

## Metadata of the chapter that will be visualized online

---

Chapter Title	Police Performance Measurement	
Copyright Year	2014	
Copyright Holder	Springer Science+Business Media New York	
Corresponding Author	Family Name	<b>Uchida</b>
	Particle	
	Given Name	<b>Craig D.</b>
	Suffix	
	Organization/University	Justice and Security Strategies, Inc.
	Postbox	P.O. 6188
	City	Silver Spring
	State	MD
	Postcode	20916
	Country	USA
	Phone	(301) 438-3132
	Email	<a href="mailto:cduchida@jssinc.org">cduchida@jssinc.org</a>

---

# **Systems of Performance Measurement**

**Dr. Craig D. Uchida**

**President, Justice & Security Strategies, Inc.**

[cduchida@jssinc.org](mailto:cduchida@jssinc.org)

**Submitted to the Encyclopedia of Criminology and Criminal Justice  
Springer-Verlag**

**October 2012**

## **Overview**

This essay discusses systems of performance measurement for police agencies. It briefly examines trends in measurement from the early years of policing to the 21st century. Common themes that run through the essay include: what are the purposes for measurement? Who is concerned about the measure (i.e., are there external or internal pressures or concerns)? Have the measures been implemented and to what degree or level of success? Have researchers conducted evaluations? What are the results?

The essay focuses on four major areas -- 1) fundamental or traditional measures; 2) the rise of Compstat as a performance management tool; 3) the use of Early Intervention or Early Warning Systems to measure police misconduct; and 4) comprehensive systems of performance measurement.

## **Fundamentals**

What should be measured by police? What is the appropriate measure of success? Is it crime reduction, clearance rates, complaints, calls for service, or use of force incidents? Is it the number of community meetings attended by officers or the quantity and quality of problem-oriented policing projects? In most instances, answers to these questions depend on the mission and goals of the agency, but are further complicated by the political problem of whether they should be measured and the technical problem of

how to measure them. Selecting and using appropriate measures is a problem shared by administrators in most public agencies that have no single, definable outcomes (Maguire and Uchida 2000; Bayley 1994; Moore and Braga, 2003).

It is well established that law enforcement officers engage in many different activities. They make arrests, process offenders, serve warrants, quell disturbances, respond to emergencies, work with the community, solve problems, form relationships with the community, and perform many other things in many jurisdictions. These activities are their outputs. Most attempts to measure and explain these outputs have relied on arrest, clearance, response times, and crime statistics.

Since the inception of municipal police departments in the early to mid-1800s, police agencies have gauged their performance primarily through arrest rates and reported crimes, though historians have shown that they also kept track of lost children and other non-crime activities (Monkkonen, 1981). The rise of professionalism in the 1900s led the International Association of Chiefs of Police (IACP) and the FBI to focus on crime-fighting rather than non-crime services. This marked a profound shift in how police characterized their work and led to the collection of national data through the Uniform Crime Reports (UCR) (Uchida, 2010).

Reformers like Chief August Vollmer of the Berkeley Police Department and the IACP urged departments to maintain records related to crime and operations. The IACP Committee on Uniform Crime Records was formed in 1927 to create a system for collecting uniform police data. Standardized definitions were formulated to overcome regional differences in the definitions of criminal offenses. By 1929, the committee finalized a plan for crime reporting that became the basis for UCR. The committee chose

to obtain data on offenses that come to the attention of law enforcement agencies because they were more readily available than other reportable crime data. Seven offenses, because of their seriousness, frequency of occurrence, and likelihood of being reported to law enforcement, were initially selected to serve as an index for evaluating fluctuations in the volume of crime. These Crime Index offenses were murder and non-negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny/theft, and motor vehicle theft. By congressional mandate, arson was added as the eighth Index offense in 1978.

There were two primary purposes for uniform crime reporting. First, police leaders saw the value of better data for the improvement of management and operation strategies. Second was the desire for data that could combat the perception of crime waves created by the media. Police managers were concerned that such crime waves were seen as a reflection of their departments' inadequacies. They believed that uniform crime data would dispel the notion of crime waves and, therefore, demonstrate the value of police efforts to control crime. The tension between developing a system that would represent a true measure of crimes known to police and developing one that would prove crime was not increasing has been present in UCR since the beginning (Wellford, 1982).

