examination, and yet disagreements over normative considerations regarding what constitutes "doing the right thing" on the front-lines often lie at the heart of public controversies over police actions (Mastrofski 1996, 2018). The implications of a patrol officer's decisions for an individual's liberty or dignity, for example, may become taken for granted over time, and thus lead to complacency about the ideals a police officer should aspire to and the obligations they should obey (Thacher 2006). As Bayley and Bittner (1984, 55) have argued about craft, "Experience may teach, but it rigidifies." Reflection-in-action could benefit officers by helping them identify, clarify, and revise their understanding of varied and ambiguous values and develop ideas about acceptable trade-offs between them (Thacher 2001). These interpretations and justifications could then support more principled decision-making in future encounters.

In sum, while "in principle" there is no domain of police practice explicitly excluded from an administrative approach that relies on rulemaking, compliance, and oversight (Bittner 1983, 2), conceptually there are some key characteristics that help distinguish a craft learning model from this approach. These resemble those identified by Thacher in his learning model for use-of-force reviews (2020). Aside from the fuller and more direct participation of patrol officers in the implementation of mechanisms for guiding discretion, the craft learning model puts greater emphasis on: (1) an open-ended, constructive, and flexible learning process in response to the challenges of dynamic situational factors, rather than a more formal disciplinary approach for enforcing general and fixed bureaucratic rules; (2) engaging directly (rather than incidentally) with relevant and oftentimes ambiguous values that shape goals worth attaining and underlie the tactics for accomplishing them (Kelling 1999, 39); and on (3) continuous rather than infrequent reviews of everyday police-public encounters, as a means to improve police officers' capacity for using their professional judgment.

Thus, our purpose is to examine whether it is possible to identify a manageable set of craft-based standards for reviewing street-level performance, and to demonstrate their potential for helping inform and channel officer decision-making as part of a systematic learning process centered on reflection. While police scholars have long recognized the possibility of using police officers' collective experiences to improve discretion (Bittner 1983; Bayley and Bittner 1984; Mastrofski 1996; Thacher 2008; Willis and Mastrofski 2017), few attempts have been made to advance this insight empirically, or to develop the logic, substance, and uses, of a craft-based learning model, especially in relation to the dominant rulemaking alternative (Thacher 2008).

RESEARCH SITE AND METHODS

In 2012, we showed a video clip of a neighbor dispute to thirty-eight patrol officers in “Newbury” (a pseudonym) and asked about the performance of the officers depicted. Similar to other approaches for trying to access intuitive knowledge that guide professionals’ decision-making but can be hard for them to articulate (Klein 2011; Mangels, Suss, and Lande 2020), we used this clip to simulate the “real-life” conditions of police work, and we relied on semi-structured interview techniques with follow-up probes (Kelling 1999; Kahneman and Klein 2009, 517). The video clip was part of a larger project examining officers’ general views about the police craft and police culture (Willis and Mastrofski 2017, 2018). We also chose this particular encounter because of its uncertainty, as the proper course of police response to the dispute did not appear obvious. Moreover, unlike a low-discretion scenario, where law, policy, and training would be foremost in an officer’s mind, our high-discretion clip increased the likelihood that experience-based decision-making came into play.² Because the purpose of our larger research project was to learn different officer perspectives on various dimensions of the culture of craft, we relied on individual interviews and did not use the kind of focus-group approach to deriving standards that we now envision as part of a craft learning approach.

At the time of data collection, the Newbury Police Department served a city of about 150,000 people. Nearly 70 percent of its residents were white, and the rest were either Black or Hispanic. Slightly less than 10 percent of residents lived below the poverty line. Newbury’s violent crime rate was significantly below the national average for cities of similar size. The department had approximately 300 officers, 100 civilians, and a long history of being service-oriented.

The selection of officers for interviews depended on the first stage of the research project, which involved the administration of a written questionnaire to officers assigned to Newbury’s patrol division. One hundred sixty-five officers were eligible to complete the survey, and opportunities arose to hand out 118 surveys (missing those with days off, training, etc.). Of these 114 were returned, for a 97 percent response rate.

Because we were initially interested in the possibility that those patrol officers recognized by their peers as “master craftspeople” performed differently to others, officers were asked to identify up to two of the top performing patrol officers employed in the department, ranking them first and second. They were allowed to consider officers of any rank and in any unit, focusing on their abilities to perform patrol officer responsibilities (for more details on this process, including the distribution of responses, see Willis and Mastrofski 2017). Approximately nine of every ten respondents indicated

2. The clip was taken from *The Police Tapes* (1977), a documentary about the New York City police directed by Alan and Susan Raymond. Although dated, the clip clearly depicts police handling a disorderly situation that is typical of the mundane, but problematic, encounters that characterize peacekeeping and that continue to confront police officers today. While undoubtedly key features of policing have changed, such as training and research on the police, much has remained the same. Several of our respondents commented on the officers’ different uniforms, but they also noted how similar this neighbor dispute was to the kinds of situations they routinely handled, and none commented that the clip was too old to be relevant. We should note that at the time of this research, we did not have access to the kinds of recordings captured through body-worn camera footage that are now available.

TABLE 1.
Characteristics of all interviewees (N = 38)

Current Position		Years as a patrol officer	
Patrol officer	33 (87%)	Less than 3 years	3 (8%)
Sergeant/Acting Sergeant	4 (10%)	3–6 years	15 (39%)
Detective	1 (3%)	7–9 years	6 (16%)
		10 or more years	14 (37%)

that they were most influenced in their selections by their direct observations of the identified officers at work. Officers were asked to describe the features of the officers' they had selected that most impressed them (open-ended question). The most mentioned traits were good work habits (56 percent of respondents): being a hard worker, thorough, attentive to detail, disciplined, positive attitude, reliable, following through, showing initiative, and following orders. Almost as frequently mentioned was the amount of knowledge or experience in doing patrol work (55 percent). Also frequently mentioned were ability to speak well, have good interpersonal skills or work with the community (38 percent), and being a good leader/helpful to other officers/a good instructor/have a team focus (29 percent). With the exception of a fairly high level of agreement on good work habits, Newbury's officers showed a fair degree of diversity in what impressed them about the top performers they identified. Nineteen officers were selected because they were named most frequently by their peers (three or more mentions). We called these "high reputation" officers. In Newbury, fifteen of the department's 313 officers (or about 5 percent of the force) accounted for 48 percent of the high reputation officer nominations. We then selected a random sample of nineteen additional officers assigned to patrol as a comparison group. Because two high reputation officers were unavailable, we decided to increase the size of the comparison group with two additional randomly selected officers. Consequently, our final interview sample consisted of thirty-eight patrol officers (seventeen high reputation and twenty-one comparison group officers).

A summary of respondent characteristics is shown in [Table 1](#).

The largest proportion of respondents was patrol officers with three -to- six years of patrol experience. In terms of the distribution, high reputation officers were more likely to have seven or more years of experience than comparison group officers (See [Table 2](#)).

