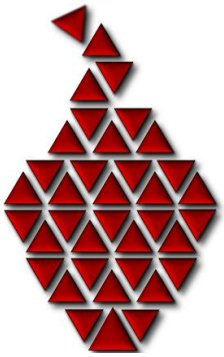

FINN

The John F. Finn Institute
for Public Safety, Inc.

Police Early Intervention Systems: The State of the Art

Robert E. Worden

Sarah J. McLean

Kelly J. Becker

Eugene A. Paoline, III

Hannah Cochran

Chris Harris

Summary Report to the National Institute of Justice

December, 2018

This research was supported by Award No. 2013-IJ-CX-0018, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice, to the John F. Finn Institute for Public Safety, Inc. The opinions, findings, and conclusions are those of the authors and do not necessarily reflect those of the Department of Justice.

421 New Karner Road
Suite 12
Albany, NY 12205
518.456.6323

The John F. Finn Institute for Public Safety, Inc., is an independent, not-for-profit and non-partisan corporation, whose work is dedicated to the development of criminal justice strategies, programs, and practices that are effective, lawful, and procedurally fair, through the application of social science findings and methods. The Institute conducts social research on matters of public safety and security – crime, public disorder, and the management of criminal justice agencies and partnerships – in collaboration with municipal, county, state, and federal criminal justice agencies, and for their direct benefit. The findings of the Institute's research are also disseminated through other media to criminal justice professionals, academicians, elected public officials, and other interested parties, so that those findings may contribute to a broader body of knowledge about criminal justice and to the practical application of those findings in other settings.

The Finn Institute was established in 2007, building on a set of collaborative projects and relationships with criminal justice agencies dating to 1998. The first of those projects, for which we partnered with the Albany Police Department (APD), was initiated by John Finn, who was at that time the sergeant who commanded the APD's Juvenile Unit. Later promoted to lieutenant and assigned to the department's Administrative Services Bureau, he spearheaded efforts to implement problem-oriented policing, and to develop an institutional capability for analysis that would support problem-solving. The APD's capacity for applying social science methods and results thereupon expanded exponentially, based on Lt. Finn's appreciation for the value of research, his keen aptitude for analysis, and his vision of policing, which entailed the formulation of proactive, data-driven, and – as needed – unconventional strategies to address problems of public safety. Lt. Finn was fatally shot in the line of duty in 2003. The Institute that bears his name honors his life and career by fostering the more effective use of research and analysis within criminal justice agencies, just as Lt. Finn did in the APD.

Contents

INTRODUCTION

The misuse of police authority and other misconduct is by many measures concentrated disproportionately among small fractions of officers. Perhaps the best-known example is the group of 44 "problem officers" in the Los Angeles Police Department (LAPD), who were featured in the Christopher Commission report in 1991 (Independent Commission on the Los Angeles Police Department, 1991: 36-37). Indicators of "risk-related" outcomes (Bobb, 2009), such as complaints and the use of force, follow a similarly skewed distribution across sworn personnel in other agencies as well. We should be circumspect in the inferences that we draw from these data, since complaints and use of force incidents are not valid measures of misconduct as such. But the 80-20 rule appears to have wide applicability in organizational settings, holding that 80 percent of the problems stem from 20 percent of the employees, and so it is quite plausible that a large fraction of police misconduct stems from a small fraction of officers. Efforts to identify and intervene with this small set of officers would therefore seem to hold a great deal of promise for preventing misconduct.

This logic underlies the adoption of early intervention (or early warning) systems by police. Early intervention (EI) systems are administrative systems designed to monitor indicators of individual officers' behavior, identify officers in whose behavior patterns of deficient performance appear, and apply interventions that enable the officers to improve. EI systems and the interventions they provide are generally conceived as non-punitive, operating independently of disciplinary processes. Indeed, part of the rationale for early intervention is that performance problems among police may manifest themselves in forms that do not permit the application of disciplinary sanctions – e.g., in unsustained personnel complaints, or uses of force that are within agency policy. Rather than sanctioning officers, EI systems provide for retraining, counseling, or other corrective measures.

EI systems share some basic structural commonalities among the many agencies that have adopted them. Generally EI systems consist of four components:

- performance indicators, such as citizen complaints or use of force reports;
- procedures for the identification or selection of officers who are exhibiting signs of problematic behavior;
- intervention(s); and
- post-intervention monitoring.

Thus an early intervention system (EIS) includes both components that are or can be mechanized, such as the information sources on officers' behavior; human components, such as the interventions; and potentially blended components comprised of both mechanical and human elements, such as procedures for identifying officers whose behavior warrants intervention. Beyond these basic structural similarities, extant research offers hints of structural variation among EI systems which may have important implications for the systems' operation and effectiveness.

The fulfillment of the promise of EI systems remains to be confirmed by systematic research, as few evaluations have been conducted and much of the empirical evidence from those studies rests on weak research designs. We discuss the evidence and its limitations below. The goal of this project was to augment the body of research on the operation and impacts of EI

systems, with particular attention to the implications of EIS structure. Our objectives included a description of the respects in which EIS structure varies, and an intensive examination of EIS processes and outcomes in a small number of agencies whose EI systems varied in key structural respects. We summarize our findings in this report. We first review and critique previous research on EI systems, and we summarize the contemporary questions about EI systems that remain unanswered. We then provide an overview of the project's methodology, including the survey of agencies and the case studies of individual agencies. We proceed to detail the structure of EI systems nationally, drawing from the survey of agencies. We next present findings from the case study of individual agencies beginning with process study examining how each of the EI systems function, and then turn to the outcome analysis and estimates of the magnitude of the impacts of the interventions on the performance of officers who are subjected to intervention. Following the presentation of brief, "snapshot" descriptions of each of eight agencies' EI systems, we conclude with a summary of what we found, overall, and discuss the implications of our findings.

EXTANT EVIDENCE

The first study of EI systems, which included a national survey of 571 agencies in 1998-99, found that about one-quarter of local police agencies serving populations of 50,000 or more had some version of an EI system, and an additional 12 percent were planning to institute such systems (Walker, et al., 2000). EI systems were more prevalent among municipal departments (33 percent) than sheriff's offices (16 percent). They were also more common among large agencies than smaller agencies: 62 percent of agencies with 1,000 or more sworn personnel had an EIS, compared with 36 percent of agencies with 500-999 sworn personnel, and only 20 percent of agencies with fewer than 300 sworn. EI systems subsequently became a key element of Justice Department consent decrees, further legitimating them as a "best practice" (Harmon, 2009; Walker, 2003a). By 2007, according to our analysis of data collected in the Law Enforcement Management and Administrative Statistics (LEMAS) survey, 40 percent of the agencies serving populations of 50,000 or more reported that they had an EIS.[1] As Walker et al. (2000) had previously found, EI systems were more common among municipal police departments and among larger agencies. EI systems have become a common feature of the organizational landscape in law enforcement.

EI System Structure

EI systems vary with respect to each component: performance indicators, identification procedures, interventions, and post-intervention monitoring. Walker et al. (2000: 4.1-4.5) found that very few departments relied exclusively on citizen complaints (1.4 percent). Instead, agencies used a variety of indicators in conjunction with citizen complaints, such as use of force (most prevalent, although no percentages noted), high speed pursuits (approximately one half), civil litigation (approximately one-third), and vehicle damage (approximately one-third). Most systems were simple "time-and-numbers" systems that provided for numerical thresholds of

[1] In the 2007 LEMAS data, 778 agencies answered the long form, which included the item about EI systems. Another 215 agencies (two-thirds of which were sheriff's offices) answered the short form, so we do not know about their use of EI systems.

specific events over specified time frames – e.g., three complaints over a twelve-month period – as the criterion by which officers were initially identified for intervention. Identification thresholds varied, though the modal criterion was three complaints or three use of force reports in a twelve-month period. Officers' immediate supervisors were commonly responsible for delivering or coordinating the intervention (62 percent), although other personnel were charged with counseling officers (45 percent) and training (45 percent). Finally, follow-up monitoring was conducted by 90 percent of responding agencies, though the duration varied: e.g., 40 percent monitored for 12 months, 47 percent monitored for 36 months.

The accuracy with which identification procedures direct attention to officers whose performance is genuinely problematic has emerged as one key issue. Identification is an exercise in prediction: in the absence of intervention, an officer will subsequently display a pattern of behavior that constitutes a problem. No consensus exists on the proper thresholds, and the validity of the predictions on which these systems are implicitly based has not been established by empirical research.[2]

One would be appropriately skeptical of such predictions. First, phenomena that occur with such a low frequency are notoriously difficult to predict with a high level of accuracy (see, e.g., Gottfredson 1987). Second, the data on which the predictions are based are at best ambiguous as indicators of police misconduct. Even when it is skillfully performed, police work involves the use of force on occasion, and citizen complaints could be expected to sometimes arise even when police officers do their jobs properly. Third, widely used "time-and-numbers" systems typically do not take into account other factors, such as the nature of officers' assignments, which may affect their exposure to situations that are more or less likely to necessitate the use of force or to yield citizen complaints. While it is probably true that the best prediction of behavior during any given time period is based on behavior during the immediately preceding time period, the predictive mechanisms on which most EI systems are based have not been validated by empirical research and likely leave some room – much room, it appears – for improvement.

Any prediction is subject to errors of one type or another: false positives and false negatives. False negatives – in this instance, officers who are assessed as lower-risk but whose performance later proves to be problematic – are a major concern for obvious reasons: they are missed opportunities to prevent misconduct for the benefit of both officers and citizens. False positives – officers who are assessed as high-risk but whose performance would be satisfactory in the absence of intervention – are also consequential. First, resources are wasted insofar as interventions are applied to personnel who do not need them. Second, officers subject to an EIS intervention may suffer an informal stigma in their selection, and so the officers who are false positives carry the burden of such a stigma unnecessarily. Third, and perhaps most importantly, a moderate to high rate of false positives is likely to reduce the legitimacy of an EIS in the eyes of line officers and the immediate supervisors who are responsible for them. If officers see peers or subordinates who, in their estimation, are solid performers and who are nevertheless selected for EIS intervention, then they are likely to see the system as broken, and that would adversely

[2] Walker (2003b, p. 4) argues that EI systems "do not attempt to predict officer performance based on background characteristics or other factors, rather they indicate that current performance levels, while not warranting disciplinary action, still warrant improvement." But the prediction on which EIS intervention is based – that recent unsatisfactory performance will continue without intervention – is unmistakable.

affect the potential of such a system for organizational reform.[3] Thus EI systems that rely on such crude selection criteria may be needlessly costly.

A few departments, such as Pittsburgh's (Davis, et al., 2002), use peer comparisons, which has the potential to adjust for officers' exposure to risk. However, it is surely difficult to define the peer groups for patrol officers properly, since exposure varies both spatially, across beats (even within precincts or districts) and temporally, across shifts or platoons. Some other departments, such as LAPD (Office of the Inspector General, 2014), specify ratios of risk-related outputs (e.g., uses of force) to activity indicators (e.g., arrests), which might better take account of variation in exposure. But not many EI systems use either of these methods, as we discuss below, and we do not know how well these forms of identification perform in making predictions.

A number of agencies provide for what Sam Walker described as a second-stage of identification, in which officers who reach EIS thresholds are reviewed by, e.g., an immediate supervisor or an EIS coordinator, to arrive at a judgment about whether and, if so, what kind of intervention is needed. Thus these EI systems do not rely entirely on the numerical criteria. One danger of such an arrangement is that, insofar as supervisors identify with their subordinates and work to "protect" them from administrative scrutiny, they may declare that a problem officer is not a problem. However, given the shortcomings of the information on which initial identification turns and the risk of a large number of false positives, some kind of second-stage review is desirable. We find some empirical support for this proposition in the reported experiences of agencies that, in effect, treat EI system indicators as screens with supervisory review as a further test of performance problems. For example, over a nearly three-year period beginning in 2009, Denver police officers reached "review limits" – 3 citizen complaints or uses of force in a quarter, 5 in 6 months, 6 citizen complaints or 7 uses of force in a year – 232 times (see Osher, 2011). They were reviewed by a board that includes managers, supervisors, officers, union officials, and one citizen. The board's review of these cases resulted in informal notification of the officer's commander 84 times, and a notification of both the officer and the officer's commander that required an informal response 28 times; in only 2 cases the board required a formal plan for remediation. In half (118) of the 232 cases, the board determined that no action was warranted.

In another example, the LAPD's EI system automatically generates an "action item" when an officer reaches a threshold. The action item requires a supervisor to determine whether corrective action is appropriate and, if so, what that action should be. Supervisors' recommendations are subject to review through the chain-of-command. The Office of the Inspector General for the Los Angeles Police Commission (2014) performed a review of the EIS, including an analysis of action items generated during a four-month period, from November of 2012 through February of 2013. Of the 748 action items generated, the dispositions of 70 percent were "no action": "no pattern of behavior posing potential risk was identified."

A further limitation of the predictions made in contemporary EI systems is their narrow focus on *recent* past behavior, i.e., the previous year or even the previous three months. Even if the indicators contained less "noise" (i.e., extraneous information) as measures of police misconduct, we would be well-advised not to discard information that may (and probably does)

[3] If officers, and the unions that represent them, are skeptical of EI systems, the process whereby officers are selected for intervention might be one very good reason. There are other reasons, to be sure: insofar as the intervention is unwelcome, and insofar as it rests on no demonstrable violation of policy, it could be reasonably perceived as an adverse action that is unjustified. Thus it is all the more important that selection rest on firm ground.

have some predictive value. Given the random elements that are intrinsic parts of indicators, we might hope that over a longer period of time, they would cancel one another out. Just as criminal histories-including offenders' age at first arrest- have predictive value in assessing the risk of future offending, so might officers' histories have value in predicting their future performance.

EI System Effectiveness

Only six agencies' EI systems have been evaluated: five in the U.S. (Bobb, 2009; Walker et al., 2000; Worden et al., 2013) and one in Australia (Macintyre et al., 2008). Despite differences in the systems' processes for identifying and selecting officers with performance problems and the interventions employed, all but one of the evaluations report positive results from EI system implementations, particularly decreases in the number and/or rate of complaints. However, the positive evaluations are of questionable internal validity: most are one-group, pretest/posttest designs, and thus are vulnerable to threats to internal validity such as maturation, regression toward the mean, and history. The stronger evaluation is also one that produced unfavorable results – no impacts on complaints, use of force, or secondary arrests. Furthermore, this evaluation found that the EI system intervention had an inhibiting effect on enforcement activity. EI systems were not intended to inhibit (good) police work and public engagement.

Maturation is a threat to internal validity that stems from several time-related patterns. Research shows that officers' activity levels decline over time with more experience; veteran officers make fewer stops and arrests when compared to less experienced officers (Crank, 1993; Friedrich 1977; Van Maanen, 1974; Worden, 1989). With declining activity, officers have fewer opportunities to engage in misconduct or otherwise generate "risk-related" outcomes. One might also expect that, as officers gain experience, they become more skilled. Learning from their peers or supervisors or from trial and error, officers might less frequently act in ways that produce complaints or attempt to resolve situations by resorting to the use of force. Thus we would expect complaints to decline across the course of officers' careers, and this expectation is confirmed by empirical analysis (Harris, 2009). These declines could appear as lower post-intervention rates of complaints or use of force, even if the intervention has no effect—a textbook example of the threat of maturation.

Insofar as all event counts fluctuate over time, we might expect that at least some (perhaps many) officers are identified by EI systems shortly after an abnormal spike in an indicator, and the subsequent regression to the mean could be mistaken for an intervention effect. Analyzing several indicators for all sworn members of one large police agency, Worden, et al. (2003, 2012) found that each year, complaints and other indicators displayed the same kind of skewed distribution found in other agencies because a small number of officers accounted for a large fraction of the events. But for the most part, it was not the same officers in the tail of the distribution from year to year, and officers who appeared to be documentable "problems" at any one point in time tended not to exhibit them later, as indicator counts regressed toward the mean.

Finally, reforms other than the introduction of an EIS could contribute to a decline in the incidence of police misconduct. To the extent that other reforms coincide with the adoption of an EIS, these effects could be mistaken for an effect of the system's intervention.