Since 1930 the UCR has evolved from a relatively small data collection effort into a large-scale national effort. In its first year of operation, the program collected data from 400 cities in 43 States, representing about 20 million people. In 2011 UCR collected data from nearly 17,000 jurisdictions in all 50 States, representing about 300 million people. UCR remains a voluntary reporting program, with city, county, and State law enforcement agencies reporting monthly to the FBI the number of part 1 offenses and part

1 and part 2 arrests that have occurred within their jurisdictions. In addition to monthly tallies of offenses and arrests, additional data are captured on particular offenses, and data on age, sex, race, and ethnicity are collected for arrests.

Another important measure is response time, the seconds or minutes that it takes for a patrol car to respond to a call for service. This measure has been used by police to show the public how efficient and effective they are in dealing with crime. But research has shown that rapid response does not lead to apprehension of suspects immediately nor is it an expectation of those who call the police (Spelman & Brown, 1981). Changes in available resources have led police to streamline their response strategies and in the process make measurement more specific. Police have employed alternative strategies, known as differential police response (DPR) where priorities are established regarding calls, civilians take reports over the telephone, or complainants are asked to use the Internet and online reporting. Since the 1980s, departments often use 'priority' call types. For example, Priority 1 may include crimes in progress or more serious violent crimes (homicides, rapes, robberies), while Priority 2 and 3 calls may include those crimes that require non-immediate responses (burglaries, alarms, etc.).

### **Changes in Measures**

With the discovery of police discretion in the 1960s, researchers and police administrators realized that order maintenance and other non-crime activities were also important to measure, as officers were not engaged in crime fighting all of the time. But it wasn't until the 1990s with the community policing movement did researchers and police begin to recognize that this new paradigm should include new performance measures. Alpert and Moore (1993) recommended the adoption of non-traditional measures that

were tied to essential goals of community policing. Yet they still focused primarily on police-related data and measures, including number of traffic tickets, crimes and arrests cleared by conviction, and police misconduct. Some of the non-traditional measures included time and dollars allocated to problem-solving initiatives, types of assistance provided to victims, and resources for strengthening police-community relationships. From this point forward, researchers and police began to develop appropriate measures that covered the spectrum of traditional policing to community policing and problem-oriented policing activities and outcomes.

### **The Rise of Compstat**

In 1994 Commissioner William Bratton and Jack Maple introduced the New York City Police Department (NYPD) to Compstat (comprehensive computer statistics), a management tool that emphasized using appropriate resources in the right place at the right time to reduce crime. A harsh and oftentimes unforgiving method, Compstat was known for publicly questioning precinct commanders about their knowledge of crime in their areas and what they were doing about the problem. Using technology and computer generated mapping, Compstat generates crime statistics for each precinct and uses them to hold precinct commanders accountable for their areas (Bratton with Knobler, 1998).

Compstat became a popular mechanism for crime reduction, and police departments across the country quickly adopted the methodology. Indeed, by 2000 nearly one-third of the country's largest police agencies had implemented a Compstat-like program (Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2001). This occurred even though evidence of crime reduction was not validated by research efforts. The explosive growth in computer power (in terms of speed and storage capacity relative to

cost), the standardization of formats across software programs, the increase in user accessibility (and decrease in cost) to these software programs, among other factors facilitated the dramatic spread of Compstat across police agencies. In addition, cities, counties and at least one state (Maryland), have established CitiStat, CountyStat, and StateStat (respectively) to assist in decision-making and accountability.

Though some scholars express doubts about the success of Compstat (especially for crime reduction), it is generally accepted that Bratton mobilized the NYPD and later, the Los Angeles Police Department (LAPD) to raise arrest productivity significantly.

