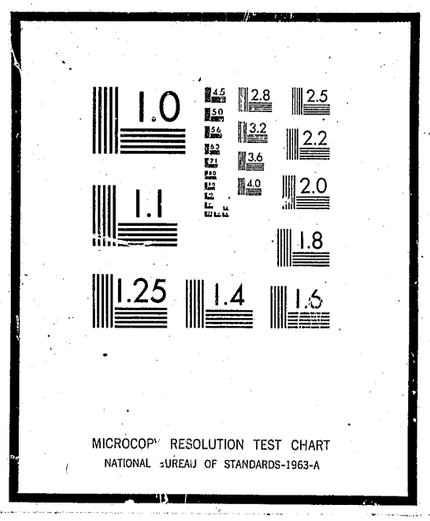

October 1974

WP-12-74

ALTERNATIVE MEASURES OF POLICE PERFORMANCE

by

GARY T. MARX

WORKING PAPER

"Innovative Resource Planning in Urban Public Safety Systems"

National Science Foundation Grant GI38004
Research Applied to National Needs
Division of Advanced Productivity, Research, and Technology

Operations Research Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Paper delivered at 1973 American Sociological Association meetings, New York. Work on this project began under the auspices of the Police Foundation. Subsequent work is being carried on as part of the "Innovative Resource Planning in Urban Public Safety Systems" project, under National Science Foundation Grant GI38004.

FOREWORD

The research project, "Innovative Resource Planning in Urban Public Safety Systems," is a multidisciplinary activity, supported by the National Science Foundation, and involving faculty and students from the M.I.T. Schools of Engineering, Architecture and Urban Planning, and Management. The administrative home for the project is the M.I.T. Operations Research Center. The research focuses on three areas: 1) evaluation criteria, 2) analytical tools, and 3) impacts upon traditional methods, standards, rules, and operating procedures. This report is associated primarily with category 1, in which current methodologies for measuring the performance of public safety systems are reviewed and new approaches explored. In this expository paper, Professor Marx sets in perspective many of the issues relating to police performance measurses. These include a categorization of currently used measures, measures on individual performance, the quality of emergency service, use of unobtrusive measures, the citizen survey, and measures for contrasting cities and the same department over time.

The work reported herein was supported by the National Science Foundation under grant GI38004.

Richard C. Larson
Principal Investigator

CONTENTS

Given the nature of their task and the means available to them in a civil libertarian society, police are bound to be more problematic than other public service institutions. But I think a strong case can be made for the argument that some problems that can be ameliorated stem from current performance evaluation practices, or perhaps better non-practices.

The public sector has generally lagged behind the private sector in developing measures of performance which would aide in evaluation and accountability. Their near monopoly, their multiple purposes and diffuse clientel, and the belief (unshaken until recently) that because their formal goals say they do good, they actually must do good, help account for this.

Social science understanding and better police performance require that we have accurate and varied indicators of performance for new recruits, individual patrolmen, and supervisors, as well as for precincts, specialized units and the department as a whole.

Such measures can indicate the extent of compatibility and trade-offs among various police goals. They can indicate the extent to which various segments of the population receive equivalent police service. They can help assess the consequences of particular programs and experiments.

They can identify areas where performance is particularly weak or strong. They can permit a more rational allocation of resources, and selection, training, placement and promotion procedures linked to actual needs. They can help establish equitable workloads among personnel. They can be important factors in developing a reward system more closely geared to the kinds of performance valued by police administrators. They can be factors in helping clarify what is expected of a police officer and the nature of the police role. It is important to provide men feedback on their performance, recognition for good work, and to communicate the seriousness with which

community service and order maintenance are viewed.

Yet few measures are available with which to make comparative ratings between departments or men, or the same departments over a period of time. The few measures that are available are used too uncritically. Police are of course not alone in this regard. It is only very recently, and partly as a result of challenges from protest groups, that public service bureaucracies have become seriously concerned with their effects and with measures of performance. The argument developed in this paper about the need for better police performance measures applied equally to other city services. Many of the same issues, problems, questions, processes and perhaps solutions, are involved, regardless of whether one considers schools, sanitation, hospitals, welfare, or police. There is much to be said for a comparative approach to the quality of publice service bureaucracies.* Yet because the police role is so crucial and the consequences of error or abuse so great, evaluation of police assumes a particularly important position.

Any discussion of measures of police performance must first consider the goals or ends of the organization. The more complex and varied an organization's goals, the more difficult finding widely agreed upon, specific operational indicators of performance is likely to be.