As for demographic characteristics, there were six female officers in our sample and thirty-two males. In terms of race and ethnicity, the largest categories consisted of five officers identifying as African American, six as Latino/a, and twenty-three as White. The interviews were voluntary and confidential, had been approved by our university's institutional review board, and lasted an average of seventy minutes. Interviewers took notes on officers' responses, which were transcribed within twenty-four hours. The interviewer did not know which officers had been identified as high performers.

We presented the video clip to respondents about half-way through a longer interview, and we asked them to judge the quality of the officers' performance. We also forewarned respondents we were going to stop the clip at a key decision point as the

TABLE 2.

Distribution of High Reputation (HR) and Comparison Group (CG) officers by years as a patrol officer

Years of experience	HR (N = 17)	CG (N = 21)
Less than 3	0	3
3–6	6	9
7–9	3	3
10 or more	8	6

encounter was still evolving, and ask them to assess these officers' decisions and what they would have done differently. This was to help ensure that respondents did not have the benefit of knowing how the encounter ended in making their assessment.

In brief, the clip showed two patrol officers responding to an actual dispute between two neighbors in an apartment building in New York City sometime in the 1970s. A complainant (C1) was upset that the woman (C2) living in the adjacent apartment had been pounding on her door with a flatiron (an iron that was heated on a stove and used to press clothes) and physically threatening her. When questioned, the woman with the flatiron said she was frustrated by the complainant slamming her door. The officers were puzzled that this should be the cause for retaliation, and for so much hostility. They tried several times to reason with her, before eventually giving up and telling her that if they had to return that night, she would be taken to the precinct. The officers then left, disposing of the flatiron through a trash chute. The entire encounter lasted about seven minutes. We stopped the clip when the second neighbor admitted to using the flatiron, and we asked the officer, "What would a highly skilled officer be doing at this point in the situation?" At the end of the clip, we asked respondents to judge the success of the officers' overall performance, and to suggest a better approach (recognizing this was informed by the benefit of hindsight).

The lead author used a combination of deductive and inductive analyses to code the interviews. Deductive codes were informed by predetermined themes based on previous research on standards for assessing the police craft (Bayley and Bittner 1984; Mastrofski 1996, 2018). At the same time, inductive coding led to the creation of emergent themes that had not been previously considered (Boyatzis 1998). This combination is consistent with a systematic flexible coding approach to in-depth interviews, designed "to support rigorous, transparent, and flexible analysis of in-depth interview data" (Deterding and Waters 2018, 1). To avoid confirmation bias during coding, the coder was blind to which officers had been identified as high performers by their peers.

The coding process began with reading through all the interviews to identify preliminary themes. Each interview was then read more carefully to articulate broad index codes that reflected the questions asked in the interview guide and concepts related to performance. This process was further accompanied with the writing of a detailed analytic memo to capture the coder's evolving thoughts and reflections and to help refine the conceptual themes that were emerging across the officer interviews (Patton 2002).

The next stage involved revisiting the themes that had now been identified and engaging in a more focused, line-by-line, analytic coding scheme of the interview text. This also led to the generation of new codes, or the combination of existing ones. Finally, the identification of respondents' attributes (namely whether or not they were a high reputation officer) was left until the prior analysis was completed. This was then the basis for going back to see if these two groups differed on any key dimensions of performance appraisal. Particular attention was paid to recording differences between individual officer responses, knowing that members of a craft community often make subtle distinctions (Kritzer 2007, 326).

Because police officers say they are able to distinguish skilled from unskilled performers (Bayley and Garofalo 1989), we initially wanted to see if these two groups judged the officers' performance in the clip of the neighbor dispute differently. However, in our analysis we were unable to detect clear differences between the views of the high reputation officers and those in the comparison group. The only notable difference was that high reputation officers were less likely to advocate for an arrest (see below). Hence, we generally do not distinguish between these groups, unless there is a marked divergence of opinion. Given that other studies have shown discernible differences between officers identified as experts and those not so identified (Bayley and Garofalo 1989; Mangels, Suss, and Lande 2020), this could be due to a limitation in our methods. Because we did not ask officers to identify those peers they considered to be most skilled in handling disputes (just those highly skilled at patrol work in general), "experts" in dispute resolution might not have been selected for our pool. It might also be easier to discern differences between experts and novices through direct observation of officer behavior rather than interviews. We simply note here that our craft model is not dependent on whether the officer is high reputation or not. Below we present the seven possible standards for assessing officer decision-making derived from the coding of these thirty-eight interviews (accountability, lawfulness, problem diagnosis, repair of harm, economy, fairness, and safety and order).

CRAFT-BASED STANDARDS FOR ASSESSING STREET-LEVEL DECISION-MAKING

Accountability

In addition to discouraging undesired conduct, a fundamental principle of U.S. criminal law is that people should be held accountable for their actions, especially when they engage in wrongdoing that harms others (Hart 1958, 410; Mastrofski 2018). In the neighbor dispute, the officers in the clips tried different approaches to make the woman accused of damaging her neighbor's door understand why her behavior was wrong, why she should not make threats and damage property, and how she was responsible for the consequences of her choices.

Among our respondents, the most popular approach was to resort to informal means of accountability with almost two-thirds of respondents suggesting that the officers try to advise, persuade, negotiate, or mediate. For example, one officer (O30) explained that he would "Talk to the neighbors separately," and that he would say

something like, “Let’s talk about the problem and try and figure out a solution.” Other officers focused on giving advice to the parties, which could include going to the building manager (seventeen respondents) or to a magistrate (five respondents). In comparison, only eight respondents suggested they would warn or threaten the second neighbor and seven respondents indicated they would arrest. There were few clear differences between the high reputation officers and those in the comparison group, but one clear pattern was around the decision to arrest: high reputation officers were much less likely to say that an on-scene arrest was appropriate (only one did so compared to six from the comparison group). Another area of disagreement among interviewees was on how best to mediate between the parties—that is whether to bring them together or manage them separately. Some recommended bringing together both parties during any mediation attempt. For example, O6 said:

You could also try sitting both of them down together and try to get them to air out their differences and divulge their dirty laundry while you’re actually sitting there attempting to be the mediator for the problem. If you separate them and talk to them separately to try to get them to air out their differences you’re actually just putting more of a distance between them, where the second that you leave everything will just go back to normal because neither knows the side of the other party.

Other officers stated that any attempts to counsel or mediate between the two parties should be conducted by *separating* the two neighbors. The following is from our interview notes:

O18 suggested having the two officers split up and talk to the two women separately at the same time. O18 said that these officers could then get back together and “compare stories” so that the two women “won’t hear what each other said.”