From an administrative perspective, Compstat represents a significant change from the past. Moore and Braga (2003) see Compstat as setting the standard for innovation for police management and as a unique system for internal and external accountability. They emphasize that as a measurement system it is "behaviorally powerful" in that it drives the NYPD and other departments in the following ways:

- It is aligned with organizational units that hold managers accountable for performance;
- The measures are simple and continuously used so that performance can be observed over time;
- The measures are aligned with those who oversee the department externally;
- Accountability is frequent, so that managers are attentive;
- Managers recognize that Compstat affects their current and future standing, pay, and promotional opportunities;
- Reviews of performance are public; and
- Comparisons can be made across situations and managers.

While the original model of Compstat has evolved since 1994, its basic principles of accountability and measurement remain intact.

### **Systems to Account for Police Behavior**

The problems of police-community relations identified four decades ago by the President's Commission on Law Enforcement and Administration of Justice (1967) continue to exist. Concern over use of force by police officers led Congress in 1994 to

order the U.S. Department of Justice (DOJ) to acquire data about the use of excessive force by law enforcement officers and to publish an annual summary of the data acquired (McEwen, 1996). The Bureau of Justice Statistics served as the repository for such data.

Concern over the practice of racial profiling by police officers making traffic stops, a long- time concern of minority communities, led to calls for measurement on the severity of the problem. During the early part of the 21st century a number of large and small departments across the country began to implement methods to measure officer encounters with the public as a result. More than twenty states have passed legislation prohibiting racial profiling and/or requiring jurisdictions within the state to collect data on law enforcement stops and searches. According to Northeastern University's Racial Profiling Data Collection Resource Center, "hundreds of law enforcement agencies" are now collecting such data (see <http://www.racialprofilinganalysis.neu.edu/index.php>).

### ***Early Warning Systems***

In addition to collecting racial profiling data, the US Department of Justice (DOJ), through the Violent Crime Control and Law Enforcement Act of 1994, imposed systems of accountability on law enforcement for "patterns or practices" of policing that violate the Constitution (Chanin, 2012). Since 1994 at least 25 law enforcement agencies have been investigated and/or reached settlements with DOJ regarding practices that violate citizens' federal rights. These agencies often agree to operate under consent decrees to improve policing practices to avoid further litigation. A common requirement of these consent decrees is the implementation of an early warning (EWS) or early intervention (EI) system designed to identify and intervene with officers at risk for future misconduct. EI systems are promoted as a specific example of practices designed to promote police

integrity (U.S. Department of Justice, 2001). One of the major assumptions behind these systems is “that a small number of police officers are responsible for a disproportionate amount of problematic police behavior. The long-term goal of such a system is to “create a culture of accountability in the agency” and thus reduce citizen complaints and litigation (USDOJ, 2001:10).

EI systems are expected to improve the capacity of police agencies to police themselves by concentrating their efforts on the people that generate the most problems. By generating ‘red flags’ on persons who engage in more misbehavior than their colleagues, police agencies can intervene more effectively in reducing that behavior. Evaluations of these systems have shown mixed results; some show success in reducing officer misconduct, while others raise questions about their efficacy.

An example of an early intervention system is that of TEAMS II in the Los Angeles Police Department. In TEAMS II (Training Evaluation and Management System 2nd phase), data are collected from 14 separate systems and compiled for all 10,000 sworn officers in the LAPD. This information is compiled in the Risk Management Information System (RMIS) and is used in two ways: employee performance assessments/continued monitoring and the automated risk management system. Like other EI systems, RMIS is designed to examine employee outcomes across five domains: use of force, citizen complaints, claims and lawsuits, preventable vehicle crashes, and vehicle pursuits. After one of these events has occurred, RMIS conducts an automated assessment of the officer’s performance against patterns established by his or her peers. These measures are then compared to the performance of the peer group

defined by similar work functions. An action item or AI is generated if the measure for the individual officer is substantially different from his/her peer group.

Once an action item is generated, the report is automatically forwarded to the employee's supervisor for review. The supervisor is then required to notify the employee of the action and to conduct a review of all relevant information to identify whether the employee's actions (both the current activity as well as prior activity) indicate a persistent pattern of at-risk behavior. If evidence suggests the possibility that at-risk behavior may be occurring, the supervisor engages in an investigation of the available information in the system reports. The supervisor documents the results of this examination by summarizing the various reports, determining whether there is a pattern of conduct, provides a comparison of the employee's performance to similar employees, explains any significant differences between the employee's performance and that of similar employees, selects and justifies a disposition (including a decision to take no action), and provides a brief summary of supervisor's discussion with the employee about this report. Cases that are canceled, judged to require no action, or end with informal counseling receive no further review up the chain of command, but those receiving other actions are reviewed up the chain to the commanding officer level.