Each of these threats to internal validity tempers the conclusions that can be drawn from most previous studies of EIS interventions. For example, the New Orleans Police Department established the Professional Performance Enhancement Program (PPEP) in the 1990s (sometime after 1995, Walker, et al., 2000: 3.24) along with other reforms. Selection for the PPEP was based on (1) incidents involving conflict in arrest situations; (2) conflict in non-arrest situations,

with 3 or more abuse complaints in 24 months; and (3) supervisory referrals. The intervention was a 4-day training session, attended by as many as 21 or as few as 6 officers (Walker, et al., 2000: 4.54). This training included several components, such as stress and conflict management, complaint avoidance, verbal judo, and sensitivity training. PPEP also required that supervisors complete evaluations of the selected officers who were under their supervision for six months. PPEP impacts were assessed only in terms of complaints – two years before and after intervention – because no other post-intervention performance data were compiled. Of 74 officers who had been subject to the PPEP intervention, analysis was confined to 27 for whom the complete four years of data were available, and who had not resigned or been dismissed. The aggregate complaints for these officers declined from 100 pre-intervention (a rate of 1.85[4] per year) to 34 post-intervention (a rate of 0.63). The findings of this and most other previous evaluations are subject to the three threats to internal validity discussed above: maturation; regression toward the mean; and history. The decline in aggregate complaints in the two-year period following intervention could be attributed to maturation (that is, officers became more skillful and/or less active over time even without an intervention) or regression to the mean (since officers selected for intervention presumably had higher rates of problem behaviors) and not to the 4-day training session. These threats to internal validity cannot be eliminated with reference to the changes in a control group, as no comparison group was established, and no account was taken of officers' exposure to risk of complaints. That the EIS was adopted along with several other reforms also makes history a plausible explanation for the study's results. (For detailed critiques of the other evaluations, see Worden, et al., 2013.)

The strength of the current research on EI system interventions is, in our estimation, no greater than a 3 on the 5-point Maryland Scientific Methods Scale (see Sherman & Gottfredson, 1997), with methodology that is rigorous in some respects and weak in others. All but one of the studies provide for either no control group or a control group that differs from the treatment group in key respects (e.g., an entire cohort of officers who were not subject to EIS intervention), with no use of control variables to adjust for those differences, and sometimes no accounting of attrition. While randomized experiments are probably not feasible, stronger designs—those that take account of temporal patterns (and not only pre- and post-intervention means) for both a treatment group and a control group, with controls for factors other than the intervention—could be reasonably achieved.[5]

Moreover, the outcomes examined in existing research have been limited to complaints and use of force, with one exception (Worden et al., 2013) that assessed the effects of EI systems on officers' activity levels. Yet there is good reason to believe that complaints and uses of force turn substantially on officers' degree of exposure to risk: the volume of calls for service that they handle and the number of contacts with citizens that they initiate (see Brandl, et al., 2001; Hassell & Archbold, 2010; Lersch, 2002; Lersch, et al., 2006). Officer-initiated contacts, particularly traffic or pedestrian stops, are "obligation encounters" (see Moore, 2002: 37) that are more adversarial in nature and thus (presumably) more likely to result in a complaint or a use of force. If EI systems reduce complaints and the use of force by deterring officers from taking this kind of initiative, with corresponding reductions in other outcomes such as searches, tickets, and

[4] Walker, et al. report a rate of 1.66 (2000: p. 4.52) but our calculations yield a rate of 1.85.

[5] Length of service, sex, and race are all correlated with one or more "risk-related" indicators, and several other characteristics appear to be risk factors for misconduct (see Cohen & Chaiken, 1972; Greene, et al., 2004; and Fyfe & Kane, 2005).

arrests, then we would have to question *how* EI systems are succeeding, and more importantly, at what cost.

To date, the only evaluation that formed a plausible control group and also examines the possible inhibiting effect of an EI system's intervention was recently completed by Worden et al. (2013). This study evaluated a four-day training course developed and implemented by a large agency in the northeastern U.S. That agency's EIS was first put into practice in 1992, and through mid-2000, 11 classes of 10-12 officers each (118 in all) had been selected and trained. The authors analyzed data on all personnel complaints (1987-2001), uses of force (1995-2001), secondary arrests (1987-2001), and all primary arrests (1987-2001), including a subset of the latter that were presumptively proactive. They formed a control group, matching each trainee by sex, race, academy class, and the number of pre-intervention personnel complaints to an officer who was not selected for EIS intervention. They compared pre-intervention levels and rates of the outcomes to post-intervention levels and rates, and also performed a panel analysis. The study's results indicate that levels and rates of risk-related outputs declined in the post-intervention period, but the declines were no greater for trainees than they were for the matched controls; the intervention did not have the expected beneficial impacts. However, the intervention had a modest inhibiting effect on officers' enforcement activity: while arrests and proactive arrests declined for both groups, they declined *more* for the trainees.

These findings are credible but hardly definitive, begging the question about their external validity. They underscore not only the need for more research on EI systems, but also the need for research that uses strong designs. The issues are very important, given the centrality of EI systems in agendas for police reform, and social science could generate empirical evidence whose credibility is more nearly commensurate with the significance of the issues. Even if further research replicates Worden, et al.'s findings, we do not believe that it would imply that EI systems are a fatally flawed concept. It would imply that we need to better understand the etiology of problem behavior, how problem behavior can be predicted (see Worden, Harris, and McLean, 2014; Worden and McLean, 2016), how interventions are applied, and how interventions affect and are experienced by the officers who are subjected to them.

EI System Context

We take as a premise that one cannot understand an EI system in isolation from its organizational context. Patterns of supervision are an important feature of that context. Good supervisors do a lot of their own early intervention without the benefit of an EIS apparatus, and it is widely thought that an EIS can backstop or even be an impetus for supervision (Walker, Milligan and Berke; 2006). Insofar as supervisors routinely monitor their subordinates' performance and play a role of mentor or "coach," they may succeed in "early early" intervention, thereby minimizing the frequency with which more formal EIS intervention is applied and/or the severity of the problems to be addressed. It is also possible that an EIS would generate gratuitous work for supervisors, particularly if it generates a large number of false positives, and could even detract from supervision if supervisory time is squandered as a result.

EI System Purposes

EI systems proliferated through the 1990s and 2000s mainly as a tool with which to prevent misconduct. Citizen oversight bodies and police "watchdog" groups expect and lobby for their use, and as noted above, the Civil Rights Division of the U.S. Department of Justice has treated the establishment or strengthening of an EIS as a staple in its mandates or

recommendations to police agencies investigated for a pattern or practice of civil rights violations (Harmon, 2009; Helsby et al., 2017; Walker and Archbold, 2014; Walker, 2003b). The object is to identify "problem" officers – officers whose performance includes a *pattern* of problem behavior –so that corrective interventions can be applied. It bears repeating that the object is not punishment, and such systems can and should be constructive for the officers, their agencies, and the communities they serve.

More recently, with a view toward improving the accuracy with which numerical identification criteria point to officers who warrant intervention, the Charlotte-Mecklenburg Police Department (CMPD) and data scientists from the University of Chicago partnered to apply more sophisticated predictive modeling to the actuarial portion of the identification process (Helsby, et al., 2017). In so doing, they may have subtly reoriented the EIS toward individual events rather than *patterns* of problem behavior. To perform the modeling, a criterion variable was necessary, and in the absence of a valid and reliable measure of behavior patterns, the analysis focused on discrete "adverse events" – that is, complaints, uses of force, crashes, etc., that were confirmed violations of agency policy. Consequently, the predictions point to officers at risk of an adverse event. Whether these same officers are the officers who have identifiable patterns of problem behavior is an empirical question, but the overlap of these two groups of officers could be incomplete in one or both of two respects. First, some of the officers at risk of an adverse event – due to their assignments – may not be officers who exhibit a pattern of problem behavior. Indeed, some of the predictors reflect the degree of officers' exposure. Second, some of the officers whose performance includes a pattern of problem behavior may not be at risk of an adverse event; part of the original rationale for EI systems, readers may recall, was to direct attention to officers whose conduct may escape discipline. The refashioned system in CMPD is hitting a target with maximal accuracy, but it might not be the right target.

Further, some in law enforcement have expressed an interest in better putting EI systems to work in promoting officers' health and wellness. EI systems have historically performed this function to some extent: insofar as officers' performance problems stem from extraordinary stress or a temporary inability to cope with stress, or from substance abuse or other behavioral health issues, EI systems would (ideally) deliver interventions that promote the officers' health and wellness. However, contemporary EI systems were not designed to identify any officer whose health and wellness would properly be of concern.

The purposes of EI systems have evolved, though the shifts may not be recognized (as in CMPD) or realized (as with the aspiration to better serve officers' health and wellness). Whether any one system can serve these multiple purposes is an open question.

Research Questions

Extant research on EI systems leave many questions unanswered. One major question is whether and to what extent they are effective in preventing police misconduct. A number of more specific questions can be posed about particular components of EI systems, answers to which might enable us to redesign systems that are (more) effective. Identification procedures are crucial pieces of EI systems, and they vary from one agency to the next. Some may be more successful in minimizing false positives and/or false negatives. Indicators, time periods, and thresholds all vary, and so do the personnel responsible for monitoring and reviewing performance indicators. The range of interventions available in an agency varies, and if EI systems are to be maximally effective, interventions will be matched to the conditions giving rise

to officers' performance problems. Supervisors must be prepared to ascertain the nature and origins of officers' performance issues and arrange for appropriate interventions. The interventions must be implemented faithfully. To ensure that the interventions were successful in individual cases, post-intervention assessment is desirable; the role of command staff – in the officer's chain of command and/or in professional standards – may be instrumental in that assessment.

THE STATE OF THE ART: PROJECT OVERVIEW

The "state of the art" project consisted of two phases: a survey of police agencies; and case studies of individual agencies' EI systems.

Agency Survey

In the first phase, we surveyed all of the 344 agencies that reported in the 2007 LEMAS survey that they had an "early warning system."[6] By the time of our survey, in 2014, these agencies' EI systems were presumably mature systems – systems that were well-established, having existed (if they continued to exist) for a long enough period to have undergone some fine-tuning and, moreover, for personnel to have become familiar with their operation. The primary objective of the survey was to learn about the details of their systems' structures:

- the number and types of indicators;
- the identification criteria, including the common "time-and-numbers" criteria as well as the use of other kinds of numerical criteria;
- whether identification for intervention was based partly on manual reviews of officers' performance by supervisors and/or others to assess whether a negative pattern of performance exists;
- the nature and range of interventions available;
- the length of post-intervention monitoring and the parties responsible for it.

The administration of the agency survey involved three waves of mailing hard copy questionnaires, and one wave of telephone calls (between waves 2 and 3) to those who did not return our survey. The initial wave involved mailing surveys to all 344 agencies, on May 30, 2014, of which 137 (39.8 percent) were completed and returned. On July 18, 2014, we sent a second wave of surveys to the 207 agencies that did not respond to the initial invitation to participate, of which 63 (30.4 percent) were completed and returned. From September 10, 2014 to October 23, 2014 we telephoned the 144 agencies that had yet to respond to our survey request. During this time period, 39 surveys were returned (27.1 percent). Our final wave of surveys were mailed on October 23, 2014, to the 105 non-responding agencies, of which 35 (33.3 percent) were completed and returned. Across the three waves of mail surveying (and our follow-up phone calls), 274 surveys were returned, for a response rate of 79.7 percent. In examining the overall response rate, across waves, nearly three-quarters (73 percent) of the 274 returned surveys occurred during the first two waves of administration.

[6] The item asked: "Does your agency have an operational computer-based personnel performance monitoring/ assessment system (e.g., Early Warning or Early Intervention System) for monitoring or responding to problematic officer behavior patterns?"

Table 1: Agency Survey Responses by Wave

Wave	Date	Surveys Mailed/Called	Number (%) Returned Per Wave	Response (%) of Overall Total (n=274)
1	May 30, 2014	344	137 (39.8)	50.0
2	July 18, 2014	207	63 (30.4)	23.0
3-calls	*Sept 10 – Oct 23, 2014*	*144*	*39 (27.1)*	*14.2*
4	October 23, 2014	105	35 (33.3)	12.8
Total			274 (79.7)	

Among the 274 responding agencies, 243 reportedly had an EI system at that time. These survey data enable us to describe the structural forms that mature EI systems take, and we report descriptive findings below.

Case Studies

The survey data also served a second purpose: they enabled us to identify among the respondents agencies that would be propitious sites for intensive case studies of process and, in a subset of those agencies, outcome evaluations. We were particularly interested in studying a set of agencies whose EI systems were different from one another in presumptively important structural respects. Among systems that used time-and-numbers identification thresholds, we sought to include an agency with fairly low thresholds and an agency with fairly high thresholds. We also sought to include an agency that incorporated peer comparisons into the identification algorithm. The durations of prescribed post-intervention monitoring varied, so we sought to select agencies that varied on that dimension, and also with respect to whether post-intervention monitoring was formally subject to review by command-level personnel. We were interested in studying agencies of varying size, but we ruled out very large agencies as organizations that we could not properly study in a few site visits. We concluded that we would include only one smaller agency (i.e., substantially less than 1,000 sworn) in the interests of statistical power. Since the number of agencies in which we could realistically conduct such intensive case studies was five to eight, we had to strike a balance among these objectives.-and of course the agencies would have to agree to participate. Given the focus on mature EI systems and the limitations of examining only a handful of them, we also undertook brief, "snapshot" descriptions of a number of other agencies whose EI systems are (or were) reputed to be either exemplary or cutting-edge systems.

Furthermore, since project resources would not permit a large volume of hand-coding data on EIS indicators, we restricted our attention to agencies that reportedly maintained computerized records of citizen complaints, use of force incidents, arrests, and either traffic stops or traffic citations. Citizen complaints and use of force incidents are, of course, common EIS indicators, and arrests and traffic stops/citations would afford us indicators of officers' exposure to risk. Because we planned to conduct outcome evaluations in as many of the agencies as possible, we looked especially for agencies whose electronic data were available for a substantial length of time – i.e., at least five years.

Ultimately, we selected and gained access to six agencies for the intensive case studies: five larger agencies, with 1,000 – 2,000 sworn, and one smaller agency, with under 300 sworn. We note that some of the agencies that we initially identified as meeting our criteria

subsequently proved to be unsuitable. The first agency we approached whose EIS used peer comparisons, for example, was interested in participating, but the EIS coordinator explained that the agency had very recently installed new EIS software and, as a consequence, the data in the software it replaced were inaccessible. Another agency that we approached was receptive to participation, but we discovered that – notwithstanding the agency's LEMAS responses about the use of computerized files – use of force data were very limited. Three agencies declined to participate. One or more of those may have proven unsuitable in the end, for one reason or another, but we have no way of knowing. Securing the agreement of the agencies to accommodate the project was a protracted process that began in late-2014 and ended only in late-2016.

In October of 2014, based on only the agency survey and LEMAS data and our selection criteria, we formulated a "short-list" of 13 agencies. The six agencies that served as case study sites include four of the agencies on the initial short-list; one of the six was initially excluded as too small. Five of the agencies on that short-list were never approached.

We conducted multiple site visits in each of the six case study sites. The number of visits per site ranged from 1 to 4. The duration of the visits ranged from 1 to 5 days. Table 2 below displays information on case study site visits.

Table 2: Case Study Site Visits

Site	Visits by Month and Year	Number of Interviews Conducted
Brewster	September 2015 June 2015 February 2018	6
Oriskany	February 2015 April 2015 June 2015 September 2015	9
Chester	August 2015 February 2016	11
Potsdam	December 2016	11
Sidney	February 2016 September 2016 June 2017	15
Caledonia	February 2017	10

During the site visits, we ascertained the availability of and arranged to obtain electronic data. In each site we also conducted interviews with executives and with others who make the EIS wheels turn. We spoke with: IA/professional standards personnel, EIS coordinators, precinct/district commanders, mid-level managers, and field supervisors. The number of interviews we conducted across the sites ranged from six to fifteen. For the most part, interviews were one-on-one and all were in person. Detail on the number of interviews per site is provided in Table 2 above.