The various approaches to accountability demonstrate some diversity among our interviewees about the best way to call these parties to account in our neighbor dispute. One obvious area of disagreement (and one we discuss later) is to what extent the officers should impose their authority on the situation. At the low end of the scale of control or intrusion, some interviewees advocated simply trying to persuade or advise both parties, while at the highest end of the scale, others wanted to arrest.

Lawfulness

Another key dimension of police performance that emerged was the lawfulness of an officer’s actions. The law sets forth “technical standards and expectations that stipulate or guide the officer’s actions in a number of domains” (Mastrofski et al. 2000, 313). In interpreting whether and how the law applies, officers must know its provisions and assess what the evidence warrants, including whether it supports the deprivation of a person’s liberty in the form of an arrest—a serious infringement on an individual’s

democratic rights. In the context of our clip, some key legal questions that arose were whether the officers did anything not permitted by law or *omitted* to do something required by it (Mastrofski 2018).

There was disagreement among our interviewees about whether the neighbor dispute qualified as a criminal or a civil matter and whether or not there was sufficient evidence to make an arrest. So, for example, O14 described the situation as mostly a civil matter with people not “playing nice in the sandbox together,” while O24 considered it a serious criminal issue stating, “All I have to say is that it might very well end up as a homicide later, because of how they left it.” We noted above that only a small proportion of officers, like O24, felt there were legal grounds for an arrest. For example, O21 said, “I would have arrested her right there. That’s a statement made in my presence about doing harm to someone else.” Others were less certain there was sufficient evidence for an arrest. O16 could not tell a crime had happened, and O23 was “embarrassed to admit” that he did not know if he could make an arrest for “destruction of property.”

From the perspective of our interviewees, there was less legal ambiguity over the officers’ decision to take the neighbor’s flatiron and dispose of it down the trash chute. Some noted this was her private property and commented that the officers in the clip did not have the right to simply remove it from her and throw it away. Other respondents emphasized its value as evidence. O10 said you have to “fill out the proper form any time you take something, otherwise it’s larceny,” before adding, “they shouldn’t have thrown physical evidence down the trash.”

Based on our interviewees’ comments, the lawfulness of an officer’s behavior in terms of both acts of omission and commission is an important dimension of police performance. Our interviewees’ uncertainty about whether this was a civil or criminal matter, and if the law permitted them to make an arrest, suggests that officers’ might struggle to interpret the law even in those situations which might not appear on the surface to be legally complex (Gould and Mastrofski 2004).

Problem Diagnosis

As other police researchers have noted, mitigating a problem depends in no small measure on an officer’s ability to make an accurate diagnosis of what *exactly* the problem is (Muir 1977). In the case of our neighbor dispute, we can envision several questions, some of which were raised in our interviews: What prompted such a hostile reaction from the second neighbor to the door slamming? Was this the first time the police had been called to this address?

Almost half of our interviewees (equally divided between the high reputation and comparison groups), remarked on this element of police performance, with most expressing dissatisfaction that the officers in the clip had expended far too little investigative effort in learning about the problem between the two neighbors. For example, O19 said:

They could have talked to C2 more to get more background. They still don’t know what the problem is. They could try to break it down and see what the underlying issues are. Why is she so mad at her? What is the noise? Where is it coming from? Through the wall? From out in the hallway?

Similarly, O15 said, “You also really don’t know what she does. It’s possible that she works midnights and maybe the noise is a serious issue in her mind. You have to ask more questions.”

Central to many of our interviewees’ diagnosis of the dispute was whether the second neighbor was experiencing some kind of mental illness. Some respondents said they would ask her directly if she had mental illness, and if she was on medication, while others said they would contact a family member, or call in mental health professionals to help. O35 said, “My approach would have been to call external services, such as mental health or the office of aging and the elderly.” He explained that the “office of mental health is better trained to dig deeper into issues.” Respondents also said they would have sought to better understand the nature of the problem by consulting other sources of information, such as checking the location’s call history, or consulting the building manager. These different approaches reveal how officers rely on diagnostic systems other than just the law, which rarely provides direction when the officer’s goal is not legal resolution through formal sanctions.

Identifying what to ask about a specific problem and knowing which diagnostic systems are available for this process takes considerable skill. Despite the crucial importance of problem identification, police departments pay little attention to assessing this dimension of police performance (Mastrofski 1996). Moreover, the time and effort that officers to take to investigate the nature of the problem can send a signal to a civilian that the issue is being taken seriously.

Repair of Harm

In their performance assessment, our interviewees criticized the officers in the clip for minimizing the seriousness of the dispute, failing to improve the immediate situation before leaving, and for doing too little to ameliorate the risk that the problem might re-occur in the future. For example, O40 said:

I always like to put myself in the victim’s shoes. If I was the first lady and I just came home from work and this other lady had attacked my door and would have attacked me if I was there, I like to think of how I would feel. I would want the situation taken care of. That is not what the officers did.

Moreover, researchers have noted that in trying to understand the nature of a problem, police officers must consider how by merely “entering the situation” they “affect the circumstances,” possibly provoking an unintended reaction (Muir 1977, 155, 165). Some of our interviewees alluded to this source of disruption when they noted the officers in the clip had actually made the problem *worse* through their response (or lack thereof). For example, O6 said:

The lady with the flatiron was the same after the police spoke with her as she was in the beginning. And in some ways it seemed as if she was even more agitated. At the beginning she answered the door in a sort of pleasant manner, but by the end she was barking out threats that she would bash the other

lady with the flatiron and that she wouldn't make any promises that she wouldn't do this after the police had left.

The criticism that the officers in the clip should have "done something more" than they did applied to both the immediate situation and to its longer-term aftermath. For some interviewees, their goal would be more modest than trying to address the deeper underlying problem that gave rise to the dispute in the first place. These respondents suggested they would have done more to *change* some key element of the situation (such as providing the second neighbor with mental health resources, or reducing the tensions between the neighbors). For O21, it was about trying to get the second neighbor to think about alternatives. He would have told her, "I understand it is frustrating always hearing that noise. I would hate it too. But all you're going to do is get yourself in trouble. You seem like a nice lady. Is there anything else we can do to deal with this in a better way?" For other officers "doing something more" involved making a more concerted effort to provide a longer-term solution that reduced the risk of the problem reoccurring in the future. This view is captured by O33's comment: "The officers should try to help resolve the problem, so they are not back there a week later."

Whether our interviewee's focus was on shorter- or longer-term outcomes, both groups of respondents were generally dissatisfied that the officers did not take more time and effort to seek alternatives that could have improved the odds of repairing the harm of the situation rather than merely warning the suspect before leaving. Hence, in Newbury at least, craft seemed to accord considerable importance to longer-term consequences of police decision-making, as well as short-term considerations.