A recent Office of Inspector General (OIG, 2012) audit found that TEAMS II corrective actions are rare. Only 64 of 1,384 AIs generated from August 1, 2010 - January 31, 2011 yielded any corrective action disposition. This audit also revealed that approximately 72% of the AI investigations were completed within 60 days of their initiation (OIG, 2012).

As part of its accountability process, LAPD has integrated the TEAMS II process into its Compstat process. This ensures that the middle and upper levels of LAPD's hierarchy are kept informed of progress in reducing undesired or questionable patterns of police practice and makes it more likely that the organization will attend to these aspects of agency, as well as individual officer performance.

### ***Studies of Early Warning/Intervention Systems***

Since 2000 a handful of studies have examined EI systems similar to the LAPD's. These include, a study of Pittsburgh (Davis, Henderson, Mandelstam, Ortiz, and Miller, 2002), a national evaluation and three case studies by the Police Executive Research Forum (PERF) and University of Nebraska Omaha (UNO) (Walker, Alpert, and Kenney, 2001), Worden, Kim, Harris, McGreevy, Catlin, & Schlieff's study (2012) in a northeastern city, and a study of a southeastern police department by Bazley, Mieczkowski, and Lersch (2009).

In 1997 the Pittsburgh Police Bureau became the first police organization to be issued a consent decree (Davis, et al 2002). The consent decree required implementation of an early warning system (EWS) and it stipulated both the specific information to be included (citizen complaints, training, transfers, arrests, etc.) as well as the categories that should be capable of being examined (individual officer, squad, shift, unit, etc.). This EWS was designed to compare officers to their peer groups using standard deviation calculations. In this way, for example, officers working the day shift in a low crime area were not directly compared to officers working the night shift in a high crime area.

The evaluation looked at the effects of the consent decree by conducting in-depth interviews, examining trends in departmental procedures, and administering a community

survey (Davis et al., 2002). The researchers did not directly analyze the effect of the EWS on officer performance, but instead reported on interviews with officers and supervisors and provided data on trends in police performance and disciplinary actions before and after the consent decree was issued. Overall the interviews showed a mix of views about the EWS. Some supervisors appreciated the system, while others said it did not provide any new information. Supervisors complained about the extra time and paperwork and said that they had less time to respond to serious incidents and to supervise their squads on the street. Officers voiced concerns about being afraid to do their jobs, stating that it was safer to do nothing.

Ultimately, the researchers concluded that “the early warning system ... all but ensures that officers heading for trouble will be identified and efforts made to straighten them out.” They also recognized problems like system redundancies, the extra time for filling out paperwork, and commented “most of the potential problem officers identified by the early warning system may already be known to supervisors” (Davis et al., 2002: 62). Lastly, they noted that in the absence of a more thorough methodology, it is difficult to determine what effects the EWS may have had on the performance, attitudes, and values of officers in the department and what effects can be attributed to the imposition of the consent decree.

Walker et al (2001) conducted a national survey of police agencies to determine the prevalence of early warning systems. They found that 27% of surveyed agencies had an EWS in 1999, with 12% planning to institute a system. Larger departments were more likely than smaller agencies to create an early warning system. The researchers also evaluated the effectiveness of three different early intervention systems: Miami-Dade,

Minneapolis, and New Orleans. They found that in all three departments the “program appeared to reduce problem behaviors significantly” and that they “encourage changes in the behavior of supervisors as well as of the identified officers” (Walker, et al., 2001).