Any discussion of police goals brings one immediately to the question of the police role in American society--both what it is and what it should be. The value laden and relative nature of the issue is apparent. The strict civil libertarian is likely to see police effectiveness rather differently than the liquor store owner held up several times or the ghetto family concerned about their child becoming an addict. Much also depends

---

* For example, M. Lipsky, "Street Level Bureaucracy and the Analysis of Urban Reform," Urban Affairs Quarterly, June, 1971.

on the priorities of police and which social groups they are most responsive to. Yet even granting the heterogeniety of American society, it is possible to specify aspects of the police role about which there appears to be fairly widespread consensus. Among these are the following police goals identified by the American Bar Association Project on Standards for Criminal Justice:

1) To identify criminal offenders and criminal activity.
2) To reduce the opportunities for the commission of some crimes.
3) To create and maintain a feeling of security in the community.
4) To protect constitutional guarantees of freedom and equality.
5) To facilitate the movement of people and vehicles.
6) To assist those who cannot care for themselves or to arrange for such assistance.
7) To identify problems that are potentially dangerous.
8) To provide, on an emergency basis, services that police are peculiarly equipped to provide or to arrange for such services.

Current measures of performance focus much more on the crime related goals (1,2), than on goals of creating and maintaining a feeling of security, protecting constitutional guarantees, and various kinds of non-law enforcement emergency service. There is a need to develop performance measures in these areas as well.

This paper suggests a number of measures of police activity and output which might serve as variables for social scientists attempting to systematically contrast individuals, units, or cities, and which might also be used as tools in police planning and supervision. The paper does not argue for a detailed new scheme of performance evaluation, nor does it offer a strategy for implementation. These are of course both crucial to any kind of meaningful change. A first step however requires awareness of some of the major issues around police performance evaluation and awareness of the range of measures that might be considered. This paper offers such a general discussion.

It is purposely speculative. It seeks to generate ideas, rather than a fixed performance evaluation scheme. Given the rather restrictive nature of the measures currently in use, there is a need to raise issues and make suggestions which might be conducive to a re-orientation, or at least a broadening of traditional measures of effectiveness and evaluation. This refers to the areas measured, as well as to the techniques of measurement.

The paper first considers and criticizes current individual performance evaluation practices. It then suggests some ways of evaluating areas not now given sufficient attention, such as the use of force and the quality of emergency service rendered, and discusses using citizen interviews as a technique of evaluation. A final section focuses attention on some means of evaluating and contrasting departments rather than individuals.

Current Measures

Individual performance evaluation is generally not well developed in police departments. The most common practice consists of an annual or semi-annual subjective rating form filled out by police supervisors, where men are rated with respect to global categories such as initiative and appearance. Many departments make no effort to assess performance at all. Because of police sub-culturae values of self-protectiveness, and because little concrete depends on the evaluations, in many departments they have been abandoned, or become empty rituals where almost everyone's performance is rated as satisfactory.

Performance measures can be classified and contrasted in many ways. Table I lists some of these. Such a classification helps frame important questions for which systematic answers are now lacking. For example:

> What assumptions are made about the measures and the broader goals of the organization? How do different types of measure vary with respect to reliability and validity? What kind of measures can least easily be "captured" by those they are intended to evaluate? What are the strengths and weaknesses of various measures? How do various measures inter-relate?

Table I suggests some ways of characterizing police performance measures. The first entry at the left tends to be predominant. For example the measures now used by police tend to be internally rather than externally generated. They are done by supervisors rather than peers, self or clients. They often involve conformity to bureaucratic standards that are only indirectly related to the actual means and ends of police work. They are based much more on subjective assesments than on objective or quantifiable indicators, and when they are objective, ask how much, rather than how well. Current

evaluation instruments refer to general qualities, rather than to behavior in specific situations and if they do involve the latter, they refer to law enforcement, rather than community service or conflict related activities. Performance evaluation is used more to punish failure, than to reward success and to make inter-individual, rather than inter-unit comparisons.

Table I: Some Ways of Classifying Performance Evaluation

Source of information:	internal (Sargeants, self-reported rates, inspections)	external (Citizens praise or complaints, attitude surveys, judicial review (exclusionary rule, entrapment)
Means of evaluation:	supervisors - peers - self - citizens - single or multiple	
Relation to performance:	indirect (internal bureaucratic regulations such as those about appearance	direct (means or goals such as number of arrests, ambulance runs, use of force, citizen feelings of safety)
Type of data:	subjective - objective	
Emphasis:	production rates; goals (how much - number of arrests, tickets, field interrogations)	quality of the process; means (how well does person relate to the public, use force, render assistance)
Degree of specificity:	general qualities (integrity, appearance; this is often part of a city-wide form)	situationally or behaviorally specific (handling family disputes, use of police radio)
Substantive area:	law enforcement - community service - conflict management	
How used:	Punishment for failure (penalties for failure to live up to bureaucratic regulations or achieve expected production rates)	Reward for success (citations for heroism, day off for spectacular arrest, a factor in promotion or choice assignments)