Economy

The expectation that police provide a "definitive resolution" to almost every encounter (Mastrofski 1996, 215) must be anchored in the reality of the resource constraints within which patrol officers must operate. The officers in the clip spent less than ten minutes handling this encounter, and while our interviewees were dissatisfied with their failure to deal with the dispute more effectively, few provided a direct estimate of how long these kinds of calls should take. Four officers made some reference to resource expenditure. O18 suggested "that it may be that the officers don't have a lot of time" and that "these types of situations require time to figure them out" and that the night "could be busy." O31 was direct in his appraisal: "They did fine, actually. You can't spend all of the time in the world on a civil disturbance," and O33 did not blame them "for not spending too much time trying to sort things out."

Police managers, similar to other legal authorities, operate under considerable budget constraints and are under pressure to deliver their services as efficiently as possible (Galanter 1974, 125). At the same time, the amount of time officers spend with a civilian can be seen as a measure of attentiveness, albeit a relatively blunt one. In a study of police contacts with civilians who were in conflict with one another at a level of at "least verbal disagreement," the amount of time that officers spent with each of these citizens "ranged from 2 minutes to 3.5 hours" with a "median of 23 minutes" (Mastrofski 1999, 6). Using time as a measure of attentiveness helps illuminate not just how much

time officers expend when interacting with civilians, but can also show what they do with that time and what the outcome is. Police organizations currently know very little about how their officers' respond to the wide range of encounters that are not crime or traffic accidents and that are resolved informally and without a formal record, such as this neighbor dispute.

Fairness of the Process

Even if police officers cannot deliver the outcome different parties desire, people care very much about how they are treated (Tyler 2004). It is important that those involved have a chance to tell their side of the story, that the officer listen attentively, and that the officer's decision is not biased toward one party or the other. People also expect the police to show that they care about whatever problem a person is experiencing and to make a good faith effort to resolve it.

Many of our interviewees felt that officers in the clip, in general, behaved professionally, not showing favoritism to one neighbor over the other and remaining calm throughout the encounter. In response to our question about whether there was an approach, with hindsight, which would have worked better, O31 said:

I don't really see it. The officers reached out to all of the resources at the time. Instead of maybe camping outside the door to make sure that nothing further happens, the officers did all that they can. They did a good job at keeping their calm, explaining the actions and consequences and if anything else happens then that is on her.

However, our interviewees also suggested there was room for improvement in how they treated the two neighbors. O13 thought that the officers were provoking the second neighbor by pointing their batons at her, and several interviewees felt that they tried to influence unduly the complainant's decision by suggesting she did not want the suspect "locked up." O24 said they were "dismissive" toward her by advising her to go to management or through the court, while O36 said that by asking leading questions, the officers were trying to talk "themselves out of work."

Modern management prescriptions and the police reforms that embody some of their key principles, such as community policing, place high value on treating people with respect. Importantly, this can contribute to the legitimacy of the police, which brings with it a host of benefits (Tyler 2004). But the comments of our interviewees also suggest that the process of treating people fairly comprises a range of officer behaviors. These include small but thoughtful demonstrations of caring for a person's well-being.

Safety and Order

The final performance standard to emerge from our interviews was how well officers in the clip attended to safety while on scene. The potential dangerousness of any situation, even the most benign, and the priority afforded preventing injury to civilians

or officers themselves is a well-documented feature of the traditional police culture (Bayley and Bittner, 1984; Paoline 2004). Failure to maintain safety and order at the scene undermines the ability of an officer to accomplish any other objectives and figures prominently in almost every field work account since Skolnick's pioneering study (1966).

Eleven respondents highlighted safety concerns during our interviews in equal proportion between the reputation and the comparison groups. In doing so they showed concern for both the officers' safety and those of the parties involved. For example, O22 said, "Talking to the offender out in the hallway was problematic because the complainant can probably hear what is being said." He continued by saying that this will cause the women to "yell at each other about what they overheard outside of the door." Other interviewees criticized the officers in the clip for allowing the offender to reach behind her door to retrieve the flatiron. O26 observed that they "asked her to get the flatiron for them. This is really dangerous. The better thing to do would have asked her where it was while one of the officers goes and retrieves the iron."

The neighbor dispute raised a number of questions about the appropriateness of the officers' response to safety concerns and whether they minimized the possibility of violence and/or risk of injury. When either an officer or a civilian is injured during an encounter, departments automatically generate an inquiry that results in significant documentation. This approach, however, overlooks those situations where injury *could* have occurred.

DISCUSSION

Egon Bittner noted decades ago that police officers' own demands for "discretionary freedom," and a lack of formal direction around peacekeeping activities requiring authoritative intervention, presented a formidable barrier to efforts to monitor, assess, and learn from patrol officers' collective experiences (Bittner 1967, 715). Even though how patrol officers use their authority in everyday encounters with the public is one of the most "fundamental aspects of their performance in a democratic society" (Mastrofski 1996, 208), few strides have been taken to create systems for reviewing and evaluating everyday performance to help promote higher quality or more democratic policing. The police mantra that "every situation is different" (Bayley and Bittner 1984) also helps perpetuate the idea that any attempts to assess an officer's skill and judgment are illegitimate, or examples of counterproductive "Monday-morning quarterbacking" (Bittner 1983, 5). It is hard to imagine how significant improvements to patrol officers' uses of their everyday discretionary authority can occur under these conditions.

Our interviews suggest that despite the "unavoidably complex" nature of policing practice (Thacher 2019, 280), it is possible to have patrol officers evaluate encounters, comment on the strengths and weaknesses of an officer's decisions, and identify areas for improvement, as a basis for deriving performance standards. While there were obviously disagreements between respondents about the best of course of action in the neighbor dispute, we still believe there could be considerable merit to using these, or similar standards, as part of a more frequent and systematic review process to help officers reflect on the choices they make, and what they could have done differently. Because this

study was part of a much broader project examining officer perspectives on the police craft, we have relied on our individual interviews, but police agencies interested in deriving performance standards could use small focus groups of police officers. Structured around a seminar format, these could be designed to provide plenty of opportunity for the sharing of different viewpoints, careful deliberation, and for building consensus (Mastrofski 1996). Although the standards we identify here are similar to those discussed elsewhere (Bayley and Bittner 1984; Mastrofski 1996), what makes the craft learning approach novel is its more comprehensive approach to devising a systematic process for reviewing and learning from routine exercises of authority in police-civilian encounters. Because standards are part of an open-ended process of inquiry and developed within individual police agencies (and in relation to local community expectations about police behavior), there is also always the possibility of new standards being developed, or old ones being refined. In addition to this creative element, researchers could subject the tactics officers identify as tools of the trade to empirical testing to learn how well they work as intended (Willis and Mastrofski 2018). In this way, tactics that have become taken-for-granted, but are ineffective and potentially harmful, could be replaced with those for which the evidence-base is more promising (Lum 2009). This validation of different officer approaches could also be an opportunity to mitigate, or perhaps even resolve, disagreements between groups (i.e., management and patrol officers) based on assumptions of what does or does not work in a specific situation (Mastrofski 1996, 232).