More recently, Harris (2009a and 2010a) and Worden and his colleagues (Worden et al, 2012) examined officer misconduct and evaluated the effects of an early intervention system in a large police department in the northeastern U.S. In evaluating the overall effects of the EI, Worden et al. found that “rates of citizen complaints, personnel complaints more generally, and secondary arrests all declined.” Yet, matched controls in the department “exhibited very similar patterns, suggesting that these changes should not be attributed to the Officer-Citizen Interaction (OCI) School intervention”. They also found a small negative effect on arrests, particularly proactive arrest, stemming from the classes. Finally, they found a marked and problematic potential for false positives in early identification systems. If false positives occur at too high a rate, this could seriously jeopardize the efficacy of these systems as well as their legitimacy among officers and supervisors.

Bazley, et al (2009) examined use of force reports as an indicator for one large (over 1000 sworn officers) urban southeastern police department’s EI program. In 2000, when the study was conducted, 33 officers qualified for the program under use of force criteria. Departmental data were analyzed to determine whether either of their independent variables—total number of officer use of force reports and weighted officer force factor value (calculated by subtracting force level from resistance level)—had a statistically significant effect on officer qualification for the early intervention program. They found that both variables were statistically significant, but only the number of use

of force reports was in the direction expected (more use of force reports indicated a greater likelihood of qualification). Weighted officer force factor, on the other hand, indicated that the program actually “flagged officers who clearly displayed a tendency to use less force when confronting higher levels of resistance” (p. 119). Moreover, the system did not flag any of the 22 officers in the department with negative weighted force factor values that indicated a tendency “to use more force than resistance encountered” (p. 120). Ultimately, then, Bazley et al. state that “these findings would seem to bring into question whether the qualifying use of force criteria, i.e., tracking the numbers of higher level use for force incidents quarterly and annually, are actually identifying the problematic officers in this department” (p. 121).

### **Comprehensive Systems of Performance Measurement**

In the 21st century researchers and practitioners have sought appropriate and comprehensive measures for police performance. The search for meaningful measures included national organizations -- the Commission for Accreditation of Law Enforcement Agencies (CALEA) and the Police Executive Research Forum (PERF) working with researchers and police executives. Ultimately, they wanted to broaden the scope of police activities and goals and attached specific measures to them. The work of Maguire (2004), Moore and Braga (2003), and Milligan, Fridell and Taylor (2006) established the theoretical and research-based groundwork for CALEA and PERF. These works showed the multi-dimensionality of police work and demonstrated the need for multiple measures and diverse data collection methods and analyses.

Maguire (2004) traced the development of performance measures within the context of police organizations. He stressed the need for formulating broad dimensions

of police performance and the use of specific measures and tools. Citing previous research by O'Neill, Needle & Galvin (1980), Hatry, Blair, Fisk, Greiner, Hall, & Schaenman (1992), Moore (2003), and Mastrofski (1999), Maguire provides examples of the dimensions of police activities. In brief, they include such areas as crime prevention, crime control (apprehension of offenders), fear reduction and the feeling of security, fairness and courtesy, police administration (financial resources), use of force, and satisfaction of customer service. Measurement tools include the use of general community surveys, specific contact surveys, employee surveys, direct observations, and independent testing or simulations for collecting information about an agency's performance. Moreover, he included methods of analysis, weighting of measures, and a detailed section on how to implement comparative performance measures.

Similarly, at PERF Milligan et al (2006) identified seven dimensions of policing that warranted measurement: 1) community safety and security; 2) perceptions of safety and security; 3) confidence, trust and satisfaction; 4) respond to crime; 5) prevent crime, 6) enhance safety and security and 7) community health. The authors found that the Prince William County (VA) Police Department followed these dimensions and had also established appropriate performance measures.

Building on all of these studies, Davis et al (2010) and CALEA developed a pilot project to establish useful and useable performance measures. Researchers and a group of police executives drafted performance indicators to assess nine dimensions of policing: 1) delivering quality service; 2) fear, safety, and order; 3) ethics/values; 4) legitimacy/customer service; 5) organizational competence/commitment to high standards; 6)

reducing crime and victimization; 7) resource use; 8) responding to offenders; and 9) use of authority.

Each of these categories contained multiple indicators that applied to a spectrum of large and small agencies, police departments and sheriff's offices. Four CALEA-accredited agencies – Dallas, TX, Knoxville, TN, Kettering, OH and Broward County, FL developed and tested 28 specific measures that were part of the nine dimensions.