Level:	individual (Number of field interrogation reports one man now writes)		unit (amount of crime or citizen attitudes in a precinct, aggregate number of field interrogation reports by all those in a precinct)
Frame of reference - is evaluation made relative to:	What others do - Comparing one man to an average for the department or all departments nationally	An Absolute Ideal Standard e.g., having every citizen report he feels safe or a rule about never shooting into a crowd.	Where an Individual is Expected to be Given his Experience and Potential Comparing a man to where he scored as a recruit

The factors on which men are rated may have little to do with what police actually do on patrol. As Egon Bittner notes "recognition is given for doing well in the department, not outside where all the real duties are."* This is related to the difficulty faced by police supervisors in evaluating their men which result from the decentralized nature of police work, a lack of clarity and conflict in police goals, and the intangible and symbolic nature of much of the police "product," particularly as it relates to deterrence. These factors, and a bureaucracy organized along quasi-military lines, result in evaluations often being based on conformity to internal bureaucratic standards, which may have little to do with how well a patrolman does his job on the street, or what he does. Being where one is supposed to be, or showing up to roll call on time with an immaculate uniform and shiny shoes may have precious little to do with a policeman's street sense, with how well he uses force, with his ability to calm tempers in a dispute, or to aide a troubled family gain the outside help it needs.

Where patrol activity is considered it is likely to be restricted to law enforcement. Two criticisms can be made here:

1) other important patrol activities are ignored

2) the indicators from which inferences about performance are drawn place undue reliance on a mechanistic tabulation of rates, rather than looking at the process though which the rates are created.

An evaluation of police patrol activity focusing only on law enforcement is too narrow to capture the diversity of the police job. As many studies have indicated only a minority of police time is spent dealing with serious crime and much evidence suggests that other general community factors are more important to crime rates than what police do. Apprehension is also

---

* E. Bittner, The Functions of Police in Modern Society. Government Printing Office, 1970.

more related to the type of offense than to actions of police. Much of the important emergency service and order maintenance work that police do goes unrecognized and unrewarded. The fact that such work, and not catching armed robbers, occupies most of the patrolman's time is ignored.

Focusing only on crime related measures may lead to the distribution of resources, or to practices which may interfere with obtaining other important police goals. It can encourage the mistaken view that crime is solely the responsibility of police, rather than of other parts of the criminal justice system and the public at large. Having police assume total responsibility for highly complex crime phenomena which they may be able to have only limited effect on, makes for police defensiveness and a degree of solidarity and isolation unbecoming a civic police.* We ironically see much attention paid to what police spend little time doing, and may not be able to greatly effect, and what police can have most effect on, and spend most of their time doing, all but ignored.

In the 1950's, the move toward a more bureaucratic police, which in some ways is the opposite of the professionalism those involved sought, meant an emphasis on rules, procedures, planning and record keeping. Objective measures productivity were sought as keys to performance. For some departments, particularly those on the West Coast, this gave rise to an emphasis on police productivity as determined by quantitative indicators. Production rates for patrolmen were determined by factors such as the number of traffic tickets written, arrests made, field interrogation cards filled out, stolen cars

* Even if the police crime role is seen as much more important than the order maintenance and social service aspects, it is possible that police effectiveness with respect to crime (and our measures of it) might improve were greater attention given to them. Community service activities can result in crime prevention. Improved police-community relations can lead to the greater citizen cooperation important in the apprehension of serious law breakers. Increased citizen confidence in police might lead to an increase in reported crime, and were crime given less internal priority, there would be less incentive for distortion of crime statistics, both resulting in better data on crime.

identified, or for detectives, crimes cleared by arrest, percent convicted, and stolen property recovered. Where these factors are not a part of the formal evaluation system, Sargeants may nevertheless have expectations regarding an informal quota system. The emphasis too easily is put on rates of production, rather than on the quality of the process through which the rates are produced. The question "how many arrests or tickets," is asked rather than "was it wise to write a ticket, or make an arrest in this context, or was the law followed in the process."

The substance of the area evaluated aside, a rather limited range of evaluation techniques are used. As noted the prime evaluation technique is internally generated--a supervisor's subjective rating. Many problems may arise when a sensitive function such as policing is only subject to internal evaluation. These include the exclusion of important alternative sources of information and definitions of problems, the tendency of an agency to protect its own and perpetuate things as they are, the ease with which self-measurements can be manipulated and lack of public confidence in the process.

Another problem lies in the use made of performance evaluation data. Police department's generally tend to put greater emphasis on punishing failure than on rewarding success. Organizations generally find it easier to do this (indeed success often becomes defined simply as the absence of failure). The failure to live up to a standard is easier to see than going far beyond the standard. Norms which have a continuous quality (rather than merely declaring "thou shall" or "shall not") and involve degree, are more conducive to rewarding success. Police supervisors tend to be seen as people who can make trouble for you, rather than people who will reward you for a job well done.

With the strong job protections offered by civil service, promotion to higher ranks based on memorization of laws and police solidarity, there has been little incentive to develop broadly based performance evaluation systems.