Once performance criteria were established, we can imagine a supervisor using reflection-in-action as part of a dialogue with one of their subordinates. This would involve reviewing body-worn camera footage of the officer's choices in a specific encounter that the officer considered problematic, because their usual responses did not seem to work well. Pausing the footage would enable the supervisor to ask what they were thinking and trying to accomplish, as the action unfolded. This would not be:

... the leisurely, systematic, ex-post-facto rehearsing and critiquing of the he-said-she-said of an interaction (the kind of "Monday-morning quarterbacking" of "And then I did this because he said that, so I did this other thing ...") that would suggest a stepping-back-and-engaging-the-question characteristic of detached reflection. Instead, we see in it the "online," "real time," in-the-midst-of-it-all seemingly "split second" judgments that lead the practitioner in different directions from his established, routine practice (Yanow and Tsoukas 2009, 1355).

As part of this reflective process, standards could help narrow the potentially dizzying array of choices available to the officer when on scene to some of the most relevant for judging how well they exercised their authority. Take problem diagnosis, for example. Several respondents suggested that the officers in the clip were not sufficiently persistent, or failed to ask the right kinds of questions, when trying to identify what the nature of the problem actually was between the two neighbors. As a result, they may have failed to appreciate the dispute's seriousness, and to identify and address any deeper underlying causes that gave rise to the second neighbor retaliating with the flatiron. Similar to Herman Goldstein's (1979) criticisms about defining incidents in vague

and bureaucratic terms, characterizing the conflict between the two neighbors as simply a dispute gives little insight to the substance of the problem, or to the complicating factors that might have been at play. Was this merely a quarrel between neighbors, or the prelude to something more violent? One of the neighbors mentioned that they used to be good friends, so what might have contributed to the current acrimony? Was there evidence that tensions between them had escalated over time? Had the neighbors previously tried to find a solution to their differences? Was the second neighbor a person with a mental health condition? Much of the focus in police organizations and by police researchers is on the outcomes of a patrol officer's decisions, and yet any actions an officer decides to take are first dependent on the officer's ability to make an accurate diagnosis (Willis and Mastrofski 2018). This scenario also helps illustrate some of the limitations of administrative policies for guiding discretion. Unlike a review process strictly constrained by rules, the process we describe uses standards that act as approximations to inform decision-making and that allow for flexibility in how officers account for the complex contingencies of street-level encounters. As such, they can help establish higher expectations for good uses of police authority by trying to "accommodate variation in the behavior that meets [these standards] while distinguishing it from behavior that does not meet them" (Cohen and Feldberg 1991, 41).

This process of assessing how well officers "measure up" to a particular standard could be aided by developing its content more fully in order to give additional guidance around the kinds and levels of behavior expected. For example, the standard of problem identification could be defined as:

Using available resources, including the parties involved, to conduct a broad inquiry into the specific nature of the problem. The purpose of the inquiry is to define the problem accurately, to identify possible underlying causes through open and persistent probing, and to assess carefully the seriousness of its effects.

Applying this standard would help supervisors identify areas for improvement and demand higher levels of performance from their officers. Moreover, some standards might lend themselves to more precise forms of performance measurement. Fairness, for example, could be assessed by a first-line supervisor's qualitative evaluation of an officer's actions in relation to each of the four elements of procedural justice (voice, dignity and respect, neutrality, and trustworthy motives). Alternatively, the evaluation could be based on a more precise scale, constructed with the help of academic researchers, for determining an officer's specific encounter-based procedural justice score (Mastrofski et al. 2016).

When reviewing BWC footage with an officer of a challenging encounter they experienced, a supervisor could use open-ended questions and probes to help the officer understand how they came to their characterization of the problematic encounter, as it was occurring. Such an exchange could help reveal any assumptions the officer made at the time, which could then be the subject of additional inquiry. In the case of an officer experiencing a similar problematic dispute to the one we showed, we can imagine them being encouraged to reconsider whether their questions to the parties involved helped solicit useful information, or needed to be asked differently, and to reflect on how, if it all, they considered the encounter to be similar or different to others, as it was occurring.

The latter opens up the exploration of alternative and potentially more effective approaches than the standard response of talking to both parties and issuing a warning.

The craft learning model review process could help structure similar deliberations around other relevant standards, such as how respectfully the officer treated both parties (fairness) and where there might be room for improvement, and whether they held the second neighbor sufficiently accountable for her actions (accountability). Discussions of this kind could also reveal to what extent an officer excelled at, or failed to consider, other performance dimensions, such as attending to safety and order at the scene in ways that best protected all parties involved, including the officers themselves. It would also be an opportunity to identify whether the officer's goals were consistent with those of the agency and, whether their actions were consistent with existing policies and procedures. The latter could be a basis for the officer to revise their goals, or for the agency to establish clearer expectations for what its officers should be trying to accomplish on the front-lines. Using this case as an example, a department might suggest its officers expend more time and effort to repair the harm that caused the vandalism of the neighbor's door than the officers depicted.

Another potential benefit of using standards is to help identify important gaps in officer knowledge and skills that might require additional training. For example, the law is sufficiently ambiguous that it is open to differing interpretations. When it came to lawfulness in our neighbor dispute, there was disagreement among respondents about whether and how this standard applied in this specific context. This is a valuable insight, as the law is an important source of self-guidance for patrol officers, and a means for expanding or narrowing the range of decision-making options available to consider. Helping officers learn how criminal or civil law may or may not apply to common situations, such as disputes, can not only help clarify what actions are lawful, but can also be a means for expanding or narrowing the range of decision-making options that then broaden the officer's repertoire for responding to future incidents. Recognizing the range of legal options available to an officer admits the possibility of a more artful response, while knowing clearly what the law does *not* permit can also set limits on police officers' sense of responsibility, giving them "a sanctuary from public blame and self-criticism" (Muir 1977, 259). This kind of performance review could also help supervisors identify what evidence-based research suggests are some of the most promising approaches for managing some of the most challenging aspects of complex social interactions involving conflict, including de-escalation tactics and procedural justice (Mastrofski et al. 2016; Wolfe et al. 2020). In turn, scientific knowledge about the likely outcomes of their decisions could help officers make more informed decisions, such as whether separating disputants is more effective at resolving interpersonal conflicts than talking to them together—a source of disagreement among Newbury's patrol officers.

Reflection-in-action could also help identify officers' more innovative and promising attempts to resolve or mitigate a problem, which could then be disseminated more widely to benefit others. Patrol officers might share their insights with one another informally, but departments have few structures in place to tap the creative possibilities of actual police practice (Thacher 2008). Rather than adhering to more routine responses, one of our interviewees suggested padding the frame of the neighbor's door with rubber to deaden the noise. Once shared, this creative insight could become part of other officers' knowledge base when seeking to repair harm in future incidents that share similar features as the neighbor dispute (Schön 1987).