The availability of computerized, event-level data on EIS indicators and on enforcement activity varied from agency to agency with the particulars of recording practices and data management.

We executed outcome evaluations in case study agencies that could make available data sufficient to analyze outcomes for officers subject to intervention and controls. The data order was a rather tall one, since in any one agency, fairly small numbers of officers are subject to intervention in any year, and we established two years as the minimum pre- and post-intervention time period over which to assess change. We considered five years as the minimally adequate span over which data on outcomes must be available to support an outcome analysis; two agencies could not provide data over so long a period of time.

We use pseudonyms for the agencies because we do not wish to penalize any of them for their agreement to participate in the study with unflattering findings. The identity of the individual agencies is not important, for each of them was selected as representing a type of EIS. None of the agencies requested anonymity, but we think it appropriate to grant it.

Finally, we selected several agencies' EI systems to describe less extensively – as "snapshots" of EI systems – and based only on publicly available information. Three of these agencies were selected because they were those on which the only previous national study dwelled: Miami-Dade; Minneapolis; and New Orleans. All three systems were favorably evaluated at that time. A fourth agency – the Los Angeles Sheriff's Department – was also the site of a positive evaluation. Baltimore's EIS was the subject of an article that appeared in The Police Chief in August of 2016. The Los Angeles Police Department was also selected, given that LAPD's EIS has been the subject of discussion in some professional circles and is also the subject of a recent evaluation. Oakland's EIS has occasionally been mentioned as exemplary, and so too has that of Charlotte-Mecklenburg.

EIS STRUCTURE: A NATIONAL PERSPECTIVE

The national survey of agencies was designed to provide a contemporary understanding of the structure and operation of EI systems. The last (and only) comprehensive picture of EI systems was drawn by Samuel Walker and his colleagues in a 1998-99 survey of 571 police agencies.[7] In the summer and early fall of 2014, we administered a survey to all agencies that participated in the 2007 Law Enforcement Management and Administrative Statistics (LEMAS) survey and reported that they utilized an EIS. Of the 883 agencies that answered the LEMAS question, 344 agencies (39 percent) reported that they had an EIS. Of these 344 agencies, 80 percent (N=274) responded to our survey, and among those 274 agencies, 243 reported they had an EIS in 2014.[8] The findings presented below are based on the 243 agencies operating these "mature" EI systems (that is, their system dates back to at least 2007).

Description of agencies

Of the 243 agencies whose systems we examined, 64 percent (N=155) had fewer than 500 sworn officers, 18 percent had a sworn force between 500 and 999, and the remaining 18 percent had a force of 1,000 or more. The majority of agencies (73 percent) were municipal police

[7] Samuel Walker, Geoffrey P. Alpert, and Dennis J. Kenney, *Responding to the Problem Police Officer: A National Study of Early Warning Systems*, Report to the National Institute of Justice (Washington: National Institute of Justice, 2000).

[8] Among the 31 agencies that reported in our survey that they did not have an EIS, 12 indicated that the agency's 2007 LEMAS response was erroneous. Five other agencies indicated that their EI systems were not computer-based. Four agencies reported that their EI systems were never implemented (e.g., existed on paper but not in practice, or in the form of a database only). Four agencies reported that their systems had been discontinued.

departments; other agencies included fifty sheriffs' offices and fifteen state police agencies. Refer to Table 3.

Table 3: Agencies by Size and Type

Agency size				Type of agency				
< 500	500 - 999	1000+		Municipal	Sheriff	State		Totals
(N=155)	(N=43)	(N=45)		(N=178)	(N=50)	(N=15)		243

We summarize the survey findings in terms of the four EIS components: (1) the performance metrics employed as potential indicators of misconduct; (2) the thresholds applied to identify officers with performance problems; (3) the interventions used to address those performance problems; and (4) post-intervention monitoring.[9]

Identification: Individual Performance Metrics

The survey provided a list of twenty-three indicators along with a space to write in additional indicators not listed; respondents were asked to designate whether or not each individual indicator was part of their EIS. For each identified indicator, respondents were also prompted to specify the period of time over which the indicator was computed and the threshold at or above which the computed value identified an officer's behavior as potentially problematic. In an EIS using a time-and-numbers algorithm, the computation was a simple count and the thresholds were integers. In an EIS using an alternative algorithm, such as ratios or peer comparisons, respondents were asked to specify the alternative formula. Table 4 below provides a listing of the individual performance indicators assessed by EI systems, broken down by agency size and also by type of agency.

The picture that emerges shows that a wide range of performance metrics are included in EI systems as indicators of potentially problematic police performance. The types of indicators assessed as a means to identify early signs of problems include those that might be indicative of problems in interactions with citizens (e.g. use of authority, complaints, discretionary arrests), abuse or misuse of personal and/or sick time, and failure to follow department rules and regulations (e.g. failure to complete trainings or annual physicals). The five most commonly considered indicators included:

- Citizen complaints
- Non-lethal force
- Internal complaints
- Vehicle damage/accidents
- Internal investigation

Fifty-four percent of responding agencies consider all five of the top five indicators and 46 percent of agencies include ten or more indicators on which selection for intervention could turn.

[9] Samuel Walker, *Supervision and Intervention within Early Intervention Systems: A Guide for Law Enforcement Chief Executives.* (Washington: Police Executive Research Forum, 2005).

Table 4: Indicators by Agency Size and Agency Type

	Agency size				Type of agency				
Indicator	< 500	500-999	1000+		Municipal	Sheriff	State		Totals
	(N=155)	(N=43)	(N=45)		(N=178)	(N=50)	(N=15)		243
Citizen complaints	138	41	42		164	43	14		221
Non-lethal force	140	37	39		161	43	12		216
Internal complaints	128	36	41		151	40	14		205
Vehicle damage/accidents	133	35	33		148	41	12		201
Internal investigations	113	38	34		138	35	12		185
Officer-involved shootings	98	34	34		125	31	10		166
Vehicle pursuits	114	25	27		124	31	11		166
Internal policy violations	100	28	32		116	34	10		160
Officer-involved firearm discharges	95	29	34		117	30	11		158
Internal reprimands	90	28	23		106	28	7		141
Failure to appear in court	47	18	18		69	12	2		83
Sick days used	49	9	13		52	16	3		71
Officer injury	48	9	14		56	14	1		71
Citizen injury	45	11	13		57	11	1		69
Tardiness	47	9	10		51	13	2		66
Loss of equipment	30	11	18		43	11	5		59
Civil suits	31	9	13		45	7	1		53
Failure to complete training	32	8	12		42	9	1		52
Unsatisfactory performance evaluation	34	3	2		29	9	1		39
Sick leave linked w days off	30	2	5		27	9	1		37
Sick leave linked w vacation	29	2	4		26	8	1		35
Discretionary (secondary) arrests	17	6	8		24	7	0		31
Failure to take annual physical	12	1	5		17	1	0		18

Identification: Time-and-numbers for Selected Performance Metrics

Both the total number of indicators and the threshold set for each indicator have implications for the number of officers who will be identified by the system. For example, an EIS

that considers a large number of indicators using a low threshold (e.g. 2 uses of force in 12 months) will "flag" more officers for review than a system that considers fewer indicators or a similar number of indicators but sets a higher threshold (e.g. 3 uses of force in 6 months). We turn now to a description of time and numbers thresholds for selected indicators.

Our survey data clearly indicate that most systems rely on a straight time and numbers approach like those shown in Table 5 below. Selection criteria take the form of thresholds associated with the indicators, and officers who exceed a count on the indicator in the specified time frame are identified for review and/or intervention. In Table 5 the numbers in the cells represent the number of agencies that set a particular combination as a basis for identifying an officer. For example, 24 agencies set 3 uses of non-lethal force in 12 months as the threshold for "flagging" officers for review and/or intervention.

Several themes emerge from the survey data. First, it is uncommon for departments to consider a time frame longer than twelve months. Secondly, departments most commonly consider the counts of events or outputs over the preceding six to twelve months. Third, the "tolerance" that departments have for the same performance metric varies across departments. For example, 8 agencies identify officers based on 3 citizen complaints in the preceding 3 months while 51 agencies identify officers who exceed 3 citizen complaints in the preceding 12 months. The threshold set and the time period over which counts are accumulated represent a tradeoff. The lower the threshold is set, the greater the number of officers who are likely to be identified by the system.

Table 5: Time and Number Thresholds

	Time Frames				
Thresholds – Non-lethal force	1 month	3 months	6 months	12 months	Totals (agencies)
1				2	2
2	4	1	5	8	18
3	5	10	17	24	56
4		5	4	8	17
5	1	5	2	8	16
6		4	5	17	26
7				1	1
8			1	1	2
10				3	3
12				2	2
15				1	1
Other	72				72
Column Totals (agencies)	10	25	34	75	216

Table 5 (cont.)					
Thresholds – Citizen complaints	1 month	3 months	6 months	12 months	Totals (agencies)
1		1		1	2
2	1	8	11	18	38
3	2	8	14	51	75
4		2	3	12	17
5	1			8	9
6			1	1	2
12				1	1
Other	77				77
Column Totals (agencies)	4	19	29	92	221
Thresholds – Officer-involved shootings	1 month	3 months	6 months	12 months	Totals (agencies)
1	2		1	9	12
2	2	2	6	20	30
3	1	2	8	15	26
4		1	1	5	7
5				3	3
6				1	1
12				1	1
Other	76				76
Column Totals (agencies)	5	5	16	54	166
Thresholds – Vehicle pursuits	1 month	3 months	6 months	12 months	Totals (agencies)
1		1	1		2
2		3	5	20	28
3	1	7	13	32	53
4		2		8	10
5				7	7
6			1	3	4
7				1	1
8				1	1
10				1	1
12				2	2
Other	57				57
Column Totals (agencies)	1	13	20	75	166

Table 5 (cont.)					
Thresholds – Failure to appear in court	1 month	3 months	6 months	12 months	Totals (agencies)
1	1			1	2
2	1	2	8	4	15
3	1	2	1	13	17
4	1			2	3
5				1	1
Other	45				45
Column Totals (agencies)	4	4	9	21	83

Post-Identification Review

Our survey data indicate that most departments provide for a second-stage review of officers who meet EIS thresholds. While initial identification is mechanical and turns on numerical counts, nearly two-thirds of departments introduce discretion into a second stage of identification that concentrates on officers who reach thresholds. Smaller departments (under 500 sworn) are somewhat more likely to provide for a review of officers who reach thresholds compared to the largest agencies (1,000 or more sworn), 64 percent and 58 percent, respectively. State police agencies are the least likely to provide for a second-stage selection process (47 percent) compared to municipal agencies and sheriff's offices. Refer to Table 6.

Among the agencies that provide for second-stage review, immediate supervisors are most commonly (42 percent) responsible for reviewing officer performance and arriving at a judgment of whether and, if so, what kind of intervention is needed. This holds true regardless of agency size or type. Internal affairs/professional standards is solely responsible for conducting the second-stage review in 21 percent of agencies.

Table 6: Post Identification Review by Agency Size and Type

	Agency size				Type of agency				
	< 500	500 - 999	1000+		Municipal	Sheriff	State		Totals
	(N=155)	(N=43)	(N=45)		(N=178)	(N=50)	(N=15)		243
Performance review conducted	99	31	26		122	27	7		156
Review conducted by:									
Immediate supervisor	40	17	8		50	14	1		65
Training personnel	1				1				1
Command staff	18	4	6		21	5	2		28
IA/professional standards	21	6	5		26	3	3		32
Other/multiple	19	4	7		24	5	1		30

Interventions

Respondents were asked to select from among seventeen various interventions those that are available in their agencies, and they were also provided the opportunity to write in interventions that were not included among the seventeen listed. We would highlight several points that emerged from the responses. First, there is much less variation in interventions compared to performance metrics. Second, the large majority (89 percent) of EI systems provide for counseling by one's immediate supervisor. Third, interventions include training in a host of topics, from review of departmental policies to human relations and stress awareness. Last, the type of interventions made available are, for the most part, similar across agencies of different sizes and types. Refer to Table 7, below.

Table 7: Types of Interventions by Agency Size and Type.

	Agency size				Type of agency				
Intervention	< 500	500 - 999	1000+		Municipal	Sheriff	State		Totals
	(N=155)	(N=43)	(N=45)		(N=178)	(N=50)	(N=15)		243
Immediate supervisor counseling	137	41	39		160	45	12		217
Professional counseling (e.g., EAP)	128	37	39		148	46	10		204
Training – departmental policy	132	38	33		152	40	11		203
Training – defensive tactics	125	38	31		142	41	11		194
Training – driving	128	38	28		141	43	10		194
Training – verbal de-escalation	119	37	34		140	38	12		190
Training – ethics	119	32	34		136	38	11		185
Counseling by other command staff	120	35	29		134	39	11		184
Training – cultural diversity	115	33	31		130	39	10		179
Training – weapons care	115	32	28		129	37	9		175
Training – communication skills	105	36	33		125	38	11		174
Training – human relations	106	30	30		123	35	8		166
Training – stress awareness	100	32	33		123	33	9		165
Reassignment/transfer	98	30	24		115	33	4		152
Peer officer support program	49	19	27		74	16	5		95
Crisis intervention teams	53	20	21		68	24	2		94

Post-Intervention Monitoring

In order to better understand accountability mechanisms in place around EI systems, we asked respondents to describe their practices for post-intervention monitoring, which are detailed in Table 8, below. Two-thirds of all agencies provide for post-intervention monitoring. The largest agencies (1,000 or more sworn) are less likely than agencies with fewer sworn to provide for post-intervention monitoring. There is wide variation in the duration of post-intervention monitoring, with a range from one month to two years. A twelve-month post-intervention monitoring period was the most commonly reported time frame (26 percent of agencies).

Table 8: Post-Intervention Monitoring by Agency Size and Type

	Agency size				Type of agency				
	< 500	500 - 999	1000+		Municipal	Sheriff	State		Totals
	(N=155)	(N=43)	(N=45)		(N=178)	(N=50)	(N=15)		243
Post-intervention monitoring	110	29	22		126	28	7		161
Duration of monitoring									
1 month	3				2	1			3
2 months	2				1	1			2
3 months	9	1	4		10	4			14
4 months	1				1				1
6 months	19	3	6		22	5	1		28
12 months	29	9	4		33	5	4		42
24 months	2	1			3				3
Other	9	4	2		11	4			15
Varies	24	8	3		32	3			35
Continuously	8	5	3		11	3	2		16
Written performance report required	86	21	18		99	22	4		125
Written report primarily by:									
Immediate supervisor	61	12	10		65	16	2		83
Training personnel	2		1		2	1			3
Command staff	11	6	3		17	1	2		20
IA/professional standards	3				3				3
Other/more than one	9	2	4		11	4			15

Among those agencies that require post-intervention monitoring, most (78 percent) require a written report as part of the follow-up process. Here again, an immediate supervisor is most often responsible for completing the written report (in 66 percent of the agencies requiring written follow-up).

Summary

The following summary statements can be made with respect to identification of potential problem behaviors. A wide range of performance metrics are considered in contemporary EI systems, with indicators ranging from citizen complaints to failure to take a physical. Almost all agencies consider citizen complaints and non-lethal force. Most systems rely on a straight time and numbers approach in which selection criteria take the form of thresholds associated with indicators. It is uncommon for contemporary systems to consider a time frame longer than 12 months, with 6-12 months being more common. The tolerance agencies afford to the same indicators varies. With respect to follow up on officers who exceed a threshold, most departments provide for second stage review, and among those that provide for second state review it is most commonly conducted by an immediate supervisor, followed by personnel in IA or professional standards. Agency responses indicate a fairly small range of interventions are made available with the large majority providing for counseling by an immediate supervisor and training. Two-thirds of agencies provide for some form of post-intervention monitoring, with fairly wide variation in the length of the post-intervention monitoring period.

EIS PROCESSES

The effectiveness of any EIS turns on its execution. Execution is one part mechanical and many parts human. The mechanical part – especially the use of performance indicators to initially identify officers who may have performance problems – deserves and has received a great deal of attention, though much more needs to be learned. The human parts have received much less attention, though no EIS will operate effectively if the human parts fail. In our case studies, we attended to both the mechanical and human parts.