There are also many technical problems noted in the literature such as: response set, varying standards and frames of reference among supervisors, lack of testing for reliability and validity, and supervisor's indifference, lack of knowledge, bias, or hesitancy to criticize other police. Appropriate training and supervision of evaluators, the use of multiple measures from various sources, and more clearly relating performance evaluation to the reward system, can help minimize some of the above.

There is a need to develop indicators for areas other than law enforcement, indicators which tell us about the quality of performance beyond sheer quantity, and techniques of evaluation which go beyond the subjective rating of a supervisor.

The discussion which follows considers some ways that these needs might be met. Three generally neglected substantive areas on which individual patrolmen should be evaluated are discussed: the use of force, arrest and civil liberties, and the quality of emergency service. An external evaluation technique, the citizen interview is also considered.

A. Measures of Individual Performance

The Use of Force and Arrest

An important aspect of the police role is to protect constitutional guarantees of freedom and equality. One component of this has to do with enforcing laws regarding freedom of speech and assembly. Yet another has to do with the means through which police obtain their goals. Police in a democratic society are differentiated from those elsewhere, by the often severe limitations they face in carrying out their goals. Probably to a greater extent than with most other agencies, the evaluation of police performance requires attention to means. Indeed police operating within the framework of the United States Constitution should be a goal in itself. A spectacular arrest achieved through physical coercion and illegal surveillance or search and seizure is a questionable gain. While apprehension of a suspected felon clearly meets one police goal, the use of such means can be seen to clearly violate another. The evaluation of police performance should be able to take account of both of these. As noted, traditional measures have focused much more on outcomes, than on the process whereby these were created.

One area where such "means" measures can be developed is with respect to police use of force. Central to the police role is the right to use force, whether this be in apprehending a felon, negotiating a family dispute, or assisting with the mentally ill, or at the scene of a traffic accident. However, measures for indicating how effectively force is used, or what constitutes the undue and unnecessary use of force, tend to be lacking. A number of gross indicators of the use of force are available, which if used jointly over a period of time would offer a systematic way of measuring this.

These indicators include the proportion of an officer's arrests in which charges of assaulting or interfering with an officer, or resisting arrest, or disorderly conduct are brought, the proportion of arrests involving the use of force, the number of times a baton or gun are drawn and used, the proportion of arrests involving by-standers as against suspected offenders, or those responsible for the initial police involvement, and the extent of injury and homicide involved in police-citizen encounters.

From considering the nature of the arrest, the actions of the person arrested, and the context, the seriousness of the threat to police and others, alternatives available to the use of force, the amount of force used by police and extent of injury and damage, judgements could be reached about how well force was used. Given the complexity of most situations and difficulties of judgement, evaluations of individuals (except in extreme cases) would be based on a number of police-citizen encounters over a period of time, such as each six months. Some cities such as Oakland have started a peer review panel with respect to the use of force. Statistics are collected on the use of force and officers using force disproportionately are reviewed.

Another important "means" area is the quality of arrest. Much work and experimentation would be required to develop appropriate measures here. One area to begin with would be some of the actions taken by prosecutors and the courts. These can be seen as rough measures of the extent of police conformity to law, or thoroughness of preparation in arrest situations. These include the proportion of those arrested who are charged, the proportion of cases actually coming to trial, and those dismissed as a result of the exclusionary rule, entrapment, or inadequate preparation. Several large cities are now monitoring how District Attorneys and Courts process their arrestees

and how well cases are prepared. In some cities a superior officer observes police testimony. There are, of course, many intermediary factors between actions of the police and judicial outcomes, and inferences drawn must be cautious. But as a rough comparative guide, in conjunction with other measures, this might be one way to deal with the "means" aspect of police performance.

Police knowledge of civil liberties and civil rights, and recent court decisions, might be periodically measured by paper and pencil means, or through role playing. While knowing what the law is, and what procedures are allowable, is no guarantee they will be followed, such information is at least a necessary pre-condition for police conformity with law.

The Quality of Emergency Service

A number of studies have shown that a majority of police interaction with citizens involves some type of emergency service--whether it be assisting at a traffic accident, returning a lost child, or arbitrating a family dispute. Yet such activities seem invisible to the public. They are not part of the TV shaped image of what police do, nor are they taken seriously enough by many police. It is ironic that those areas where police are most directly helpful to people and probably are most effective (compared at least to effectiveness against crime) are those least rewarded and most hidden from public view.