Data were collected using five instruments, including:

- ✓ *A self-assessment instrument* that asks agencies to provide traditional data -- basic crime rates, clearance rates, arrests, citizen complaints, and budgetary information.
- ✓ *Community surveys and business surveys* to gauge people's and local retailers' perceptions of police effectiveness;
- ✓ *Contact surveys* that assess the quality of interaction between police and citizens who request police assistance or who are stopped by the police; and
- ✓ *Officer surveys* that measure a broad array of perceptions of police including morale, integrity, agency leadership, and knowledge of policies.

After successfully implementing the pilot phase of the project five participants were added to the original four -- Raleigh, NC, Avon, CT, Boca Raton, FL and Las Vegas, NV Police Departments, and the Arapahoe County, CO Sheriff's Office.

The results from the analysis of the nine agencies show the value of this comprehensive set of measures. As the authors indicate, the set of measures or subsets can be used for a variety of purposes (with appropriate caveats). Some examples include:

- ✓ Comparing performance across patrol areas to determine “whether some areas stand apart from the others in how satisfied citizens are with the handling of a crime complaint or a traffic stop;
- ✓ Making year-to-year comparisons of agency of patrol area performance to determine if performance is trending in a positive direction; and
- ✓ Comparing performance with other agencies to provide a benchmark against which to assess how an agency is doing in various areas.

Lastly, the authors provide detailed information about their measures, particularly the survey instruments. They give specific instructions about implementing the surveys and describe their results in the nine agencies. Further, they provide the range of scores for the specific measures that they use. These explanations and tables serve as a guide as to whether an agency's performance on each dimension is within the bounds established by agencies involved in the field test. The value of the research allows another agency or researcher to use the data collection instruments and analyze the data appropriately. The results from an individual department can be compared to the nine CALEA agencies to obtain an overall context for its performance.

### **Key Issues/Controversies**

While a range of performance measures for police has been discussed, created, and implemented for over 150 years, adequate measures have yet to be fully agreed upon and validated by researchers and practitioners. This essay has attempted to identify those approaches and to describe research and evaluation that have been conducted on those measures. This section discusses key issues and on-going controversies about measurement.

### ***On traditional measures***

Traditional measures of arrests, crimes known, clearance rates, and response times have been criticized for their lack of comparability and validity. In addition, using crime rates as a measure of police success ignores other factors such as the social, economic, and political influences within jurisdictions. Reported crime rates also do not reflect the number of actual crimes that occur. Underreporting of crime is commonplace,

especially among high-risk populations, immigrant groups, those fearful of retaliation, and those who do not believe that police can do anything about a particular crime. Arrests and clearance rates are also problematic. Legal definitions of arrests vary across jurisdictions, making them difficult to compare. Clearance rates are indicators of investigative effectiveness, but often times can be manipulated and subject to measurement error (Cordner 1989).

### ***On Compstat***

As indicated above, Compstat has been a popular method as a management tool and to measure performance. From a measurement perspective, Compstat relies on the same crime data that have been criticized as inadequate measures of police performance for the past thirty years. The difference in their use today is that computer generated maps, trend analysis, and other visualizations assist decision makers in 'seeing' what is occurring on a week-to-week, month-to-month, and year-to-year basis.

Compstat has received accolades and criticism from various circles, including researchers and practitioners. Evaluations of Compstat focus primarily on its effectiveness (or lack thereof) regarding crime reduction (see Willis et al 2007). Evaluations of Compstat as a management tool are virtually non-existent.

### ***On Early Intervention Systems***

Early Warning/Intervention Systems focus primarily on officer misconduct, though the LAPD's system includes performance measures of employees. Researchers have found mixed results about the effectiveness of EWS/EI's in deterring that misconduct. Problems with the studies abound and are self-reported by the researchers themselves. For example, Walker et al could not determine the most effective aspects of

intervention or whether certain aspects are more effective for certain types of officers. This problem occurs with the studies by Davis et al and Bazley et al. Another problem is the inability of the early warning systems to collect data on positive police officer performance, so measuring the deterrent effect of an early intervention is difficult to assess. Third, researchers could not disentangle the effect of the general climate of rising standards of accountability on officer performance from the effect of the intervention program itself.