As we explained above, in the project overview, we conducted case studies of EI systems in six police agencies. We make no pretense that these six systems are a representative sample. To the contrary, there is a good chance that the agencies that were willing to accommodate us have above average EI systems. We wanted to learn from systems in which agencies invest the effort necessary to make them work.

We briefly describe the agencies and their EI systems. Then we summarize the processes in numerical terms: the numbers of officers who meet EI thresholds and are initially identified, and the number to whom interventions of one kind or another are applied. Finally, we summarize the findings of the site visit interviews, describing the findings of our qualitative analysis in terms of major themes.

The Agencies

Table 9 summarizes the key structural features on which the case study sites varied. All of the names are those of villages in New York State. Agency sworn personnel levels are based on 2013 LEMAS survey responses and are approximate, to preserve the agencies' anonymity.

Table 9: Key Structural Features by Case Study Site.

	Brewster	Oriskany	Chester	Potsdam	Sidney	Caledonia
Full-time sworn personnel	2,000	1,500	1,000	1,000	1,250	250
Time & numbers – high thresholds		X				X
Time & numbers – moderate thresholds					X	
Time & numbers – low thresholds	X					
Peer comparison threshold			X			
Hybrid thresholds				X		
Selection review by EIS coordinator/IA	X			X	X	X
Selection review by supervisor		X		X	X	X
Selection review by mid-level manager			X			
Selection review by committee					X	
Post-intervention monitoring subject to higher review	X		X			

Several agencies have time and numbers indicators and thresholds, but they vary in the height of the thresholds. Oriskany's thresholds are fairly high – e.g., 3 personnel complaints in 3 months. Brewster's are fairly low: 2 complaints in 12 months or 2 uses of force in 12 months. Sidney's complaint threshold is between those: 3 complaints in 12 months. Police vehicle crashes – 2 in 12 months – and vehicle pursuits – 2 in 6 months – also form time-and-numbers indicators in Sydney.

But Sydney also applies thresholds within each of three categories of force that are based on the distribution of force. For category 1 force, the lowest level of force, the officers in the top 1 percent in a 6-month period define the threshold. For category 2 and 3 force, respectively, the officers in the top 5 percent define the threshold.

Chester provides for peer-based thresholds, such that each officer's counts are compared to the mean of their organizational group (precinct/unit and shift). All risk activities tracked are assigned weights. Officers who are one standard deviation above the mean for their organizational group (ie. unit and shift) and who also have activity in a second category, or who have a sustained complaint are flagged for further review. Chester also provides for multiple categories of outcomes and a couple of levels of intervention including supervisory observation, supervisory monitoring, supervisory review and supervisory intervention.

One agency – Potsdam – uses hybrid thresholds, i.e., ratios of force to arrests, as well as time-and-numbers thresholds. Officers may be identified for EIS review based on time-and-numbers counts: 3 or more uses of force in a 30-day period; 3 or more complaints in a 6-month period; or 2 or more complaints that each contain at least 1 allegation in the same complaint category (such as 2 complaints that allege conduct violations) in a 6-month period. In addition, however, officers may be identified based on a ratio of use of force to arrests that is at least 1:5 in a 6-month period, or a ratio of force to arrests that is at least three times the average of his/her shift for a 6-month period.

Three agencies provide for rolling 12-month thresholds and three assess officers' numbers on a quarterly schedule. All provide for a manual review of the performance of officers who reach thresholds. However, the parties who perform the review vary. In Oriskany, the review is conducted primarily by an officer's immediate supervisor. In Brewster, by contrast, the review is conducted only by internal affairs personnel. In Chester, unit captains (or Lieutenants if that is the highest ranking unit officer) perform the initial review. In the other agencies, the reviews are conducted by multiple parties: an initial review by the EIS coordinator, who may screen cases out or forward them to the officer's immediate supervisor for review. In Sydney, the supervisor's review is subject to review by a high-level committee.

Post-intervention monitoring is done by only the immediate supervisor in Oriskany, or by a higher authority in addition to a supervisor, in Brewster and Chester. In Brewster, the higher authority is the commander of internal affairs. In Chester, the higher authority is the chief, who convenes quarterly meetings of unit commanders for the purpose of checking on the status of officers subject to intervention.

Chester convenes quarterly command review sessions where the Chief of police and unit commanders are present and at which commanding officers are expected to report on the status of their subordinates who are in an active intervention plan.[10] Sidney convenes semi-monthly

[10] The focus of the meetings is dictated by the Chief. For example, some past Chiefs would question unit commanders about officers who were flagged and not subject to intervention as well as those who were subject to intervention while other chiefs limit discussion to only those under intervention.

meetings chaired by a deputy chief at which individual intervention plans are reviewed. In contrast to Chester, an EI coordinator discusses the intervention plans, not unit commanders.

Processes by the Numbers

Given the structural variation in these agencies' EI systems, the likelihood that an officer will reach or exceed an EIS threshold also varies from agency to agency. The diagrams in Figure 1 show in an average year how many officers worked in patrol, in blue, and the numbers who reached EIS thresholds, in orange.[11] The numbers of patrol officers and front-line supervisors are estimated based on personnel records. Though any sworn member can reach a threshold, officers and front-line supervisors at an elevated risk for reaching a threshold by virtue of both their exposure to the public and the nature of the behavioral indicators that can be measured. The diagrams also show, in yellow, how many officers were subject to some form of intervention. (Two shades of yellow, in Potsdam and Caledonia, are indicative of two different forms of behavior: problem behavior and commendable behavior.) The smaller, yellow "bubble" represents the officers assessed as true positives – those who would benefit from some form of intervention. The orange bubble, exclusive of the yellow bubble within it, represent the officers assessed as false positives – those who reached an EIS threshold, but whose behavior failed to exhibit a pattern warranting intervention.

EI structures that cast a wider net, much like Chester's or Potsdam's, generate a larger pool of officers subject to the second stage of review. While this allows for a more inclusive consideration of officer behavior, it also generates more false positives. EI structures that cast a smaller net, like Oriskany's and Caledonia's, generate fewer false positives, but they may be more susceptible to false negatives – officers whose behavior is problematic but is not detected by the EIS.

[11] Our quantitative analyses of process are, at the time of this writing, limited to four agencies.

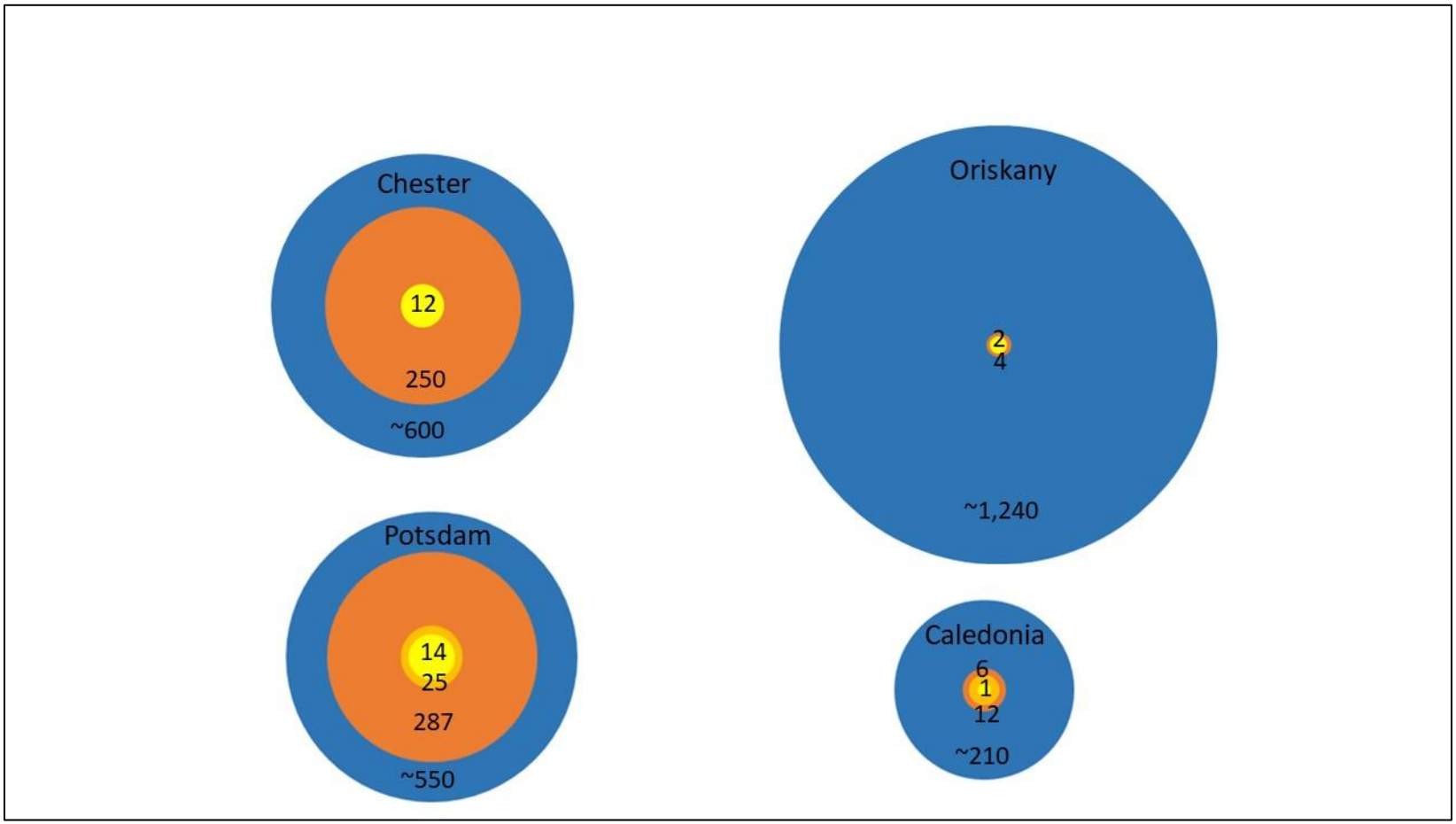
Chester
12
250
~600
Oriskany
2
4
~1,240
Potsdam
14
25
287
~550
Caledonia
6
1
12
~210

Figure 1

Perspectives from the Field

We turn next to a summary of the key themes that emerged through the interviews. We begin with themes that recurred across all of the agencies, followed by themes that were unique to – or at least more pronounced – in certain cities and seemed to vary with EI structure. . Our analysis of interview data revealed four issues/topics about which we found fairly widespread agreement: inadequacy of indicators in current EI systems; the importance of front-line supervisors; symbolic value of EI systems; and agency failure to institutionalize EI systems.

Inadequate Indicators

Both academics and practitioners alike have long recognized that current EI systems fail to take into account potentially valuable indicators of problematic behavior. We asked interviewees to share with us (1) indicators included in their agency's system that they believed were valuable, (2) indicators that were not included, but that they believed should be, and also (3) indicators that were included in their agency's system, but which they felt introduced "noise" and offered little to no value in identifying problem behavior.

Through these discussions, three main sources of data inadequacies emerged. First were the family of indicators viewed as invaluable but not quantifiable. Much of these discussions highlighted a focus on what is traditionally viewed as early warning systems, though these were not the terms in which interviewees framed their responses. The most common example shared with us focused on the importance of supervisors knowing the baseline or "norm" of their individual officers. For example, if a normally outgoing officer becomes withdrawn, or an officer who habitually shares meals with others on the shift no longer does so, a good supervisor should become aware that the officer is "off." By watching for changes in normal patterns in behavior, a supervisor can identify earlier the signs of problems and intervene. More broadly, these indicators were described as "the human factor," "organic knowledge," or even an "inkling" that could never be translated into a numeric value for incorporation into an algorithm.

The second class of indicators that most believed should but were seldom accounted for were seen as "off limits", even though they could be quantified. These indicators were generally housed in separate systems and considered protected human resource or administrative data. For example, we routinely heard that monitoring tardiness, counting sick days, or monitoring *when* days off were taken (repeatedly on a Friday or a Monday) were actions supervisors took to identify officers with personal problems that could eventually lead to on-duty issues. Supervisors reported that while these indicators were not part of the formal EI system they personally took them into consideration and viewed them as vital to understanding subordinates' performance. It bears noting that while our case study agencies did not incorporate such indicators into their systems, some EI systems have done so (e.g. sick time usage). .

The final inadequacy reported in most of the case study systems stemmed from the lack of a group baseline (different from the individual baseline/norm discussed above). A common concern was that absent a proper baseline against which to compare officers, active but good officers were inappropriately flagged and less active officers, some of whom could be exhibiting problem behaviors, were able to escape notice. The most common example provided was the active officer who worked a busy shift and/or a high crime area, and who as a result encountered the people and problems most likely to generate indicators that lend themselves to being quantified – such as complaints, uses of force, or accidents.

Importance of Front Line Supervisors

That supervision is crucial to the success of a police department is recognized by scholars and practitioners alike. The view that front line supervisors were critical, if not necessary, with respect to early, early intervention outside a formal system as well as under the umbrella of a formal EIS was widely shared by those with whom we spoke in case study agencies. EI systems should properly be conceived as only one piece of a larger system of accountability within law enforcement agencies. Indeed, when we spoke with field supervisors about performance management and accountability, most did not immediately speak of the department's EIS. In fact, discussion of the EIS as a means to monitor and manage performance often only followed our specific prompting about the formal EIS.

In the context of our conversations about accountability and monitoring officer performance, three main ways that front line supervisors could impact the success or failure of informal and formal measures to address problem officers emerged: identifying problem behavior; intervening where problems arise and monitoring subsequent changes in behavior; and in framing the intent of an agency's formal and informal efforts, thereby influencing officer trust in agency practices. Front-line supervisors were widely regarded as the gatekeepers for early identification and intervention with officers exhibiting signs of performance issues.

Symbolism

EI systems were viewed as symbolically valuable – either for internal or external constituents. Two schools of thought emerged. The first was the belief that EI systems had real or potentially favorable symbolic value, while the other adopted a cynical view of the symbolism behind an EIS. The balance struck between internal and external constituents shaped the extent to which the symbolic value of an EI system was viewed favorably or cynically. Having an EIS, regardless of its actual ability to change behavior, symbolized the commitment of the agency to its officers. If (and most acknowledged this as a big *if)* properly "marketed" to internal constituents, agency executives could demonstrate an internal orientation and commitment to supporting officers. The more cynical viewpoint was that the agency failed – either in marketing or in real intent – to package the EIS as internally oriented and merely adopted a formal EIS as a means to appease external constituents.

Failure to Institutionalize

The final theme that emerged – and about which there was widespread agreement – was the view that interviewees' respective agencies had failed to institutionalize their EIS. Training-related failures were viewed as key impediments to institutionalization. The lack of training, or the content of training when it was offered, were two such failings. In many cases, we were told that the agency had offered EIS training when the system was first introduced, but failed to offer additional trainings. As a result, personnel with many years of experience had only a rusty memory of the training, while newer personnel had never been exposed to EIS training at all.

Even where training was offered once or more frequently, the general view (which was corroborated in our review of policies and interviews with executives) was that the EIS training typically focused on the mechanics of the software behind the EIS, as well as on indicators and broad strokes of the process to be followed by supervisors. The trainings failed to focus on the intent of the system and on pattern identification. The failure to offer any training, refresher training, or training that goes beyond system mechanics was seen as causing ancillary negative consequences beyond limiting institutionalization.

The failure to incorporate the EIS into training curricula offered to line officers curbed general awareness of its existence, especially in agencies with high thresholds. Additionally, where officers knew of the system, the failure to offer training on its intent left a blank slate for officers to fill, and many with whom we spoke inferred that officers presumed it was a disciplinary system. In some agencies, supervisors indicated they became aware of the existence of their agencies system only when they were studying for the promotional exam from officer to supervisor. Even at that point, the promotional exam study guides focused on the role of front line supervisors in completing paperwork associated with the EIS rather than on the purpose of the system.