Discussion of whether police should, or should not, be dispensing emergency services often seem unduly academic. One argument for their involvement is that in emergency situations the coercive power of police is occasionally needed to force help, or gain compliance, from the non-cooperative. But beyond this, in the most rational and organized of societies with unlimited resources, perhaps the police might be restricted to violations of the law where the use of force is relevant to deterence and apprehension. Yet we are rather far from such a state of affairs. By default, police are left with many emergency service tasks and conflict situations. There is a major city-wide 24-hour need here not met by other public service agencies or helping professionals, who generally sit and wait for those in need of help to come to them. Police are in a unique position not shared by other professionals. Being in homes and neighborhoods where the need for help is often great, and where awareness of potential sources of help is often limited, they can play a crucial role in identifying problems and linking those in need of help with appropriate agencies.

Public confidence in, and satisfaction with police might increase were such services given more explicit formal recognition. If chances for advancement and more favorable assignments depended partly on how well such service tasks were carried out, they might be done more conscientiously.

As a result of the extreme demands on their time, limited training, and the diversity of situations they must deal with, most police cannot be expected to fully do the job of psychiatrists, legal advocates, employment counselors, doctors, etc. But they can offer support on an emergency basis and link people with difficulties with those in a position to give help on a more sustained basis, and at the very least their intervention should not exacerbate the problem.

Given the variety of service tasks police face and the need for quick action in unsupervised settings, finding measures of performance is difficult. Yet to begin with, a simple counting of such service would bring greater understanding of it and attention to it. We should be able to know how departments and people within a given department differ from each other with respect to the amount and kinds of service rendered.

A number of means, including supervisor's review of service activity, follow-up interviews with a sample of those having contact with police, and periodic assement of patrolmen's awareness of referral resources, and perhaps response to simulated service situations could be used.

Issuing a receipt, or official notice after such police encounters with citizens, as is done with traffic tickets, would give greater visibility and attention to this function and make tabulation of such service contacts easier.*

---

* For example see the suggestion by A. Reiss, Police and Public, Yale University Press, New Haven, 1971.

Evaluation efforts could be directed at how well the symptoms of the problem in the immediate situation were handled. This would include things such as the quality of the inter-action with citizens, expressed interest in the persons' problems, and the type of action taken or suggested by the patrolman (e.g., extent and quality of referrals made, whether assaults and certain types of arrest were avoided).

Taking or encouraging action which helps get at the source of a problem such as (unemployment, alcoholism, housing code violations) in addition to dealing with the symptoms (angry people) should also be a positive factor in performance evaluation. For example, in a landlord-tenant dispute over a rent increase, and the failure of the landlord to fix plumbing, and faulty exposed electrical wiring, does the policeman simply leave after indicating that this is a civil matter over which he has no authority, or does he refer the tenant to a legal aide society or rent control board and notify the appropriate city agency about building code violations? In a family conflict does the policeman threaten to arrest an unemployed drunken husband if he doesn't leave the house and go for a walk, or does he also tell the family about community agencies that can offer help with employment, alcoholism, and family relations and notify a given agency.

Additional factors in evaluating service might be the ratio of situations initiated by the dispatcher vs. the patrolman, and a patrolman's knowledge of a community's helping resources and procedures for drawing on these such as better business bureaus, mental health clinics, legal aide service, charitable organizations, government inspection and regulatory agencies, etc.

Discovering Unobtrusive Measures

In addition, there are no doubt many subtle indicators lost to the outsider which experienced police supervisors use in evaluating (or at least in forming impressions of) their men. It would be worth trying to discover some of these. Just as a sergeant may know who the more aggressive men are with respect to crime, he may also know which men are most effective with respect to community service and order maintenance, but how does he know this? What kinds of indicators tell him that some men have more patience, tolerance, empathy, understanding or coolness under pressure than others? Which men are though to relate best to minorities, deviants and those with atypical political beliefs, have the ability to resolve disputes and make arrests with a minimum of force, and are most helpful in referring people with problems and needs to appropriate agencies? How does type of organization and training effect the above? On what grounds is it thought that a man would be good for community relations or juvenile or crime prevention work? Does knowing that a man's uniform is in order, that he is infrequently late or absent, that he writes a large number of traffic tickets and field interrogation reports say anything about the above, or are other indicators more appropriate?

Experienced police supervisors no doubt make use of many obtrusive and unobtrusive indicators of personnel characteristics, though these are not all at a conscious level, highly codified or a part of the formal evaluation process. It would be useful to try and identify some of these indicators and means of codifying them, as well as the organizational contexts which illicit such behavior.

An External Indicator: The Citizen Survey

Thus far we have considered areas important to police work that have not received sufficient attention with respect to evaluation. We now turn to one technique of evaluation that would be useful to the above, as well as to other areas of performance. This technique is externally generated. Rather than the direct actions of police, it involves the attitudes, experiences and behavior of the consumers of police service, whether this be those arrested, those victimized, those calling police in social service and order maintenance situations, or the public at large.