### ***On Comprehensive Performance Measurement Systems***

While researchers and practitioners recommend different designs for comprehensive performance measurement systems, the actual implementation of those systems is rare. O'Neill et al (1980) and Hatry et al (1992) suggested multiple dimensions and multiple measures, but those were not implemented. Maguire and Moore and Braga's systems were adopted in part by CALEA and used in nine agencies (Davis et al 2010). It remains to be seen whether more agencies will implement the system.

The CALEA/Davis et al approach (2010) is noteworthy for its use of multiple surveys and data instruments, but the difficulties lie in their actual use. A representative sample is necessary to obtain accurate perceptions of the jurisdiction and decisions need to be made about the sampling frame and sample size. Similarly, the use of contact surveys with victims of crime or those who were involuntarily stopped by police include problems with sample size, administering the survey, and completion rates. If these can be overcome, however, the information gathered from the surveys is particularly valuable for the police and the jurisdiction that they serve.

## **Future Directions**

Despite the problems associated with performance measurement, the future of this topic is bright. Police agencies are striving to become more efficient and effective in their work because of the budget crises and economic downturns in their jurisdictions. In addition, police executives are looking for ways to continue to reduce crime, particularly violent crime. Measuring performance and justifying or maintaining their budgets are important to these executives. As a result, using Compstat-like methods, obtaining feedback from residents and businesses, keeping track of officer performance, and continuing to make managers accountable for their areas, divisions, and precincts are methods that will continue to be used and expanded.

Technological developments will assist in the ease of use of measurement tools. Traditional data of crimes, arrests, clearance rates, and response times will be managed more efficiently because of improvements in computer systems and will be analyzed and visualized more quickly than before. Computer-generated "dashboards" that show charts, graphs, maps, and tables of crime are available now and will be used extensively by executives, commanders, and managers for a variety of purposes including performance measurement. Tablets, iPads, smart phones, software applications, and the Internet will enable agencies to conduct surveys of officers, citizens, and businesses on a routine basis.

## **Related Entries**

COMPSTAT

Control of Police Misconduct

Development of the UCR and the NCVS and their Contribution to Criminology

Evaluation Police Performance -- Methodological Issues

History of the Statistics of Crime and Criminal Justice

Measuring Police Officer Performance

Measuring Police Unit Performance

Police Use of Excessive Force

## Recommended Readings and References

Alpert, GP and Moore, MH (1993). Measuring Police Performance in the New Paradigm of Policing, in DiIulio, JJ, Alpert, GP, Moore MH, Cole, GC, Logan, C, Petersilia, J, & Wilson, JQ eds. *Performance Measures for the Criminal Justice System: Discussion Papers from the BJS-Princeton Project*. Washington, DC: US Department of Justice, Office of Justice Programs, and Bureau of Justice Statistics.

Bayley, DH (1994). *Police for the future*. New York: Oxford University Press.

Bazley, T, Mieczkowski, T, & Lersch, K (2009). Early Intervention Program Criteria: Evaluating Police Use of Force. *Justice Quarterly*, Volume 26, No.1, 107-124.

Bratton, W, with Knobler, P (1998). *Turnaround: How America's top cop reversed the crime epidemic*. New York: Random House.

Chanin, JM (2012). Negotiated justice? The legal, administrative, and policy implications of 'pattern or practice' police misconduct reform. Research report submitted to the U.S. Department of Justice.

Cordner, GW (1989). Police Agency Size and Investigative Effectiveness. *Journal of Criminal Justice*, 17(3), 145-155.

Davis, RC, Cordner, G, Hartley, C, Newell, R & Ortiz, C (2010). *Striving for Excellence: A Guidebook for Implementing Standardized Performance Measures for Law Enforcement Agencies*. Washington, DC: CALEA. Retrieved from <http://www.calea.org/content/striving-excellence-guidebook-implementing-standardized-performance-measures-law-enforcement>.