The inadequacy of indicators, the importance of front line supervisors, the symbolic value of an EIS (potential or realized, positive or negative), and the failure to institutionalize EI systems encompassed four main areas about which there was general agreement. Through the interviews, we distilled three additional areas that stakeholders viewed as important in understanding EI systems, but about which we detected more divergent points of view. The issues centered on "false positives," the demands placed on supervisors, and on the best ways to operationalize EI systems.

False Positives

Among practitioners and academics alike there has for some time been an understanding that false positives will emerge in data driven systems – with many limitations in the data sources tapped – designed to flag officers whose behavior may be problematic. Not surprisingly, false positives were seen by those with whom we spoke as inevitable products of EI systems. However, viewpoints differed with respect to the consequences associated with system generated false positives. The most common arguments regarding the negative consequences of a false positive centered on the detrimental impact they had for officer morale and productivity. Officer's enforcement activity would be inhibited out of a desire to avoid being flagged by the EIS.

On the flip side were those who acknowledged that systems would inevitably identify some officers incorrectly as potentially engaging in problematic behavior, but then went on to frame this as an opportunity. EI systems whose structure prompted front line supervisors to both review records of officers who crossed a threshold and look for *patterns* would reinforce the agency's expectation that front line supervisors exercise oversight over subordinates. Even in instances where, upon review, supervisors determine there is no pattern of problem behavior, it offers positive consequences – they hold front-line supervisors accountable for monitoring officer behavior and reinforce the agency's expectation that this function is a part of their role. Additional benefits that could yield from these opportunities were premised on the quality of the supervisor and their own trust in the EIS. Supervisors who believed that the EIS was introduced to help officers (not as a means to discipline them) should use the "flag" (whether ultimately found true or false) as the opportunity or "excuse" to have a conversation with subordinates, talk with them about how they are doing, and to frame the intent of the EIS in a positive light, thereby promoting officer trust in the system.

Demands on Supervisors

The next area about which there was agreement was that EI systems placed demands on front line supervisors. There were two schools of thought about whether the demand placed on supervisors was a drawback of the system. Some with whom we spoke appeared to see EIS

demands on supervisors as both inevitable and negative (even in agencies where very few officers were flagged and therefore came to the attention of front-line supervisors). It bears noting that this view was not exclusive to front-line supervisors – we heard it from mid-level managers and EI coordinators as well. Those who viewed the impact on front-line supervisors in a negative light shared two common, broad justifications: responding to EIS alerts took supervisors off the streets where they were needed more, and supervisors had to speak to the same officers repeatedly, despite the fact these officers were also repeatedly determined not to be exhibiting problem behavior (often to the same officers). This ultimately resulted in a loss of trust and legitimacy in the supervisor—subordinate relationship.

At the other end of the spectrum were those who agreed that EI systems would and should inevitably place demands on front-line supervisors, as this was the point of having an EIS and a benefit. We identified several themes from those who saw front-line supervisors as appropriately tasked by EI systems. One was that reviewing and monitoring subordinate behavior is a supervisor's job, not an "extra" demand. Moreover, having a formal EIS was viewed as a potentially useful tool in performing this part of the supervisory role. For example, EI systems that housed officer performance and activity in one place makes it easy for supervisors to learn about new officers coming into their unit, especially when review of new officers is a formal part of policy. Second, responding to EI alerts/flags provided a necessary prompt to many supervisors to review subordinates' behavior, which ultimately guarded against officers "slipping through the cracks" to the point where the opportunity to provide intervention had passed. The third theme centered on the idea that front-line supervisors played a specific role within the EIS, and it offered an opportunity to reinforce the agency's more general expectation that supervisors be engaged with subordinates in an ongoing and meaningful way.

Operationalization

How best to operationalize an EI system revealed divergent views. Before turning to the findings, we would note that during interviews we would explain to respondents how other agencies structured a key component of the EIS, thereby giving respondents the opportunity to juxtapose the example with their own system and reflect on the pros and cons of different approaches. Different schools of thought emerged on who should respond to alerts and where this should be housed, who should oversee interventions and perform post-intervention monitoring, and, last, on the periodicity of alerts.

The EI systems we studied differ with respect to the rank and placement of who completes the second-stage review of officers who cross a threshold. The proximity to the officer of the person conducting second-stage review cut two ways for those we interviewed. For most, the front-line supervisor closest to the officers was seen as uniquely positioned to separate problem behavior from the inevitable consequences of police work. For others, the closeness of the supervisor to the subordinate represented a problem, rather than a strength, based on the belief that immediate supervisors would be less objective and, rather, protective of subordinates. A more removed party – like an EI coordinator - who specialized in second-stage review was presumed to be more objective and better equipped to identify patterns, as well as able to distinguish noise from incidents that warrant further inquiry. Using staff other than first-line supervisors was a structural design viewed by some as having the additional benefit of reducing the number of cases going down the chain of command for further review. This resulted in minimizing the demand placed on front-line supervisors while bolstering the legitimacy of the

system in the eyes of those who make the wheels turn, as "noise" was not sent to front-line supervisors for review.

Post-intervention monitoring is another key structural piece of EI systems. While not common, some agencies include review by an executive level committee. Those with whom we spoke agreed post-intervention monitoring was crucial, and two schools of thought emerged with respect to the use of a higher level committee to undertake macro-level oversight. The perceived benefits of such an arrangement included the symbolic value in showing buy-in from the highest levels of the agency, the ability to detect agency-wide patterns of problems, and the ability to detect inequity in interventions for like patterns of problems.

For those who were less enthusiastic about the use of review committees, their concern stemmed from several sources. Until and unless EI systems could accurately identify true problem officers, it was unfairly detrimental to have the highest levels of the agencies privy to "potentially" problematic officers as this could lead to unfair stigmatizing or chances for officers' promotion to be negatively affected. This concern was especially pronounced for those who believed that their own or the typical EIS was better-tailored for identifying active officers than it was for identifying officers exhibiting potentially problematic behaviors. The notion that a review committee could monitor interventions across officers for similarities in the type or length of the intervention was a non-starter for some, insofar as they were of the mindset that this was against the grain of interventions. In their minds, if interventions were to be effective at bringing about change, they had to be tailored to the individual problem officer and the root causes of their problems. From this perspective, any disparity across officers was preferable to generic interventions, and also the relationship between supervisors and subordinates was a strength to be exploited, rather than a drawback to be avoided.

The final component in how to operationalize EI systems that was recognized as a key component to address was the periodicity of generating EI alerts. Upon reflecting on the pros and cons of rolling versus periodic alerts, we heard drawbacks to both arrangements. The drawback of rolling counts is that the same officer is more likely to be flagged repeatedly in a short period of time. Quarterly reviews help curb this issue, but introduce a tradeoff: if an officer genuinely requires intervention, 90 days or more could be lost if reviews are held only quarterly. Respondents recognized the real tradeoff between maximizing the timeliness of the intervention and minimizing the frequency of reviewing alerts of the same officer in a short time frame. If pushed to choose among the two options, the decision turned largely on faith in front-line supervisors. Those who felt front-line supervisors were both able to identify early signs of problems and willing to informally intervene with officers were more comfortable with quarterly alerts.

Summary

The views described above cross-cut the different case study sites. While there was consensus on four key themes - the inadequacy of current indicators, the importance of front-line supervisors, the symbolism tied to EI systems, and the failure to institutionalize EI systems - the local saliency of each varied somewhat. The perspectives of those with whom we spoke highlighted additional areas where distinct schools of thought emerged around the same issues – the consequences of false positives, the demands on supervisors, and the pros and cons of different structural arrangements. In making sense of why some areas of consensus were more or less salient locally, and of the viewpoints on either side of the same issue, we identified several forces that shaped perspectives.

Executive level emphasis on the goals of the EIS shaped local views. The impetus for introducing the EIS had an impact. The length of time the EIS had operated locally mattered. The larger culture of supervision within the agency was of importance. The balance the EIS struck with respect to high and low thresholds and the resultant number of alerts shaped conversations. The EIS within the larger system of departmental policies and practices mattered. Last, efforts to institutionalize the system shaped how respondents thought about their own EIS, as well as EI systems more generally.

The lessons we learned in the field made it clear that merely having an EIS is insufficient. We repeatedly heard or surmised that in most agencies there was very little general awareness of the system. Like most organizational change efforts, EI systems require agencies to direct real, sustained attention and resources toward the system. The IT side of the EI house has been awarded much more attention than the human side.

EIS OUTCOMES

Our outcome evaluation is designed to tell us whether interventions by each EI system are associated with detectable changes in officers' performance over time, before and after intervention, *and* whether those changes are attributable to the intervention.[12] Thus we ask whether the changes among officers subject to intervention are greater than the changes that we see among comparable officers not subject to intervention. In Chester we consider officers subject to only supervisory review as having been "treated," but we isolate them from those subject to a conventional EIS intervention; thus we have two treatment groups in Chester.

In each agency, we formed a control group matched one-to-one as nearly as possible to the officers subject to intervention. (In Chester, with two intervention groups, we have two corresponding control groups.) We took account of pre-intervention counts of complaints against the officers and their length of service, sex, race, and ethnicity. Table 10 displays comparisons of the control groups to the treatment groups in each agency with respect to length of service, sex, race, and ethnicity. Among these factors, length of service is arguably the most important, given the likely arc of police activity across an officer's career, and on length of service, the control groups are closely matched to the intervention groups. The matches are identical on officers' sex, and very close on race/ethnicity, though some control groups include a slightly disproportionate number of white officers.

For each officer subject to intervention, we formed annual rates of the outcome indicators for two years prior to the intervention – operationalized as the date on which the officer was "flagged" by the EIS – and for two years following. We formed corresponding outcome measures for each officer's matched control. Our estimation of intervention impacts turns on a "difference in difference": the magnitude of the pre-/post-intervention change for the intervention group relative to the pre-/post-intervention change for the control group.

Within the parameters of the available data – two years pre- and post-intervention – and the number of officers subject to intervention, the numbers of officers included in the analyses are quite modest, affording limited statistical power. In drawing inferences from the differences in differences, we cannot ignore the dispersion around the mean values, but neither should we rigidly adhere to a conventional .05 standard of statistical significance in rejecting the null

[12] Our quantitative analyses of outcomes are, at the time of this writing, limited to three agencies.

hypothesis. In fact, none of the differences reported below achieve statistical significance at the .05 level.

Table 10: Control and Treatment Group by Site.

	Length of service (mos)	% Male	% White	% Black	% Latino	% Other Race	N
Oriskany							
Intervention	104.8	94.7	63.2	0	31.6	5.3	19
Control	104.2	94.7	68.4	15.8	15.8	0	19
Chester							
Intervention	112.7	78.9	31.6	68.4	0	0	19
Control - Intervention	115.8	78.9	36.8	63.2	0	0	19
Review	118.4	96.4	71.4	17.9	0	7.1	28
Control - Review	121.0	96.4	75.0	25.0	0	0	28
Potsdam							
Intervention	90.3	90.9	81.8	9.1	0	9.1	11
Control	89.4	90.9	90.9	9.1	0	9	11

Oriskany

In Oriskany, complaints may be investigated by precinct supervisors or by internal affairs, depending on the nature of the alleged violations. Regardless of who investigates a complaint, it is given a tracking number if internal affairs is notified, though an unknown number of complaints are investigated by the precincts without notifying internal affairs. With respect to use of force reporting, Oriskany requires that officers complete reports on the use of firearms and less-lethal weapons, including chemical agents and electronic control devices, whether or not an injury is reported or suspected; in the cases of reported or suspected injuries, uses of physical force – e.g., pain compliance, take-downs, counter-moves – must also be reported.

We see declines in all of the indicators for both intervention and control groups in Oriskany; see Table 11. Complaint rates decrease to the same level, albeit from a somewhat higher baseline in the intervention group. The use of electronic control devices also drops to the same level, albeit from a somewhat higher baseline in the control group. There is no evidence of depolicing among the officers subject to intervention, for though their rates of arrests and stops decrease, the decreases are greater still among the control officers.

Potsdam

Potsdam has a citizen oversight authority that performs complaint intake, though complaints may be filed directly with internal affairs, any police facility, or any officer. Every complaint, by policy, is recorded and tracked regardless of whether it is filed with the citizen oversight authority of the police department. The lowest level of reportable force is force "that is intended to establish control of a resistant subject, though not intended or reasonably likely to cause persistent pain or physical injury," such as "non-striking use of baton, takedown performed

in completely controlled manner with no injury, handcuffing against resistance, and pointing of a firearm."

In Potsdam, both complaints and use of force decreased more for the officers subject to intervention more than they did for the controls; see Table 12. The pre-intervention baselines were somewhat higher for the treatment group, and the post-intervention levels were lower for the treatment group. Both groups followed a roughly similar decline in arrests and stops, leading us to infer that the decrease among treatment group officers is not attributable to the intervention. These findings are consistent with the hypothesis of intervention impacts, though the small N and the small differences relative to the variance in the indicators caution against a strong conclusion.

Table 11: Outcomes: Oriskany Interventions, 2001-2012

	Intervention		Control	
Outcome Indicator	Pre	Post	Pre	Post
Complaints*	1.6	0.7	1.1	0.5
Use of force*	1.3	0.6	1.5	0.6
ECD use*	0.8	0.2	1.3	0.3
Arrests**	32.9	26.8	40.6	23.8
Stops**	147.4	125.7	171.2	111.9
N	19		19	

* Data from January 1999, through June, 2015
** Data from January 2004, through June, 2015

Table 12: Outcomes: Potsdam Interventions, 2014

	Intervention		Control	
Outcome Indicator	Pre	Post	Pre	Post
Complaints*	2.01	0.83	1.64	1.3
Citizen initiated*	1.75	0.72	1.48	1.23
Use of force*	4.57	3.42	4.4	4.1
Pointing of firearm*	2.06	1.13	1.71	1.74
Arrests*	85.1	58.7	80.2	59.5
Citations*	27.8	16.8	29.9	10.3
N	11		11	

*Data from January, 2012 through December, 2016

Chester

Like Potsdam, Chester has a citizen oversight authority that performs complaint intake, though any police facility or officer must accept citizen complaints. Every complaint, by policy, is recorded and tracked regardless of whether it is filed with the citizen oversight authority of the police department. Reportable force includes the use of "hard hands," any physical strike or

"instrumental contact," the use of chemical weapons and electronic control devices, canine bites, firearm discharges, and any action that results in serious injury.

In Chester, only very small differences in differences appear; see Table 13. The largest difference of the hypothesized nature is in citizen complaints for the intervention group, which was halved from a rate of nearly 1 per year. Other differences in anticipated outcomes are negligible. Stops increased for all groups, but with smaller increases for the two treatment groups.

Table 13: Outcomes: Chester Interventions, 2012-2013

	Intervention		Control - Intervention		Supervisory review		Control - Review	
Outcome Indicator	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Internal investigations*	0.62	0.38	0.51	0.20	0.65	0.53	0.51	0.25
Use of force*	1.10	0.91	0.89	0.64	1.73	1.62	1.31	0.91
Vehicle pursuits*	0.21	0.19	0.10	0.08	0.23	0.29	0.22	0.20
Citizen complaints*	0.95	0.51	0.52	0.48	0.93	1.17	0.57	0.56
Civil suits*	0.02	0.11	0.02	0.01	0.02	0.04	0.03	0.02
Firearm discharges*	0.01	0.08	0.01	0.05	0.02	0.03	0.01	0.04
Stops**	123.6	130.7	61.3	63.8	120.7	172.1	93.3	147.0
N	19		19		28		28	

*Data from January, 2006 through December, 2015 **Data from January, 2009 through December, 2015

Conclusions

We find in these data very little evidence that the interventions had either salutary or detrimental impacts on the officers to whom they were applied. In general, these officers' use of force, complaints, and other risk-related outputs declined from pre- to post-intervention. In two of the three agencies, these officers' enforcement outputs also declined. If we had examined data on only these officers, as treatment groups, we would guardedly infer that the interventions had (small) impacts, though we might also suspect that the effect was achieved by discouraging enforcement activity. With the analytic benefit of the matched control groups, however, it is fairly clear that the outcomes observed for the treatment groups follow a trend that is comparable to that of the control groups. We therefore infer that the interventions had either no impacts or (as in Potsdam) very small and limited impacts at best.