One of the unfortunate aspects of professionalism and bureaucratization is that the client is often ignored or taken for granted. All expertise and decision making rest with the professional. Communication too often goes exclusively downward. Unlike the case of business, the performance of a public service agency cannot for long be judged apart from the citizens it serves, however varied these citizens may be. External measures in the form of citizen feedback can help departments better understand what citizens see as their needs, what priorities they have, what experiences they have had in police encounters, and how they view particular situations.

Such feedback can better gear service to needs, help breakdown the gap between citizens and police, give administrators a broader picture of performance, and one which is more difficult to manipulate than internally generated performance measures.

Surveys can give a voice to those who are usually unheard and can reveal different attitudes and needs of various segments of the community. For example, the needs of the aged are usually not a direct factor in setting policy, yet surveys reveal the greatest concern about safety and

wish for increased police services among this group. It is true some people may misperceive, consciously distort or value practices that are illegal and wrong. Yet this is not an argument for ignoring public attitudes. What people think and believe is perhaps more crucial--than what really is. It is vital to understand public attitudes, to educate the public to what police can and cannot legally and realistically do, and to gear service to felt needs.

Baltimore and several other departments have begun random follow up interviews with a small proportion of those calling police for service. Citizens are asked to rank police on things such as courtesy, understanding, capability, explanations of their actions, whether police were thought to have done everything they could to handle the problem and whether they would call police again if a similar problem arose.*

Another device is to have the internal investigation unit play a pro-active rather than a re-active role, by asking citizens on a random basis about the performance of specific officers. Citizens are often hesitant to make complaints because they don't know the procedure, fear retaliation, or feel that it does no good to express a grievance. It would seem very different when an organization is open enough and has the courage and confidence in itself to not simply wait for citizens to complain, but to actively seek out their opinions about the service it provides and experiences they have had with it. This fact alone might increase public confidence and satisfaction independent of organizational and operational changes.

---

* This is described in F. Furstenberg and C. Wellford, "Calling the Police: The Evaluation of Police Service," Law and Society Review. Those interviewed are told "every day a large number of people call the police department for service. Routinely, we follow up some of these calls in order to check on the quality of the service that was provided. Our records indicate that you requested assistance from the police department recently. If I may, I would like to ask you a few questions about the service that was provided."

Such pro-active investigations would apply best to men on foot, but could also apply where motorized patrolmen have a regular beat. While asking 20 people of varied background (merchants, youth, clergy, etc.) impressions about a man might be expected to turn up some complaints on almost anyone of a personal or idiosyncratic nature, a consistant pattern of negative responses (perhaps at several points in time) would be more meaningful. Knowing that this was being done might also effect the performance of some patrolmen.

A patrolman in a city close to Boston was recently involved in an altercation with a youth who later died. Residents and merchants had long been very disappointed with this man's performance, yet this was unknown to his supervisors. If this had been known through the pro-active investigation suggested above the man might have been taken off the street and the incident avoided.

B. Some Measures for Constrasting Cities and the Same Department Over Time

Let us turn from measures that can be used for evaluating individual patrolmen to the organizational level. The focus here is on a total Department (or units within it), or on one city as against another.

The individual measures discussed can of course be aggregated to form group measures. Thus the ratio of disorderly or resisting arrest charges to total arrests can be determined for an entire department as well as for patrolmen. Other measures relate best to the police department as a unit (or geographical segments of it) and can not be usefully disaggregated to the level of the individual patrolmen. This is the case for crime rates and general citizen attitudes about crime and police.

Police can now give limited answers to questions such as the extent to which various segments of the community (either geographical or social such as race, age, class, etc.) receive equivalent police service, or have equivalent needs for police service. Decisions about resouce allocation and tactics, and understanding the consequences of new forms of organization and operation, require a better array of information than is currently used by most departments.

Another phase of this NSF project is considering measures of an organizational nature for the patrol force. Various measures such as response time, patrol frequency, and internal work load are among factors considered.* This section considers some additional measures of an organizational nature.

---

* Publication P1 Police Performance Measures Literature Survey.

Homicides by Police

Another measure of use of force which is probably best treated as an aggregate figure at the group level, is homicides by police while on duty. There appears to be tremendous variation between departments, and no doubt within them. For example, between 1950 and 1960 the "justifiable homicide" rate per 10,000 officers was about 45 times as great in Akron, Ohio, as in Boston (Table 1).

Rates for some other cities are shown in Table 2. On the other hand, there was greater equality between races with respect to a citizen's chances of being a homicide victim through intervention of police in Akron than in Boston, though there is still great variation between racial groups. Thus, in Akron, the Negro to White ratio is 5.8 to 1 (rates of 16.1 for Negroes and 2.7 for Whites per 1,000,000 population) while in Boston a Negro had a 25.2 times better chance of being killed by a policemen than did a white (rates of 3.2 for Negroes and .1 for Whites per 100,000 population). (Table 2.) Patterns of justifiable police homicide are partly due to regional and sub-cultural variations in the resort to violence, crime patterns, the availability of weapons, and differences in reporting. These figures would be clearer if they were based on extent and type of contact with police. But it is partly a consequence of training, supervision and departmental policies. Such rates should be computed on a regular basis, many of such justifiable homicides while within the framework of the law, are unnecessary and represent poor judgement on the part of the officer. Information on number of police homicides by city is collected annually and an aggregate national figure is published in the United States Public Health Department's Vital Statistics. But data is not presented by city. If it were made public by city it might serve as a strong incentive for the more careful use of force.