Finally, in assessing the officers in the clip, our respondents helped identify important normative concerns. One of the most obvious was the legality and advisability of making an arrest, and there was some disagreement among our officers about the appropriateness of this response. The majority of officers felt an arrest was not warranted, but this was the one area where there was a clear difference between the high reputation officers and the comparison group (one versus six supporting an arrest, respectively). It is difficult to know whether our more experienced high reputation officers seemed less likely to make an arrest because of skepticism or sagacity, and this is a potential avenue for future research. They may have considered an arrest too time consuming, or more likely to exacerbate than mitigate the problem.

Whatever might explain these differences, officers' differing opinions on whether to arrest helps raise a central concern in policing. In a democracy which seeks to minimize government intrusions on individual liberty, officers must strive to be as judicious as possible when deciding to take a person into custody (Harmon 2016). Decisions around important public values, such as liberty and order, are often controversial, but policing is a moral enterprise. As such, it is essential to consider and debate these values in order to help advance normative understanding (Dagan 2017). Using a review process organized around performance standards to also encourage officers to reflect on the normative basis of their decisions could help them identify and clarify the values that matter in a given context and apply relevant criteria for deciding "the right thing to do" (Mastrofski 2018). In the case of the neighbor dispute, these considerations should go beyond simply assessing the lawfulness of an arrest to include its substantial "physical, financial, psychological, or social" costs, whether these are "to suspects, families, officers or communities" (Harmon 2016, 320). Officers might also weigh other values relevant to justifying a decision, such as the role of retribution or just deserts. In the case of the second neighbor, did the suspect deserve to be arrested for the harm she caused in damaging the complainant's door, and how much should an assessment of her blameworthiness depend on a prior record of similar outbursts?

While researchers and scholars may have no superior claim to moral expertise than a practitioner, they could play a role in helping to systematically uncover and clarify which values were at play, and by assisting in their meaningful critique (Thacher 2001). While there may not be consensus on whether or not to arrest, a craft learning model might be able to contribute to more principled choices. It would do so by helping patrol officers engage more fully with their own moral intuitions (which are often made quickly) (Haidt 2001), and by helping them refine the reasons for their value judgments (or even revise the judgments themselves).

CONCLUSION

We recognize that the standards we derived from the experiences of Newbury's patrol officers are not exhaustive, that others could be added, and that the role of the community in this process requires fuller treatment. We also recognize that while the standards identified here would seem to have considerable face validity, they are likely to vary across different police organizations in both substance and application, depending at least on a department's leadership, on broader community expectations,

and on the quality of police supervision (Wilson 1968; Mastrofski et al. 1987). Regarding the latter, first-line supervisors often resist reform and vary considerably in their supervisory styles (Brown 1981). Even though some research suggests that officers might appreciate greater guidance from first-line supervisors in highly discretionary situations (Engel 2000: 284), there are significant barriers to supervisors' abilities to influence their subordinates' actions. These include the solidarity of patrol officers and the transitory nature of supervisor-officers relations.

Moreover, there are several limitations to our research, including that it involves only a single agency, and that it sketches a conceptual model that awaits empirical testing. Moreover, we do not address the nettlesome issue of deciding which performance standards should receive priority over others in a particular situation, especially when these might conflict (Thacher 2001). We also do not focus on the formidable barrier presented by a police subculture where police supervisors are generally reluctant to subject patrol officer decision-making to close scrutiny (Engel and Worden 2003; Willis and Mastrofski 2017). Perhaps with this background, our vision for a form of police professionalism grounded in street-level experience looks naïve, and is so unlikely to succeed that it is not even worth attempting.

On the other hand, the long history of reformers' struggles to effectively control police discretion suggests that it might be time to try something more novel to supplement administrative rulemaking. As Lincoln once wrote about the willingness needed to embrace a better future: "The dogmas of the quiet past are inadequate to the stormy present, and we must rise with the occasion. As our case is new, so we must think anew and act anew." Moreover, there may be good reasons to be sanguine about the potential benefits of a craft learning model. Some research suggests that police supervisors can affect the development of their officers and the judiciousness with which they exercise their discretionary authority (Muir 1977), and that reform from the bottom up can be successful in changing behavior (Toch 1980). Recently, a study showed that even brief meetings between officers and sergeants for talking about encounters on the street could lead to "more measured responses to later incidents" (Owens 2020: 1). Those officers who met regularly with their supervisors were less likely to resolve incidents with an arrest than members of a control group, and to be involved in incidents where force was used. The authors attributed this change to discussions prompting officers to reflect on their thought processes and actions. These meetings helped "slow down" officers' thinking on the street, so that they paid closer attention to the reasons for, and possible implications of, their decisions (Owens 2020, 2). Perhaps then this is the time to fully recognize and appreciate the complexities of police discretion in everyday encounters, to use standards to raise expectations about its use, and to experiment with bolder and more innovative review processes for making it more deliberate, transparent, and principled.

REFERENCES

- Bayley, David H., and Egon Bittner. "Learning the Skills of Policing." *Law & Contemporary Problems* 47, no. 4 (1984): 35–59.
- Bayley, David H., and James Garofalo. "The Management of Violence by Police Patrol Officers." *Criminology* 27, no. 1 (1989): 1–26.
- Bayley, David. H. *Police for the Future*. Oxford, UK: Oxford University Press, 1994.