These analyses are limited in several respects. First, the estimated impacts of EIS interventions are average effects, but if the effects are contingent on the nature of officers' problems, the nature of the intervention, and the match of intervention to problem, then the average might underestimate the effects under some conditions.

Second, the outcome measures are limited to those based on data that are routinely captured in police records. The effects of the interventions on behaviors that are not reflected in these indicators are beyond our reach. Considering the vagaries of citizen reporting of police misconduct, complaints are not a reliable indicator of officers' behavior. Considering the many forces that affect the incidence with which officers become involved in situations that prompt the

use of force by police, counts of force may better reflect the demands placed on officers than their behavioral propensities. In other words, the same "noise" in these indicators that compromise their value in identifying problematic behavior also compromise their value as outcome measures.

Third, the statistical power of our analyses is limited by small Ns. The estimated effects of intervention may not be sufficiently precise to detect small effects, which could be lost in the variance. Our findings may reflect not statistical limitations but rather programmatic realities. First, in these agencies, the identified officers with confirmed patterns of problematic conduct/performance are not, on average, egregiously problematic. In Potsdam, for example, the officers subject to intervention who had most frequently used reportable force reported fewer than 20 such events in two years; compared with the EIS threshold of 3 uses of force in 30 days, less than one per month on average represents a fairly low baseline.

Second, the EIS not the only or even primary mechanism for regulating performance and correcting performance deficiencies. Day-to-day supervision is (and always has been) the principal mechanism. In these agencies, day-to-day supervision may be preventing many of the kinds of problems to which EI systems would otherwise direct attention.

Agency Snapshots

The Appendix includes brief descriptions of EI systems in each of eight agencies. Here we highlight several themes that we can draw from the documented experiences of those agencies, about many of which we will have more to say in the conclusions.

First, the snapshots underscore that EI systems require an administrative investment. All of the systems on which Walker et al. (2000) conducted case studies, with complimentary findings, subsequently ceased to function. Minneapolis, it appears, reestablished an EIS in 2009, though several years later it was not considered to be operating satisfactorily. New Orleans reestablished an EIS under the terms of a consent decree. Miami-Dade had no EIS in 2007, and we could find no trace of an EIS there in 2018.

Second, information technology is a recurring and costly challenge. Any but the simplest EIS requires the integration of disparate information systems housing different types of police records. This requires technical expertise that may exceed that of a police agency's personnel or even of a city's information technology unit. Oakland may be the poster child for this problem, but OPD is hardly the only police department that has struggled with information technology in building (or sustaining) its EIS.

Third, and relatedly, many stakeholders see value in building a data warehouse that serves the information needs of an EIS as well as other human resource (HR) purposes. Insofar as the EIS may benefit from a larger effort to assemble human resource information, the initial and maintenance costs of the EIS itself is reduced, and as a component of an HR system, it may be less susceptible to an image as punitive.

Fourth, the technological challenges may have obscured the equally or more important human elements of any EIS. No mechanized system – not even Charlotte-Mecklenburg's machine-learning algorithm – can obviate the role of police personnel in reviewing the details of individual officers' performance to determine whether intervention is warranted. Moreover, regardless of the predictive success that any actuarial assessment of risk can achieve, it will be (for the foreseeable future) humans who diagnose the nature of an officer's performance

deficiencies, prescribe the form(s) of intervention, and apply the intervention(s). Much less attention has been devoted to the execution of these human functions.

Fifth, and relatedly, training is essential, and it is probably inadequate in most instances. New Orleans may be an exception that proves the rule, having developed and delivered to its supervisors training not only in the basics of the policy and procedure but also in pattern identification.

Sixth, EI systems exhibit many structural permutations of oversight and accountability – who is responsible for what, and who ensures that responsible parties fulfill their responsibilities. Some of those provide for a minimum of oversight and accountability. We think it likely that forms of oversight and accountability have important implications for the institutionalization of an EIS.

Seventh, concerns about identification errors – false positives and false negatives – are ubiquitous. Many internal stakeholders are interested in sparing good cops from administrative poking. Many external stakeholders – including especially but not limited to court-appointed monitors – are interested in throwing a wide net around potentially problematic conduct to better ensure that officers who need intervention get it. At any level of predictive accuracy, however, reducing the incidence of one kind of error leads to more of the other kind.

Eighth, Charlotte-Mecklenburg's system raises the question about what it is that an EIS is designed to prevent: patterns of problematic conduct or individual adverse events. CMPD's system, and its risk ranks, is designed to prevent the latter, though it remains to be seen how successfully it prevents adverse events. Predicting adverse events can be accomplished to some extent by capitalizing on the fact that some officers have much greater exposure – i.e., contacts with citizens – than others, putting them at higher risk of mistakes or bad judgments. Whether individual events can be prevented is an open question. Predicting patterns of problem behavior is, we suspect, a different proposition.

Finally, and most generally, there are no evidence-based practices, leaving agencies to make educated guesses as they fit (or refit) an EIS to their organizational environments (unless a monitor is making the guesses).

Conclusions and Recommendations

The concept and practice of EI seems fairly well-established, particularly among larger agencies. Be that as it may, our agency survey and case studies have confirmed that EI systems exhibit a lot of structural variation despite their common components. The number and nature of the indicators that are incorporated into EI varies from agency to agency, and with respect to any one indicator, the thresholds and the time periods over which counts or other computations are formed also vary. Some agencies' EI thresholds are so high that few officers reach them; other agencies' thresholds are much lower, resulting in a much larger number of officers who appear on the EIS radar.

The case studies revealed that, even when the thresholds are fairly high, the behavior of many officers who reach thresholds is, upon closer inspection, not indicative of a pattern that calls for a corrective or preventive intervention. We knew that the available indicators are not valid measures of police misconduct. Moreover, the most common time-and-numbers types of thresholds are very crude predictive instruments. They do not honor some basic principles of risk assessment (Worden, Harris, and McLean, 2014), and they are unnecessarily restricted to a

small 3- or 12-month slice of officers' career histories. We believe that many agencies EI systems leave quite a bit of room for improvement in this regard. They could better advantage of the accumulated information about officers' performance, and they could better take account of variations in officers' exposure. On their face, peer comparisons are superior to time-and-numbers thresholds, though defining peer groups is devilishly difficult, since even within a police precinct and a single platoon or shift, exposure will vary across beats. Ratio-based thresholds may be more effective in controlling for officers' exposure.

Even so, until much more is learned about the etiology of police misconduct and the forms that it takes, any actuarial assessment of risk will yield errors of two kinds, false positives and false negatives. Manual review of records of officers' behavior is required to differentiate the false positives from the true positives. In most instances, immediate supervisors perform this function, though some agencies assign this responsibility to other parties, such as an EI coordinator or unit, and in others the responsibility is shared between an EI coordinator and an identified officer's supervisor. We see some virtue in dividing the duty, such that most of the false positives are screened out centrally while supervisors remain involved in this process when doubts cannot be resolved without their participation. In some agencies, this initial judgment is subject to review by others – e.g., in the officer's chain-of-command, or an executive-level committee – and we also see virtue in such oversight of the second-stage identification process.

Whoever it is that performs this task, they should be adequately trained for it. They should be instructed about the breadth of the review – what kinds of records should they consult, and over how long a period of time? For what should they be attentive in assessing officers' behavior for identifiable patterns? The people with whom we spoke in the case study agencies had not received such training. Some of them, such as the lieutenant in internal affairs who had been conducting all of his agency's reviews for years, may have learned by doing. If so, then his expertise should be captured in a set of procedures, so that when he retires or otherwise moves on, his successor can climb the learning curve much more quickly. If this task is delegated to and among field supervisors, each of whom may have occasion to perform it only episodically, then a set of guidelines and training would be essential. It would also serve to better institutionalize the EIS.

Regardless of the part that field supervisors play in the identification process, the nature of supervision surely has an important bearing on the operation of an EIS. It is sometimes said that an EIS can and should serve to improve police supervision, and we heard words to that effect in some of the case study interviews. We do not disagree; if first-line supervision in an agency is weak, then an EIS that is properly managed can be a stimulus for better supervision. However, if first-line supervision in an agency is strong, as it appeared to be in some of the case study agencies, then the EIS plays a smaller role in regulating officers' performance. Effective supervisors become acquainted with their subordinates' strengths, weaknesses, and predilections, monitor how they do their work, intervene informally or formally when needed. Not all police organizations enable field supervisors to perform their role optimally. Supervisory training may be lacking, or the division of supervisory labor and/or assignment practices may undermine supervisors' capacities to supervise effectively. We can only speculate based on the limited research on police supervision and our impressions formed in our field research, but we allow as how the effect of supervisory practices on EI systems is much greater than the effect of EI systems on supervision. We would further speculate that investments in supervision would pay dividends as great as or greater than investments in more elaborate EI systems.

Larger patterns of police management can likely affect how supervision is done. Police supervision is not merely a matter of constraining officers' choices and sanctioning officers who violate policy, of course. It is also about developing officers' judgment and decision-making skills so that they can better serve their community, and about providing or arranging for assistance when – as the human beings they are – officers hit one of life's rough patches. In principle, an EIS serves both the purposes of controlling misconduct by officers and of supporting officers, but in practice, the organizational tone suggests that one or the other purpose is predominate. In one of our case study agencies, the executive had made it clear that every sworn member of the agency owed it to his/her colleagues to intervene – if only to speak up – when someone was putting themselves at risk. The executive cited the career-ending drunk-driving arrest of one officer, who was widely known in his precinct to have an alcohol abuse problem, as an object lesson for the entire agency. The message that this was a preventable outcome was clearly received by members of his agency: we are responsible for one another. Supervision, and the agency's EIS, was infused with this orientation. Neither was perceived as punitive. In other agencies, especially but not only those whose EI systems were part of a consent decree and shaped around a concern with controlling misconduct, the EIS tends to be seen by officers as a threat rather than as a support.

These human parts of early intervention deserve much more attention that they normally receive. When we first struck up conversations about agencies' EI systems, the natural inclination of people in the agencies was to tell us about the information system on which the EIS drew: sources of data about indicators, how the data were organized, how/whether the software generates automatic alerts when thresholds are reached. As difficult as it is to form comprehensive information systems, it may be the more tractable and less important feature of EI systems. Once a pattern of problematic conduct is identified, how the sources of the problem are ascertained, how an intervention or set of interventions determined, and how they are applied are all tasks that are inescapably human. Preparing them to perform the tasks and ensuring that they do so are essential to early intervention.

Finally, EI systems are not unlike any of a number of administrative or programmatic innovations that require on-going commitment among potentially competing priorities. Thus they are vulnerable to administrative turnover: if an executive or a key administrator is replaced by someone who does not share the same level of commitment to the EIS, then it is liable to atrophy. As one agency survey respondent indicated, in explaining that the agency no longer had an EIS: "it faded into obscurity." No one need make a judgment that an EIS is not needed (though that also happened in a few cases). It might simply wither.

APPENDIX
EIS Snapshots

Miami-Dade

Walker, Alpert, and Kenney's (2000) examination of early warning systems featured the Miami Dade Police Department's EIS as among the earliest mechanisms tracking officer behaviors with the aim of identifying potentially problematic behavioral patterns. MDPD's EIS tracked complaints, use-of-force incidents, internal investigations, commendations, and disciplinary actions taken toward officers. Supervisors reviewed reports quarterly and annually, and after its first full year in 1981, monthly. When supervisors received flagged officers' reports, they would discuss the incidents in question with the officer and assign a disposition. This may have included no further action, referrals or internal or external department services, supplementary training, or, most frequently, informal counseling with the immediate supervisor. Supervisors were encouraged to think of these reports as a more general "resource" for their supervisory duties, reflecting the MDPD's uniquely integrated EIS viewpoint and revealing the system's broader utility for the department. The MDPD had no written directives regarding follow-up monitoring for EIS officers, nor a formal policy about record-keeping. Walker et al. (2000) note the consistent history of the MDPD's EIS, citing the strong administrative support and organizational integrity as significantly responsible the system's overall success. The authors distinguished the MDPD's system from the others of its time by remarking that the "EW system [is] found to be simply one part of a larger personnel data system that… is part of a broader commitment to accountability" (2000: 5.21). However, according to Miami-Dade's responses to the 2007 LEMAS survey, the MDPD's EIS was by then no longer operational.

Minneapolis

The Minneapolis Police Department's Early Intervention System was implemented during a turbulent period for the agency, which had attracted attention for allegations of police brutality and misconduct. The system was introduced in conjunction with the election of a new mayor, appointment of a new chief, and development of a Civilian Review Authority, which as Walker et al. (2000) note, renders an evaluation of the independent effects of the EIS difficult. The system relied solely on citizen complaints to identify officers in need of intervention, and employed a threshold of three in a span of twelve months. This selection criterion, however, was not consistent throughout the system's history, and Walker et al. (2000) remark that at one point in time the MPD collected information on only the 20 officers who garnered the most civilian complaints, and at another point in the mid-90s, the MPD's EIS was effectively nonexistent. It is unclear how frequently supervisors reviewed EIS reports at the time of Walker et al.'s study, but following the release of the NIJ report, the MPD installed a new policy of bi-weekly reviews of all "potentially problematic officers." Interventions included informal counseling sessions with direct supervisors, and the MPD had no explicit post-intervention follow-up or documentation procedures. Walker and Archbold critique the MPD EI system's neglect of supervisory focus, which "paid little explicit attention to the behavior of supervisors" (2014: 144).

More recent evidence of the MPD's EIS is scarce, but we infer that the EIS ceased to operate sometime after Walker, et al.'s study. In its 2007 annual report, the Minneapolis Civilian Police Review Authority (known by the acronym CRA) observed that

> The MPD is proposing to establish an early warning system to identify officers who may be experiencing job performance difficulties. The CRA is seeking to have its data included in

the MPD early warning system. CRA data combined with MPD data would provide a more complete picture of an officer's conduct. The CRA would advocate for such a combination of data. A final determination of this matter has not been made

In or around 2013, then-Chief Janee Harteau requested an independent review of the department, including its EIS. Technical assistance offered by the Department of Justice's Diagnostic Center was directly sought to assess the MPD's EI practices. The Diagnostic Center's report noted that "the department developed a police Early Intervention System (EIS) in 2009, but it has gaps and is not used uniformly" (2015: 5). At that time, according to the report, the MPD's EIS suffered from insufficient buy-in, too-few interventions offered, minimal training, and "lack of automation."

New Orleans

Following a series of federal investigations of the New Orleans Police Department (NOPD), a 2012 consent decree included a mandate to implementat an EIS, whose stated purpose was to "promote constitutional and professional police practices; manage risk and liability, to evaluate the performance of NOPD employees across all ranks, units, and shifts."[13] The early 1990's version EIS, evaluated by Walker, et al. (2000), had fallen out of use. In 2016, the NOPD invested $4 million in a newly designed EIS called INSIGHT. The new EIS is defined as a human resources and personnel management system and data warehouse. Department supervisors receive and integrate member information in order to facilitate close and effective supervision and identify patterns or series of incidents that might indicate at-risk behavior. Training is heavily emphasized in the implementation of a new EI system. The training for the current system involves 8-hour sessions during which supervisors and commanders are taught to identify patterns of behavior by analyzing data, looking for "broader, pattern-based reports at least quarterly" (NOPD Operations Manual, 35.1.9). NOPD supervisors recently underwent additional training at the monitor's request. INSIGHT includes a dashboard that allows supervisors to "look at everybody in their command in 30 seconds and get a summary of what they have done over the past 24 hours."[14]

NOPD created an Early Intervention Unit (EIU) in conjunction with INSIGHT. This unit is comprised of three commanders: an EI Unit director, IT director, and Deputy Chief of Staff, who are tasked with tracking, reporting, and liaising with the various officers and units flagged or otherwise involved in EI processes. This unit constitutes the INSIGHT Executive Committee, who create peer groups and threshold measures, and whose duty it is to continually review the effectiveness and functionality of the system, as well as individual officer's flags if necessary. This committee also maintains updated policies based on mental health professional's recommendations and current practices regarding interventions. INSIGHT's thresholds are based on ratio measures, such as use of force to total arrests, that utilize peer-comparisons to identify either positive or negative patterns of behaviors, as well as count thresholds for certain indicators (including vehicle pursuits and citizen complaints).