Davis, RC, Ortiz, CW, Henderson, NJ, Miller, J, & Massie, MK (2002). Turning necessity into virtue: Pittsburgh's experience with a federal consent decree. Vera Institute of Justice. Retrieved from <http://www.vera.org/content/turning-necessity-virtuepittsburghs-experience-federal-consent-decree>.

Eck, J & Maguire, ER (2000). Have Changes in Policing Reduced Violent Crime: An Assessment of the Evidence. In Alfred Blumstein and Joel Wallman (Eds.), *The Crime Drop in America* (pp. 207-265). New York: Cambridge University Press.

Harris, C (2010a). Problem officers? An analysis of problem behavior patterns from a large cohort. *Journal of Criminal Justice*, 38(2), 216-225.

Harris, C (2010b). Longitudinal patterns of internally generated complaints filed against a large cohort of police officers. *Policing & Society*, 20 (4), 401-415.

Hatry, HP, Blair, LH, Fisk, DM, Greiner, JM, Hall, JR, Jr., & Schaenman, PS (1992).

How effective are your community services? Procedures for measuring their quality (2<sup>nd</sup> ed.). Washington, DC: Urban Institute and International City/County Management Association.

Maguire, ER (2004). Measuring the Performance of Law Enforcement Agencies - Parts 1 and 2. *CALEA Update*.

Maguire, ER & Uchida, CD (2000). Measurement and Explanation in the Comparative Study of American Police Organizations. In Duffee, D (ed.), *Criminal Justice 2000: Vol 4. Measurement and Analysis of Crime and Justice*. Washington, DC: National Institute of Justice.

Mastrofski, SD (1999). Policing for People. Ideas in American Policing, Washington, D.C.: U.S. Department of Justice, Police Foundation (March).

McEwen, T (1996). *National data collection on police use of force*. NCJ 160113. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics and National Institute of Justice.

Monkkonen, E (1981). *Police in urban America: 1860–1920*. Cambridge: Cambridge University Press.

Moore, MH & Braga, A (2003). *The Bottom Line of Policing: What Citizens Should Value (and Measure!) in Police Performance*. Washington, DC: Police Executive Research Forum.

Moore, MH & Braga, A (2003). "Measuring and improving police performance: The lessons of Compstat and its progeny," *Policing: An International Journal of Police Strategies & Management*, Vol. 26, No. 3:439-453.

Office of the Inspector General (2012). "Training Evaluation and Management System (TEAMS) II audit, phase I. 2010-2011." Office of the Inspector General. Los Angeles Police Commission.

O'Neill, MW, JA Needle, & RT Galvin (1980). Appraising the performance of police agencies: The PPPM (Police Program Performance Measures) System. *Journal of Police Science and Administration*, Vol. 8, No. 3, 253-264.

President's Commission on Law Enforcement and Administration of Justice (1967). *Task Force report: The police*. Washington, D.C.: U.S. Government Printing Office.

Spelman, W, & Brown, D (1981). *Calling the police: Citizen reporting of serious crime*. Washington, DC: Police Executive Research Forum.

Uchida, CD. (2010). The development of the American police: An historical overview. In *Critical issues in policing: Contemporary readings*. 6<sup>th</sup> ed. Edited by Roger G. Dunham and Geoffrey P. Alpert. Prospect Heights, Illinois: Waveland Press.

United States Department of Justice (2001). "Principles for promoting police integrity: Examples of promising police practices and policies." Retrieved from <https://www.ncjrs.gov/pdffiles1/ojp/186189.pdf>

Walker, S, Alpert, GP & Kenney, DJ (2001). "Early Warning Systems: Responding to the Problem Police Officer." *Research in Brief*. Washington, DC: National Institute of Justice.

Weisburd, D, Mastrofski, SD, McNally, A, Greenspan, R, and Willis, JJ. (2001). *COMPSTAT and Organizational Change: Findings from a National Survey*. Washington, DC: Police Foundation.

Wellford, CF (1982). Redesigning the Uniform Crime Reports. *American Journal of Police* 1 (2): 76–93.

Worden, RE, Kim, M, Harris, C, McGreevy, M, Catlin, S, & Schlieff, S (2012). Intervention with problem officers. An impact evaluation of one agency's EIS. *Criminal Justice & Behavior*.