- Bazon, E. "Police Reform is Necessary. But How Do We Do It?" *New York Times*, June 13, 2020. <https://www.nytimes.com/interactive/2020/06/13/magazine/police-reform.html>.
- Bittner, Egon. "The Police on Skid-Row: A Study of Peace-Keeping." *American Sociological Review* 32, no. 5 (1967): 699–715.
- . *The Functions of the Police in Modern Society*. Chevy Chase, MD: National Institute of Mental Health, 1970.
- . "Legality and Workmanship: Introduction to Control in the Police Organization." In *Control in the Police Organization*. Edited by Maurice Punch, 1–11. Cambridge, MA: MIT Press, 1983.
- . *Aspects of Police Work*. Boston, MA: Northeastern University Press, 1990.
- Black, Donald. "The Social Organization of Arrest." *Stanford Law Review* 23, no. 6 (1971):1087–111.
- Boyatzis, Richard E. *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: Sage, 1998.
- Brown, Michael K. *Working the Street: Police Discretion and the Dilemmas of Reform*. New York: Russell Sage, 1981.
- Charles, Michael T. *Police Training All the Rules: Implementing the Adult Education Model into Police Training*. Springfield, IL: Charles C. Thomas Publishing, 2000.
- Cohen, Howard S., and Michael Feldberg. *Power and Restraint: The Moral Dimensions of Police Work*. New York: Praeger, 1991.
- Christopher, Steve. "The Police Service Can Be a Critical Reflective Practice . . . If it Wants." *Policing: A Journal of Policy and Practice* 9, no. 4 (2015): 326–39.
- Dagan, David. "The Baltimore Cops Studying Plato and James Baldwin." *The Atlantic*, November 25, 2017.
- Davis, Kenneth C. *Discretionary Justice: A Preliminary Inquiry*. Baton Rouge, LA: Louisiana State University Press, 1969.
- . *Police Discretion*. St. Paul, MN: West Publishing, 1975.
- Deterding, Nicole M., and Mary C. Waters. "Flexible Coding of In-Depth Interviews: A Twenty-First-Century Approach." *Sociological Methods & Research*, 2018. <https://doi.org/10.1177/0049124118799377>.
- Dewey, John. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Boston: D.C. Heath & Co Publishers, 1933.
- Engel, Robin S. "The Effects of Supervisory Styles on Patrol Officer Behavior." *Police Quarterly* 3, no. 3 (2000): 262–93.
- Engel, Robin S., and Robert E. Worden. "Police Officers' Attitudes, Behavior, and Supervisory Influences: An Analysis of Problem Solving." *Criminology* 41, no. 1 (2003): 131–66.
- Ericson, Richard V. "Rules in Policing." *Theoretical Criminology* 11, no. 3 (2007): 367–401.
- Friedman, Barry. *Unwarranted: Policing Without Permission*. New York: Farrar, Straus, and Giroux, 2017.
- Galanter, Marc. "Why the 'Haves' Come Out Ahead: Speculations on the Limits of Legal Change." *Law & Society Review* 9, no. 1 (1974): 95–160.
- Goldstein, Herman. "Police Policy Formulation: A Proposal for Improving Police Performance." *Michigan Law Review* 65 (1967): 1123–46.
- . "Improving Policing: A Problem-Oriented Approach." *Crime and Delinquency* 25, no. 2 (1979): 236–58.
- Gould, Jon B., and Stephen D. Mastrofski. "Suspect Searches: Assessing Police Behavior Under the Constitution." *Criminology & Public Policy* 3, no. 3 (2004): 315–62.
- Gouldner, Alvin. *Patterns of Industrial Bureaucracy*. New York: The Free Press, 1954.
- Haidt, J. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgments." *Psychological Science Review* 108, no. 4 (2001): 814–34.
- Harmon, Rachel A. "Why Arrest?" *Michigan Law Review* 115, no. 3 (2016): 307–64.
- Hart, H.L.M. "The Aims of the Criminal Law." *Law & Contemporary Problems* 23 (Summer) (1958): 401–41.
- Kahneman, Daniel, and Gary Klein. "Conditions for Intuitive Expertise." *American Psychologist* 64, no. 6 (2009): 515–26.

- Kelling, George L. *Broken Windows and Police Discretion*. Washington, DC: U.S. Department of Justice, National Institute of Justice, 1999.
- Klein, Gary. *Streetlights and Shadows: Searching for the Keys to Adaptive Decision-making*. Cambridge, MA: MIT Press, 2011.
- Koen, Marthinus C., Willis, James J., and Stephen D. Mastrofski. "The Effects of Body-Worn Cameras on Police Organization and Practice: A Theory-Based Analysis." *Policing and Society* 29, no. 8 (2018): 968–94.
- Koen, Marthinus C., and Brooke Mathna. "Body-Worn Cameras and Internal Accountability at a Police Agency." *American Journal of Qualitative Research* 3, no. 2 (2019): 1–22.
- Krantz, Sheldon. *Police Policymaking: The Boston Experience*. Lexington, MA: Lexington Books, 1979.
- Kritzer, Herbert M. "Toward a Theorization of Craft." *Journal of Social & Legal Studies* 16, no. 3 (2007): 321–40.
- Lum, Cynthia. *Translating Police Research into Practice*. Washington, DC: The Police Foundation, 2009.
- Lum, Cynthia, Christopher S. Koper, David B. Wilson, Megan Stoltz, Michael Goodier, Elizabeth Eggins, Angela Higginson, and Lorraine Mazerolle. "Body-Worn Cameras' Effects on Police Officers and Citizen Behavior: A Systematic Review." *Campbell Systematic Reviews* 16, no. 3 (2020). <https://onlinelibrary.wiley.com/doi/10.1002/cl2.1112>.
- Lvovsky, Anna. "Rethinking Police Expertise." *Yale Law Journal* 131, no. 2 (2021): 475–572.
- Mangels, Laura, Joel Suss, and Brian Lande. "Police Expertise and Use of Force: Using a Mixed-Methods Approach to Model Expert and Novice Use-of-Force Decision-Making." *Journal of Police and Criminal Psychology* 35, no. 3 (2020): 294–303.
- Mastrofski, Stephen D. "Community Policing as Reform: A Cautionary Tale." In *Community Policing: Rhetoric or Reality?* Edited by Jack R. Greene and Stephen D. Mastrofski, 47–67. New York: Praeger, 1988.
- . "Measuring Police Performance in Public Encounters." In *Quantifying Quality in Policing*. Edited by Larry T. Hoover, 207–41. Washington DC: Police Executive Form, 1996.
- . *Policing for People. Ideas in American policing*. Washington, DC: Police Foundation, 1999.
- . "The Police in America." In *Criminology: A Contemporary Handbook*. Edited by Joseph F. Sheley, 405–43. Belmont, CA: Wadsworth/Thomson Learning, 2000.
- . "Controlling Street-Level Police Discretion." *The ANNALS of the American Academy of Political and Social Science* 593 (2004): 100–18.
- . *Do the Right Thing: Evaluating Police Performance at the Street Level*. Inaugural Stephen D. Mastrofski Lecture, March 19. Fairfax, VA: George Mason University, 2018.
- Mastrofski, Stephen D., and R. Richard Ritti. "Making Sense of Community Policing: A Theory-Based Analysis." *Police Practice and Research* 1, no. 2 (2000): 183–210.
- Mastrofski, Stephen D., R. Richard Ritti, and Debra Hoffmaster. "Organizational Determinants of Police Discretion: The Case of Drinking-Driving." *Journal of Criminal Justice* 15, no. 5 (1987): 387–402.
- Mastrofski, Stephen D., Jeffrey B. Snipes, Roger B. Parks, and Christopher D. Maxwell. "The Helping Hand of the Law: Police Control of Citizens on Request." *Criminology* 38, no. 2 (2000): 307–42.
- Mastrofski, Stephen D., Tal Jonathan-Zamir, Shomron Moyal, and James J. Willis. "Predicting Procedural Justice in Police-Citizen Encounters." *Criminal Justice & Behavior* 43, no. 1 (2016): 119–39.
- Meares, Tracey, Tom R. Tyler, and Jacob Gardener. "Lawful or Fair? How Cops and Laypeople Perceive Good Policing." *Journal of Criminal Law & Criminology* 105, no. 2 (2015): 297–344.
- Moore, Mark H. "Learning While Doing: Linking Knowledge to Policy in the Development of Community Policing and Violence Prevention in the United States." In *Integrating Crime Prevention Strategies: Propensity and Opportunity*. Edited by Per-Olof Wikstrom, Ronald V. Clarke, and Joan McCord, 301–33. Stockholm, Sweden: The National Council for Crime Prevention, 1995.
- Muir, William Ker, jr. *Police: Streetcorner Politicians*. Chicago: University of Chicago Press, 1977.
- Nix, Justin, Natalie Todak, and Brandon Tregle. "Understanding Body-Worn Camera Diffusion in U.S. Policing." *Police Quarterly* 23, no. 3 (2020): 396–422.