Interventions include a range of non-punitive measures that are specifically aimed to address an individual pattern of behaviors. A supervisor might determine that an informal meeting with the officer would be sufficient to address these behaviors, or they might refer the

[13] United States v The City of New Orleans, Civil Action 12-CV-01924.

[14] Steen, M. (2017, April 10). *Do Intervention Systems Amount to Scrutiny of Police Departments?* Retrieved from: www.governing.com/templates/gov_print_article?id=421376514.

officer to remedial training. If an informal meeting is determined to be insufficient, a formal meeting with command staff may be scheduled, or the officer's duties may be altered to reduce stress. Transfers to another unit to better suit professional strengths or counseling services are available, or a supervisor may devise an evaluation period in which review periods for officer behaviors (e.g., 30- 60- 90- day increments) would allow for expanded supervision and intermittent mentoring.

NOPD has not been immune to technological difficulties. The April 2018 Consent Decree report noted that all supervisors were not able to view all personnel under their command in INSIGHT; each sergeant was able to view only a fraction of the officers assigned to the shift the sergeant supervises.

LASD

In *The New World of Police Accountability*, the Los Angeles Sheriff's Department's Personnel Performance Index (PPI) was highlighted as an exemplary EIS whose implementation was "widely regarded as arguably the best EIS in the country" (Walker, 2014: 165). The PPI operated in conjunction with the Performance Mentoring Program (PMP), which placed eligible officers in an intervention plan to be mentored by three supervisors: one Commander, one Captain, and one Sergeant. In 2003, the PPI's greatest shortcoming, in the judgment of the Special Counsel to the LASD (an LASD-appointed oversight counsel) and described by Walker, was human error. Sloppy reports and inconsistent report submission or review were cited as reasons for the system's halting success.

The 2009 27th Semi-Annual Report of the Special Counsel to the LASD provided a thorough history of the EI system and an evaluation of its contemporary efficacy, concluding that the PPI "could be better." The Counsel's request to incorporate certain performance indicators into the system to match other contemporary EI systems had not yet been fulfilled, and the Performance Mentoring Committee held too few meetings. By the time of this report, the LASD tracked a handful of performance indicators, including complaints, all uses of force, founded administrative investigations resulting in suspension, and shootings. These indicators were analyzed with the aim to identify patterns; that is, complaints en masse would not be flagged, but rather types of complaints that form a pattern of behavior would trigger an alert. The Counsel encouraged the department to give greater attention to allegations of false statements or information, uses of derogatory language, absences, performance standards, and unresolved administrative investigations. The Counsel also urged an incorporation of "detention and arrest-related data, investigations, inmate complaints, and data on warrantless stops and seizures" (p. 38). The Special Counsel to the LASD was disbanded in 2014, and the 34th and final report stated that the LASD PPI "must be overhauled, updated, and expanded," comparing the system to "your grandfather's Cadillac or Oldsmobile" (p. 7).

More recent accounts of the LASD's EIS are difficult to find, as the department does not post their PPI policy in their publicly available manual of policy and procedures. A recent Settlement Agreement with the Department of Justice outlined reforms for the LASD in Antelope Valley (which constitutes a large swath of the LASD's geographic jurisdiction) after finding a "pattern or practice by law enforcement officers that deprives individuals of rights, privileges, or immunities secured by the Constitution or Federal Law."[15] The DOJ investigation also found LASD violations of the Fair Housing Act. The Settlement Agreement proposes

[15] CV 15-04174.

reforms to the PPI and PMP, which were expected to be complete by this year (2018). Some of the modifications include:

- Accounting for peer comparison for LASD deputies and unties
- Including data about stops and searches, compliance with community engagement requirements
- Modifying data entry for PPI so that it is maintained electronically
- Expansion of the Performance Mentoring Program

Baltimore

The Baltimore Police Department's EIS is currently undergoing an overhaul as part of the recently implemented consent decree. Several high-profile incidents contributed to a Department of Justice investigation of the agency. The DOJ investigation concluded "BPD engages in a pattern or practice of conduct that violates the Constitution and federal law, including: (1) making unconstitutional Stops, Searches, and Arrests; (2) using enforcement strategies that produce severe and unjustified disparities in the rates of Stops, Searches and Arrests of African Americans; (3) using excessive force; and (4) retaliating against people engaging in constitutionally-protected expression."[16] The DOJ report characterized BPD's EIS as "effectively nonfunctional" at the time (Herron, 2016).

The current BPD EIS policy, dated about a year post-consent decree in April 2018, states that the system is intended to be a non-disciplinary, confidential, managerial tool to identify possibly problematic behavior and provide individualized interventions for those members who require them. Though the policy explicitly holds that "EIS does not list specific threshold limits per type of incidents in order to prevent unwarranted interventions, or the late identification of a member's negative behavior," an earlier description of the EIS explained that three complaints in a rolling 12-month period generates an EIS alert (Herron, 2016). The policy alludes to alerts produced by IAPro, which suggests that a numerical threshold exists but is not advertised and is subject to change. The policy lists indicators that might trigger a member for an alert as: excessive force complaints, misconduct allegations, preventable departmental accidents, substance abuse, domestic-related incidents, and any violations of departmental policy. It also notes that these indicators do not represent a comprehensive account of items that may trigger an alert.

One unusual element of this policy is the role of the Early Intervention Unit (EIU). The policy explains that the BPD EIS "relies on thoughtful consideration of EIU members tasked with reviewing the system" (Policy 1707). When alerts are generated, it is the EIU that "reviews the narrative in the alert(s), and any other relevant documents to the incident or the member's past behavior, to determine whether a referral for intervention is appropriate." Thus the determination of whether a pattern of problematic conduct exists is performed not by an officer's immediate supervisor, but rather by the EIU.

Vernon Herron, the Commander of Command Investigations, EI, and Court Liaison Unit, penned an article in *The Police Chief* detailing the system, citing the familiarity of EIU members as a strong benefit:

> …the detectives currently assigned to the Early Intervention Unit have worked with or, on many occasions, are familiar with the officers who are selected to participate in the early intervention process. Many times, these detectives can provide a historic

[16] United States of America v The Police Department of Baltimore City et al. A.4.

> perspective on the officers that is not always captured in BlueTeam [a common software used for for EI systems]. For instance several of the officers who have been identified for interventions have been involved in traumatic events during their careers, such as police-involved shootings, serious injuries in the line of duty, or sustaining injuries during the civil unrest. (2018:7)

Another unusual feature of BPD's EIS policy is the phased implementation of interventions and of EIU involvement. When the EIU first determines that intervention is necessary, the officer enters Phase 1, during which the officer's behavior is reviewed by the first-line supervisor, and is monitored for a 30 day period. A Performance Improvement Plan (PIP) is also optionally assigned to the officer at this stage, created individually for the officer's needs in an effort to "clearly define the performance problem/issues…and recommend specific solutions" (Policy 1707). PIPs range from supervisor guidance (including sergeant ride-alongs, reviewing member's training history, "identifying and supporting positive behavior"), trainings, assignment change, counseling, or referrals to other supportive service programs. If an officer triggers another alert during Phase 1, s/he enters Phase 2, in which the PIP becomes mandatory, and the first-line supervisor again monitors the officer for 30 days. If another alert is triggered during the 12-month rolling case of a Phase 2 intervention, the member reaches the Phase 3 of the Intervention Process. The stated goal of the final phase is to "permanently resolve persistent issues that a member displays" (Policy 1707). A PIP is developed by the member, their entire command, Director, and EIU, and the officer is again monitored for 30 days.

Compared with current policy, the BPD consent decree outlines a much more specific and detailed list of triggering incidents and behaviors, including all uses of force, all injuries and deaths, vehicle pursuits and traffic collisions, all instances in which an officer issues 3 or more citations in a single encounter, use of force used when the subject is charged with resisting arrest, all misconduct complaints, training records, and officer sick leave usage.

Los Angeles Police Department

A 1996 review of the LAPD's then-existing EIS – the Training Evaluation and Management System (TEAMS) – performed for the Police Commission found the system ineffective and anemic, and encouraged the department to seek reforms. The Commission applied for and presented the LAPD with $175,000 in federal grant money to accomplish this goal. In the wake of the Rampart scandal, however, it was discovered that the LAPD had not yet touched this grant money and that the agency's existing EIS was not functional.

The LAPD entered into a consent decree in 2001, which ensued partly from "…a pattern or practice of unconstitutional or otherwise unlawful conduct that has been made possible by the failure of the City defendants to adopt and implement proper management practices and procedures." The consent decree outlined the requirements for a remodeled EIS, resulting in the development of TEAMS II and the Risk Management Information System (RMIS). It specified a list of performance measures to incorporate into the system, including but not limited to: all non-lethal uses of force, all officer-involved shootings or discharges, all vehicle pursuits and traffic collisions, civil and administrative lawsuits, assignment and rank history, all arrest reports, citations, and stops. In addition to these requirements, the consent decree directed that protocol would provide for TEAMS II reports to be used for supervisor reviews, performance evaluations, reviews for officer transfers, and would be available at the officer's request. Included in the stipulations regarding supervisor's role in pattern analysis, LAPD managers are likewise to

perform regular reviews of their subordinate supervisors' behavior to ensure that proper oversight was maintained.

TEAMS II is a data sharing and analysis tool that contains an EIS as well as a separate system that tracks use of force incidents. The modern version of TEAMS II utilizes peer comparisons to identify individual officers or groups of officers who exhibit at-risk behavior. Peer groups are created based on the officer's type of work or contact frequency with the public. The largest peer group includes all patrol officers who are not in a supervisory role – about 3,000 officers (Swatt, Uchida, and Solomon, 2017). When an officer reaches a threshold, an "action item" is sent to the officer's supervisor, who is to perform an independent review of the officers' history in addition to a comparison of the officer's history against his or her peers and thereupon give the officer a performance assessment. If the supervisor deems the officer to be displaying a pattern of at-risk behavior, available dispositions include: no action, commendation, informal meeting, training, special evaluation reports, modified field duties, Risk Management Committee referral, Directed Behavioral Sciences Services referral, notice to correct, or a complaint. As noted above, a review of TEAMS II by the Office of the Inspector General for the Los Angeles Police Commission (2014) found that most action items culminate in the conclusion that "no pattern of behavior posing potential risk was identified."

TEAMS II is the subject of outcome evaluations. Uchida, Swatt, and Solomon (2017) estimated the department-wide impact of the initiation of TEAMS II on the use of force and citizen complaints. Applying an interrupted time series design, they found that the rate of use of force events decreased immediately after TEAMS II was implemented and slowly increased every month subsequently. They also found that the rate of citizen complaints decreased immediately after TEAMS II was implemented and continued to decrease subsequently. They note, however, that in the absence of a control series, these changes may be due to other factors. Swatt, Uchida, and Solomon (2017) estimated the individual-level effects of action items on patrol officers' use of force and citizen complaints, finding mixed effects. Within four 28-day deployment periods of an action item, no impact on either use of force or citizen complaints was detected; within 13 deployment period (i.e., about one year), action items were associated with a slight decrease in use of force and a slight increase in citizen complaints. Their analysis did not isolate officers in whose behavior supervisors identified a pattern signifying potential risk, however, in effect treating action items – most of which result in no action – as a homogeneous set of interventions.

<u>Oakland</u>

Following an investigation by and a settlement agreement with the Department of Justice in 2003, the Oakland Police Department was placed under a court-appointed independent monitor. The settlement agreement required that OPD develop and implement an EIS, which was initially known as the Personnel Information Management System (PIMS):

> This computerized relational database must maintain, integrate, and retrieve data necessary for supervision and management of OPD and its personnel. Specifically, this data must be used by OPD to promote professional police practices; manage the risk of police misconduct; and evaluate and audit the performance of the OPD members of all ranks, employees and OPD units, subunits, and shifts. (OPD Negotiated Settle Agreement, 3rd Semi-Annual report)

The saga of OPD's EIS is largely one of technological hurdles.

As of the monitor's 3rd annual report in 2004, PIMS was "progressing without concern" and notable progress had been made. By the 7th annual report in 2006, PMIS had been replaced

by the Personnel Assessment System (PAS), which was cited as a notable achievement by then-Chief Wayne G. Tucker. The PAS was an EIS similar to many threshold-based systems of that time, intended to be a non-disciplinary intervention program created to meet the monitor's settlement agreement criteria. The PAS's extensive performance dimensions included uses of force, all officer-involved discharges (firearm and OC canisters), canine deployment, civil suits, citizen complaints, in-custody deaths and injuries, reports of financial claim, assignment history, training history, sick leave usage, among others. The system used peer-group based thresholds, with groups determined by a Personnel Assessment Unit. Dispositions for officers determined to be exhibiting concerning behavioral patterns were either a supervisory monitor, in which a supervisor would develop a system of oversight over a certain period of time or monthly basis, or an intervention. Intervention strategies widely ranged from mentoring, training, peer counseling, referral to the EAP or counselor, substance abuse rehabilitation, required use of a recording device, consultation with the District Attorney's office, or a performance deficiency notice. With the exception of a voluntary Chaplain referral, all interventions were mandatory.

Despite the strides made in the development and implementation of the PAS by this time, periodic reviews revealed inconsistencies in officer alerts, moderately high numbers of false positives, and disagreements about dispositions for flagged officers among commanders, deputy chiefs, and PAS Review Panel. During its search for a replacement system, the OPD was not able to remain in compliance with the Settlement Agreement. The first mention of the revamped PAS system, the Performance Reporting Information Metrics Environment (PRIME), is contained in the 26th monthly report from November of 2015. The new system, which was originally developed by Microsoft, was implemented in May of 2017. The department struggled during the early implementation stages to maintain risk-assessment and pattern-identifying productivity. This slow start impeded the EIS review process because of data complications and programming bugs, greatly reducing the number of reviews. In total, the City spent roughly $6 million on PRIME.

In the 51st monthly report in February 2018, the monitor noted that many of the issues that emerged from PRIME stemmed from the fact that it was a singular database that "replaced multiple files that could be and were easily audited by the separate division within OPD." Issues with unnecessary data replication and omission were discovered weeks after the fact. It was revealed that PRIME was not capturing arrest data for a 6-week period in November 2017, and was also not equipped to create alerts based on that missing information. In that same year, the City IT Department worked to hire two full time non-sworn staff members to manage the PRIME system, as well as attempted to contract a project manager to re-engineer the system.

The 54th monthly report reflects optimism for improvements in PRIME, expecting "that a redesigned system will be implemented in summer of 2019." In an itemization of the issues involved with the system, the monitor noted three main components of their efforts are to

1. redesign PRIME;
2. include in PRIME links to the Human Resources Management System, officer training data, BWC footage; and finally
3. Develop a dashboard for supervisors in order to assess stop data (as part of a project with Stanford), and improve overall functionality.

In this same report, a newly hired consultant identified an additional issue with PRIME that was previously overlooked: new arrest records were being loaded into PRIME on a nightly basis, creating new records for all arrests rather than only updating new or changed arrest data, producing extra records in PRIME not related to any activity. The monitors report also notes that

the city is expected to hire two full-time staff members to replace the impromptu IT unit that had formed in OPD to deal with data issues as they arose with the EI system.

In August of 2018, the 55th monthly report notes that there are communication difficulties between the commanders and supervisors with PRIME venders and techs regarding how data would be collected and used. In order to address this issue, the monitor provided a list of suggestions to address the ongoing issues with PRIME implementation and revisited several definitions and objectives of the system:

1. Track PRIME data on a monthly basis;
2. To create a system for reporting to assist with the management process (to identify those with unusually high levels of stops, or who are exemplary in their productivity);
3. To decide what the data should look like (for threshold development);
4. To identify the questions that the database can help answer; and
5. To determine how the data in PRIME can be used to assess how well the risk management system is working.