Table I

Rates of Justifiable Homicides by Police Officers per 10,000 Officers - 1950-1960*

City	Annual Rate per 10,000 Officers
Boston	1.05
Buffalo	4.76
Milwaukee	5.50
Philadelphia	6.08
Washington, D.C.	10.65
Chicago	22.53
Cincinnati	24.82
Kansas City, Mo.	35.41
Miami	38.15
Akron	48.50

* From G. Robin, "Justifiable Homicide by Police Officers," Journal of Criminal Law, Criminology and Police Science, Vol. 54, 1963.

Table II

Negro and White Justifiable Homicide Rates by City 1950-1960*

	Rates per 1,000,000 Population		
	Negro	White	Negro: White Ratio
Akron	16.1	2.7	5.8 to 1
Chicago	16.1	2.1	7.4 to 1
Kansas City, Mo.	17	2.2	7.5 to 1
Miami	24.2	2.7	8.8 to 1
Buffalo	7.1	.5	12.2 to 1
Philadelphia	5.4	.2	21.9 to 1
Boston	3.2	.1	25.2 to 1
Milwaukee	13.5	.4	29.5 to 1

* From G. Robin, "Justifiable Homicide by Police Officers," Journal of Criminal Law, Criminology and Police Science, Vol. 54, 1963.

Citizen Attitudes and Behavior

Periodic surveys of the public at large (rather than only of those having a direct encounter with police) are also an important means. Such surveys can measure general attitudes and images of police, feelings of security, felt needs, satisfaction and dissatisfaction with various aspects of police service, willingness to report various kinds of crime, to testify in court and serve on jurys, and knowledge of the law, civil liberties, and police powers. Differences between groups and changes over time could all help in evaluating performance and guiding policy. We have many such surveys but they are done on an ad hoc basis and usually in no way figure into departmental policy.

The means are readily at hand for systematically asking people about police and city services generally. For example, the HUD Urban Observatory Program asked a large sample of citizens in 10 cities their beliefs regarding an array of public services--from garbage collection to schools, to police. In the case of the latter, for example, representative samples were asked how fast police were thought to respond, how effective they were in controlling crime, what the most serious law enforcement problems were, whether the enforcement of traffic regulations should be stricter, whether the person questioned had recently been victimized and how safe he or she felt. Various indices based on such measures can be constructed and the same city compared at two points in time, such as before and after a major change in policy.

Such a comparative approach also can yield interesting results between cities and can be a cue to the effect of various forms of police organization and policy. For example, Tables 3A and B show citizen's feelings of safety and satisfaction with police service for each of the 10 Urban Observa-

tory cities for both blacks and whites.* There is a noticeable variation between these cities. The variation of a racial nature is particularly interesting. For example, blacks in Kansas City are more satisfied with the quality of their police service than are whites in Boston and Baltimore. Kansas City also is one of the most highly professionalized American police departments.

When the 10 cities were ranked in terms of their degree of police professionalism, the more professional the department, the greater the degree of citizen satisfaction.** This runs contrary to much current thinking about police, wherein the supposed greater universalism, efficiency, aggressiveness, bureaucratization and centralization of the more professional departments is thought to produce poorer police-community relations. Such a pattern must be interpreted with caution since the number of cases is small, but it does suggest the kinds of questions that can be partly answered when data on citizen attitudes is considered.***

---

* Data for blacks are shown only for seven cities since the number of cases in the other three cities was too small for valid statistical inference. A fuller account of this data is given in G. Marx, "On the Inter-Relationship of Police Organization and Operations, Crime, City Characteristics, and Citizen Attitudes: Data from 10 Cities," paper presented at the annual meetings of the American Political Science Association; 1972.

** The professionalism measure is based on things such as degree of centralization, ratio of supervisors to men and car to foot patrols, and selectiveness of recruitment.

*** Among other questions this comparative attitude data permitted analyzing were: Are citizen satisfaction and feelings of safety higher in cities where there are more footpatrolmen and local precincts, greater per capita expenditures on police and more police, and in cities with "reform" governments? The answer to the above, at least for the 10 cities under study, was "no." In fact, the reverse was true. Beyond the professionalism measures mentioned above, citizen satisfaction with police increased as did the percent of civilians and minority group members on the force, the amount of police community relations training, and size of the juvenile division.