- Owens, Emily. "How Do You Build a Better Cop? By Making Them Slow Down." *The Washington Post*, June 2, 2020. <https://www.washingtonpost.com/outlook/2020/06/02/police-misconduct-reform-training-floyd-protests>.
- Paoline, Eugene. A. "Shedding Light on Police Culture: An Examination of Officers' Occupational Attitudes." *Police Quarterly* 7, no. 2 (2004): 205–36.
- Patton, Michael Q. *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: Sage, 2002.
- Phelps, Joshua M., Jon Strype, Sophie Le Bellu, Saadi Lahlou, and Jan Aandal. "Experiential Learning and Simulation-Based Training in Norwegian Police Education: Examining Body-Worn Video as a Tool to Encourage Reflection." *Policing: A Journal of Policy and Practice* 12, no. 1 (2018): 50–65.
- Ponomarenko, Maria. "Rethinking Police Rulemaking." *Northwestern University Law Review* 144, no. 1 (2019): 1–64.
- President's Task Force on 21st Century Policing. *Final Report of the President's Task Force on 21st Century Policing*. Washington, DC: Office of Community Oriented Policing Services, 2015.
- Reiss, Albert J., jr. *The Police and the Public*. New Haven, CT: Yale University Press, 1971.
- Schön, Donald. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books, 1983.
- . *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. San Francisco, CA: Jossey-Bass Publishers, 1987.
- Searcey, Dionne. "What Would Efforts to Defund or Disband Police Departments Really Mean?" *New York Times*, June 8, 2020. <https://www.nytimes.com/2020/06/08/us/what-does-defund-police-mean.html>.
- Sennett, Richard. *The Craftsman*. New Haven, CT: Yale University Press, 2008.
- Sklansky, David A., and Monique Marks. "The Role of the Rank and File in Police Reform." *Policing and Society* 18, no. 1 (2008): 1–6.
- Skolnick, J. *Justice Without Trial*. New York: Wiley, 1966.
- Slobogin, Christopher. "Policing as Administration." *University of Pennsylvania Law Review* 165, no. 1 (2016): 91–152.
- Thacher, David. "Policing is Not a Treatment." *Journal of Research in Crime & Delinquency* 38, no. 4 (2001): 387–415.
- . "The Normative Case Study." *American Journal of Sociology* 111, no. 6 (2006): 1631–76.
- . "Research for the Front Lines." *Policing and Society* 18, no. 1 (2008): 46–59.
- . "Channeling Police Discretion: The Hidden Potential of Focused Deterrence." *University of Chicago Legal Forum Article* 13 (2016): 533–77.
- . "The Aspiration of Scientific Policing." *Law & Social Inquiry* 41, no. 1 (2019): 273–97.
- . "The Learning Model of Use-of-Force Reviews." *Law & Social Inquiry* 45, no. 3 (2020): 755–86.
- Tiftt, Larry. "Control Systems, Social Bases of Power and Power Exercise in Police Organizations." *Journal of Police Science and Administration* 3, no. 1 (1975): 66–76.
- The Police Tapes. 1977. Directed by Alan and Susan Raymond. Video Vérité.
- Toch, Hans. "Mobilizing Police Expertise." *The ANNALS of the American Academy of Political and Social Science* 452 (1980): 53–62.
- Tyler, Tom R. "Enhancing Police Legitimacy." *The ANNALS of the American Academy of Political and Social Science* 593 (2004): 84–99.
- Waddington, P.A.J., Kate Williams, Martin Wright, and Tim Newburn. "Dissension in Public Evaluations of the Police." *Policing and Society* 25, no. 2 (2015): 212–35.
- Walker, Samuel. *Taming the System: The Control of Discretion in Criminal Justice, 1950–1990*. Oxford, UK: Oxford University Press, 1993.
- Walker, Samuel E., and Carol A. Archbold. *The New World of Police Accountability*. Thousand Oaks, CA: Sage Publications, 2014.
- Walker, Samuel E. "Governing the American Police: Wrestling with the Problems of Democracy." *University of Chicago Legal Forum Article* 15 (2016): 615–60.
- White, Michael D., Michaela Flippin, & Aili Malm. "Key Trends in Body-Worn Camera Policy and Practice: A Four-Year Funded Policy Analysis of U.S. Department of Justice-Funded Law Enforcement Agencies." *Office of Justice Programs*, 2019. https://bwctta.com/sites/default/files/inline-files/BWC%20Policy%20Analysis%20-%204th%20edition%20FINAL.Dec_.2019.pdf.

- White, Michael D., and Aili Malm. *Cops, Cameras, and Crisis: The Potential and the Perils of Police Body-Worn Cameras*. New York: NYU Press, 2020.
- Willis, James J. *Improving Police: What's Craft Got to do With It?* Washington, DC: The Police Foundation, 2013.
- Willis, James J., Stephen D. Mastrofski, and David Weisburd. "Making Sense of Compstat: A Theory-Based Analysis of Organizational Change in Three Police Departments." *Law & Society Review* 41, no. 1 (2007): 147–88.
- Willis, James J., and Stephen D. Mastrofski. "Understanding the Culture of Craft: Lessons from Two Police Agencies." *Journal of Crime and Justice* 40, no. 1 (2017): 84–100.
- . "Improving Policing by Integrating Craft and Science: What Can Patrol Officers Teach Us About Good Police Work?" *Policing and Society* 28, no. 1 (2018): 27–44.
- Willis, James J. "Culture Eats Strategy for Breakfast": An In-Depth Examination of Police Officer Perceptions of Body-Worn Camera Implementation and Their Relationship to Policy, Supervision, and Training." *Criminology and Public Policy* (2022). <https://doi.org/10.1111/1745-9133.12591>.
- Wilson, James Q. *Varieties of Police Behavior*. Boston: Harvard University Press, 1968.
- Wolfe, Scott, Jeff Rojek, Kyle McLean, and Geoffrey Alpert. "Social Interaction Training to Reduce Police Use of Force." *The ANNALS of the American Academy of Political and Social Science* 687 (2020): 124–45.
- Worden, Robert E., and Caitlin J. Dole. "The Holy Grail of Democratic Policing." *Criminal Justice Ethics* 38, no. 1 (2019): 41–54.
- Yanow, Dvora, and Haridimos Tsoukas. "What Is Reflection-in-Action? A Phenomenological Account." *Journal of Management Studies* 46, no. 8 (2009): 1339–64.