In the 56th and most recent report from September 2018, the monitors note that additions to the database are being developed, which includes the incorporation of the newly developed Human Resources Management System (replacing the old personnel database), stop-data analytics in their risk management assessments, and the creation of "dashboards" for supervisors and commands staff (developed by Stanford as a part of understanding implicit bias). There are lingering concerns about PRIME excluding training information or allowing supervisors access to body worn camera footage, the incorporation of which into the program is not as far along as the personnel database. These encompassing improvements would require external expertise, and the monitor warns of the dangers of relying on external consultants (data scientists) without the development of internal expertise.

Charlotte-Mecklenburg

The stated purpose of the Charlotte-Mecklenburg EIS is to "enhance communication, coaching, and mentoring for supervisor-employee relationships" (CMPD GO 800-010) in a non-punitive format, using predictive analytics and machine learning to identify individual employees who might either warrant recognition for exemplary performance, or require interventions in order to correct behaviors and improve interactions with citizens.

Instead of relying on simple threshold measures to generate alerts regarding patterns of employee behavior, the Charlotte-Mecklenburg EIS flags officers through a "machine-learning" method that uses an algorithm to calculate daily "risk ranks" for officers rather than a "pure binary classification" (Helsby et al., 2017: 5). The risk ranks reflect the estimated probability of "adverse interactions," defined as "a potential violation of law or policy that may occur in the future if proper interventions or assistance are not offered to the employee" (CMPD GO 800-010). The algorithm incorporates a large volume of information from multiple departmental data sources concerning officers' histories, demographics, and external employment – much more in terms of the breadth of information and its time span than most EI systems, including CMPD's previous EIS. Information on officers' behaviors include 1-, 3- and 15- year counts of arrests, citations, uses of force, secondary employment, training history, field interviews, traffic stops, citizen complaints and criminal complaints. Information on the neighborhoods in which officers work include crime rate, average voter rate, publicly available data from 2010 and 2012 neighborhood quality of life studies, and the area's the 311 non-emergency call-rate. The EIS also compares officers to the mean frequencies for other officers assigned to similar divisions or

beats. The machine learning model uses these sets of indicators to create risk ranks by determining how each indicator should be weighted to predict the risk of an adverse event. The machine leaning approach reportedly allows the system to be "easily retrained on new data and with new variables" (Helsby et al., 2017: 5), and feedback from the EIS Administrator, supervisors, and commanders further allows the model to "train itself."

To demonstrate the predictive power of the machine-learning model, Helsby, et al. (2017) used data from 2009 to simulate the performance of the algorithm, making predictions of risk for 2010 incidents and comparing predicted outcomes with actual events. Their results showed that the model would have correctly predicted 12% more officers at risk of an adverse event compared to the previous EI system, and decreased the incorrect identification of officers by 32%. The model was determined to be an improvement upon the previous EIS system, whose predictive success was only marginally better than random selection. Using a machine learning binary classification method, authors concluded that three factors were the most indicative of risk for officers. The first is past behaviors: officers who are routinely found to have been engaged in an adverse incident are likely to engage in another. The second is stress, and authors noted that officers who respond to suicide calls are at higher risk for adverse incidents, as are officers who respond to domestic violence and calls with children. The third most predictive factor was patrol neighborhood.

At least monthly, the EIS Administrator reviews the top 5% of officers (or a percent determined at the EIS Administrator's discretion) determined to be at risk, and for any one officer who is predicted to be at risk, either dismisses the risk rank or creates an alert for his/her chain of command. Alerts are automatically assigned to a direct supervisor, and after initial review are moved up the chain of command for approval or rejection. Supervisors are not dependent on the system to issue alerts and may generate alerts themselves based on their own observations of officers' behaviors or personal information.

The types of interventions provided include counseling, remedial training, referral to the Employee Assistance Program (EAP), peer counseling referral, or other recommendations of a CMPD psychologist in the Community Wellness Division. Counseling would entail a meeting or scheduled meetings with a supervisor to discuss and review an employee's understanding of a policy, behavior, safety or tactical issue. If a supervisor recommends remedial training, the officer is to successfully complete a course or instruction or one-on-one training. However, an EAP or Peer counseling referral is voluntary.

Because the CMPD EIS is in its early stages, the full effectiveness and functionality of the program is yet to be determined. CMPD leadership has stated that they intend to supplement the risk ranks generated by the EIS with body worn camera footage in order to observe differences between those officers who lie on higher risk ranks to officers that fall on the bottom risk scores. Additionally, NOPD expects to use group-level predictions for certain groups of officers, allowing them to provide interventions for a broader target, though these advancements are still in developing stages.

REFERENCES

Bobb, Merrick J., and staff. 2009. *Los Angeles County Sheriff's Department: 27th Semiannual Report.* Los Angeles: Police Assessment Resource Center.

Bond, Graham, D. (2018 January 31). "Oakland Police Spent $6 Million on a Computer System to Track Officer Misconduct, and It Doesn't Work." *East Bay Express,* Retrieved from https://www.eastbayexpress.com/oakland/oakland-police-departments-compliance-with-federal-court-ordered-reforms-in-jeopardy-due-to-failure-of-key-computer-system/Content?oid=12877068.

Baltimore Police Department Monitoring Team. 2018, July 18. Baltimore Police Department Monitoring Team First Semiannual Report. Venable LLP and 21CP Solutions, LLC.

Bobb, Merrick J. 2009. *27th Semiannual Report to the Las Angeles County Sheriff's Department*. Police Assessment Resource Center.

Brandl, Steven G., Meghan S. Stroshine, and James Frank, 2001. "Who Are the Complaint-Prone Officers? An Examination of the Relationship between Police Officers' Attributes, Arrest Activity, Assignment, and Citizens' Complaints about Excessive Force," *Journal of Criminal Justice* 29: 521-529.

Burgess, Rachel, Kelli M. Evans, Charles A. Gruber and Christy E. Lopez. 2004, *Third Quarterly Report of the Independent Monitor: Delphine Allen, et al,v. City of Oakland, et al.*

Bustamante, Alexander A. 2014, August 20. *Los Angeles Police Commission Review of the Department's Early Warning System.* From the Office of the Inspector General.

Cody, Crystal, Estella Patterson, Kerr Putney, Jennifer Helsby, Joe Walsh, Lauren Haynes, Rayid Ghani, Samuel Carton, Kenneth Joseph, Ayesha Mahmud and Youngsoo Park. 2016. *Building better early intervention systems.* Police Chief Magazine.

Cohen, Bernard, and Jan M. Chaiken. 1972. *Police Background Characteristics and Performance.* New York: RAND Corporation.

Crank, John P. 1993. "Legalistic and Order-Maintenance Behaviors among Police Patrol Officers: A Survey of Eight Municipal Police Agencies," *American Journal of Police* 12: 103-126.

Davis, Robert C., Christopher W. Ortiz, Nicole J. Henderson, Joel Miller, and Michelle K. Massie, 2002. *Turning Necessity into Virtue: Pittsburgh's Experience with a Federal Consent Decree* (New York: Vera Institute of Justice).

Davis, Ronald, Anthony Rachal, Carolyn Marsh and Donna Howell. 2004. *Oakland Police Department Negotiated Settlement Agreement, 3rd Semi-Annual Report.* Oakland Police Department and Office of the Inspector General.

Figuero Paul, Andrea Van Peteghem and Carolyn Marsh. 2006. *Oakland Police Department Negotiated Settlement Agreement, 7th Semi-Annual Report.* Oakland Police Department and Office of the Inspector General.

Freund, Helen. 2017. *Monitor: NOPD 'early warning' Program for Problem Cops Shows 'Significant Progress,' More Training Needed.* The Advocate, Retrieved from https://www.theadvocate.com.

Friedrich, Robert J., 1980. "Police Use of Force: Individuals, Situations, and Organizations," *Annals of the American Academy of Political and Social Science* 452: 82-97.

Fyfe, James J., and Robert Kane, 2005. *Bad Cops: A Study of Career-Ending Misconduct among New York City Police Officers*, Report to the National Institute of Justice (New York: John Jay College of Criminal Justice).

Greene, Jack R., Alex R. Piquero, Matthew J. Hickman, and Brian A. Lawton. 2004. *Police Integrity and Accountability in Philadelphia: Predicting and Assessing Police Misconduct*, Report to the National Institute of Justice. Philadelphia: Temple University.

Harmon, Rachel. 2009. "Promoting Civil Rights through Proactive Policing Reform," *Stanford Law Review* 62: 1-68.

Harris, Christopher, 2009. "Exploring the Relationship between Experience and Problem Behaviors: A Longitudinal Analysis of Officers from a Large Cohort," *Police Quarterly* 12: 192-213.

Hassell, Kimberly, and Carol Archbold. 2010. "Widening the Scope on Complaints of Police Misconduct," *Policing: An International Journal of Police Strategies & Management* 33: 473-489.

Helsby, Jennifer**,** Samuel Carton, Kenneth Joseph, Ayesha Mahmud, Youngsoo Park, Andrea Narvarrete, Klaus Ackermann, Joe Walsh, Lauren Haynes, Crystal Cody, Estella Patterson, and Rayid Ghani, 2017. "Early Intervention Systems: Predicting Adverse Interactions between Police and the Public," *Criminal Justice Policy Review*: 1-20.

Herron, Vernon, 2016. "A New Approach to Early Intervention: BPD Case Study," *The Police Chief.*

Independent Commission on the Los Angeles Police Department, 1991. *Report of the Independent Commission on the Los Angeles Police Department* (Los Angeles: Author).

Kelly, Kimbriell. 2016, August 21. *Can Big Data Stop Bad Cops?* The Washington Post. Retrieved from https://www.washingtonpost.com.

Lane, Emily. 2017, September 1. NOPD Program Aims to Stop Officers with Problems from Becoming Problem Officers. Retrieved from https://www.nola.com/crime/index.ssf/2017/09/nopd_program_aims_to_stop_offi.html.

Lersch, Kim Michelle, 2002. "Are Citizen Complaints Just Another Measure of Officer Productivity? An analysis of citizen complaints and officer activity measures," *Police Practice and Research* 3: 135-147.

Lersch, Kim Michelle, Thomas Bazley, and Thomas Mieczkowski, 2006. "Early Intervention Programs: An Effective Police Accountability Tool, or Punishment of the Productive?" *Policing* 29: 58-76

Macintyre, Stuart, Tim Prenzler, and Jackie Chapman, 2008. "Early Intervention to Reduce Complaints: An Australian Victoria Police Initiative," *International Journal of Police Science and Management* 10: 238-250.

Moore, Mark H. 2002. *Recognizing Value in Policing: The Challenge of Measuring Police Performance*. Washington: Police Executive Research Forum.

Office of the Inspector General for the Los Angeles Police Commission, 2014. *Review of the Department's Early Warning System*. Los Angeles: Author.

Osher, Christopher N., 2011. "Out of 232 Red Flags on Denver Officers, Two Led to Formal Remediation," *The Denver Post*, November 11.

Rector, Kevin and Jessica Anderson. 2018, August 6. *One Year into Consent Decree, Baltimore Police Reforms Move Forward Slowly and Behind the Scenes.* The Baltimore Sun.

Steen, Margaret. 2017, April 10. Do Intervention Systems Amount to Scrutiny of Police Departments? Retrieved From www.governing.com/templates/gov_print_article?id=421376514.

Sherman, Lawrence W., and Denise Gottfredson, 1997 "Appendix: Methodology for this Report," in Lawrence W. Sherman, Denise Gottfredson, Doris MacKenzie, John Eck, Peter Reuter, and Shawn Bushway, *Preventing Crime: What Works, What Doesn't, What's Promising* (College Park, MD: University of Maryland).

Swatt, Marc L., Craig D. Uchida, and Shellie E. Solomon, 2017. "The Impact of Action Items on Use of Force and Complaints Among Patrol Officers," unpublished paper, Justice & Security Strategies.

Uchida, Craig D., Marc L. Swatt, and Shellie E. Solomon, 2017. "The Impact of an Early Intervention System on Use of Force and Civilian Complaints: An Evaluation of LAPD's TEAMS II," unpublished paper, Justice & Security Strategies.

Van Maanen, John. 1974. "Working the Street: A Developmental View of Police Behavior," in Herbert Jacob (ed.), *The Potential for Reform of Criminal Justice.* Beverly Hills: Sage.

Walker, Samuel. 2003a. "The New Paradigm of Police Accountability: The U.S. Justice Department 'Pattern or Practice' Suits in Context," *St. Louis University Public Law Review* 22: 3-52.

Walker, Samuel. 2003b. Early *Intervention Systems for Law Enforcement Agencies: A Planning and Management Guide.* Washington: Office of Community Oriented Policing Services.

Walker, Samuel, Stacy Osnick Milligan and Anna Berke, 2006. *Strategies for Intervening with Officers through Early Intervention Systems: A Guide for Front-Line Supervisors.* Washington, DC: Police Executive Research Forum.

Walker, Samuel and Carol A. Archbold. 2014: *The New World of Police Accountability (2nd ed., pg. 165).* Thousand Oaks, CA: Sage.

Walker, Samuel, Geoffrey P. Alpert, and Dennis J. Kenney, 2000. *Responding to the Problem Police Officer: A National Study of Early Warning Systems*, Report to the National Institute of Justice (Washington: National Institute of Justice).

Warshaw, Robert S. 2015, November 12. *Twenty-Sixth Report of the Independent Monitor for the Oakland Police Department.* Independent Monitor for the Oakland Police Department.

Warshaw, Robert S. 2018, February 23. *Fifty-first Report of the Independent Monitor for the Oakland Police Department.* Independent Monitor for the Oakland Police Department.

Warshaw, Robert S. 2018, July 19. *Fifty-fourth Report of the Independent Monitor for the Oakland Police Department.* Independent Monitor for the Oakland Police Department.

Warshaw, Robert S. 2018, August 23. *Fifty-fifth Report of the Independent Monitor for the Oakland Police Department.* Independent Monitor for the Oakland Police Department.

Warshaw, Robert S. 2018, September 13. *Fifty-sixth Report of the Independent Monitor for the Oakland Police Department.* Independent Monitor for the Oakland Police Department.

Worden, Robert E. 1989. "Situational and Attitudinal Explanations of Police Behavior: A Theoretical Reappraisal and Empirical Assessment," *Law & Society Review* 23: 667-711.

Worden, Robert E., Mary A. Pratte, Shelagh E. Catlin, Chris Harris, and Shelley L. Schlief, 2003. "Problem Officers, Problem Behavior, and Early Warning Systems," presented at the Annual Meeting of the Academy of Criminal Justice Sciences.

Worden, Robert E., Christopher J. Harris, and Sarah J. McLean, 2014. "Risk Assessment and Risk Management in Policing," *Policing: An International Journal of Police Strategies & Management* 37: 239-258.

Worden, Robert E., MoonSun Kim, Christopher J. Harris, Mary Anne Pratte, Shelagh Dorn, and Shelley S. Hyland, 2013. "Intervention with Problem Officers: An Outcome Evaluation of an EIS Intervention," *Criminal Justice and Behavior* 40 (April): 410-438.

Worden, Robert E., and Sarah J. McLean, 2016. "Prediction and Intervention to Prevent Police Misconduct," in Ronald Burke (ed.), *Stress in Policing: Sources, Consequences and Interventions.* London: Routledge.

Policies:

Oakland PD General Order D-17 (2013)
Baltimore PD Policy 1707 (2018)
New Orleans PD Operations Manual, Chapter 35.1.9 (2016)
New Orleans PD Operations Manual, Chapter 35.1.10 (2016)
Charlotte-Mecklenburg PD General Order 800-010 (2017)
Los Angeles PD General Order 668 (date not specified)

Consent Decrees:

United States of America v The City of New Orleans, Civil Action 12-CV-019240.
United States of America v Police Department of Baltimore City, et al., Civil Action 1:17-CV-00099-JKB.
United States of America v The County of Los Angeles and The Los Angeles County Sheriff's Department, Civil Action CV 15-03174.
United States of America v County of Los Angeles and Los Angeles County Sheriff Jim McDonnel, Civil Action 15-5903.
United States of America v City of Los Angeles, CA, Board of Police Commissioners of the City of Los Angeles, and the Los Angeles Police Department, Civil Action 00-11769.
Delphine Allen et al., v City of Oakland et al., Master Case No. C00-4599 THE (JL).