Table 3A

Feelings of Safety by City and Race*

Whites	Mean Score	% very safe	% very unsafe
Albuquerque	2.02	32	10
San Diego	1.95	30	11
Milwaukee	1.70	24	18
Atlanta	1.61	22	22
Nashville	1.60	20	23
Kansas City, Missouri	1.49	17	25
Denver	1.48	16	25
Kansas City, Kansas	1.39	9	26
Baltimore	1.32	13	30
Boston	1.26	12	34
Spanish speaking			
San Diego	1.92	31	12
Albuquerque	1.81	25	10
Denver	1.11	12	41

Blacks	Mean Score	% very safe	% very unsafe
Milwaukee	1.45	18	31
Nashville	1.31	7	23
Kansas City, Kansas	1.17	11	34
Atlanta	1.12	8	38
Baltimore	0.97	6	42
Boston	0.68	4	60
Kansas City, Missouri	0.65	8	65

* From G. Marx, "On the Inter-Relations of Police Organization and Operations, City Characteristics, Crime and Citizen Attitudes: Data From 10 Cities," paper delivered at 1972 American Political Science Association Meetings.

Table 3B

Satisfaction with Police by City and Race*

Whites	Mean Score	% very satisfied	% dis-satisfied	
San Diego	7.31	37	5	(325)
Alburquerque	7.16	31	10	(223)
Milwaukee	7.14	33	7	(299)
Denver	7.14	33	7	(193)
Kansas City, Missouri	7.03	26	9	(215)
Kansas City, Kansas	6.99	29	10	(105)
Baltimore	6.82	21	15	(220)
Atlanta	6.60	21	14	(166)
Nashville	6.56	14	12	(224)
Boston	5.80	18	21	(277)
Spanish speaking				
San Diego	6.94	21	5	(19)
Alburquerque	6.14	21	17	(96)
Denver	5.25	22	23	(27)

Blacks	Mean Score	% very satisfied	% dis-satisfied	
Kansas City, Missouri	5.96	16	21	(62)
Kansas City, Kansas	5.21	16	37	(38)
Baltimore	5.04	7	38	(255)
Milwaukee	4.68	19	42	(47)
Nashville	4.58	6	44	(51)
Atlanta	4.49	8	41	(159)
Boston	3.35	5	62	(79)

* Based on a three item index

Considering just the area of crime, surveys and observation can be very useful in giving police an idea of how secure citizens feel. An argument can be made that one of the most essential services police provide is a feeling of security. Any evaluation dealing with crime must consider subjective feelings and perceptions as well as the actual "objective" facts of crime. In City A for example, if crime is relatively low compared to other cities, yet citizens have a high degree of fear, certain actions of an educational or symbolic nature may be called for on the part of police. If people in certain contexts have a high chance of being victimized yet seem unaware of this, a different type of educative action is required.

Periodic surveys could ask how secure people feel, how concerned they are about crime, how worried about personal victimization, what security precautions they take, how they may have altered their behavior in response to crime, how and when they are in public, and about recent victimization experiences of their own, of others they know, or instances of victimization they have heard about in their neighborhood or city. The same people could be interviewed each six months for several years, or different people interviewed, with an emphasis on retrospective questions asking them whether they feel more, or less, safe walking in their neighborhoods than at a previous point in time.

Less obtrusive measures might also be used such as the number of people out on the street at night in a given area, as measured each six months (controlling for density, proximity of area to commercial and entertainment centers, availability of public transportation, etc.) or increase or decrease in the sale of weapons, watch dogs, property insurance, special locks and alarm systems, the use of private guards, letters to newspapers about crime, community police patrols, etc.

While any single measure would have to be used with the greatest caution, a consistent pattern of increase or decrease in a number of these measures over a six month period would permit reasonable inferences about citizen feelings of security. A weighted composite could be developed and standardized which would permit contrasts within as well as between cities.*

Just how closely such an index of citizen feelings of security is related to police behavior could be discovered by altering police practices (increased patrols, more men walking, or giving more publicity to police success with crime, crime prevention programs, etc.) in equivalent areas, or in the same area in different time periods, and relating these to changes in a composite measure of citizen feelings of security.

It is important to discover how citizen feelings of security, victimization rates and experiences, reported crime, and various police practices interrelate. Considering these factors <u>together</u> gives a far richer and more meaningful assessment than the number of crimes known to police or cleared by arrest or a victimization survey considered separately. One can look at absolute increases and decreases in crime known to police, and in victimization and the ratio between them. For example, if the amount of reported crime is considered in ratio to the amount of actual perceived crime, as measured by a victimizaiton survey, we have the hypothetical framework shown in Figure 1. If such measures are taken at two points in time, such as each six months, a number of outcomes are possible each implying a different evaluation of performance and the need for different policies.

---

* Such a measure would be crude, especially initially and a great many methodological criticisms could be directed at it, denial or inability to remember victimization, saturation points in possession of weapons, insurance and locks might be reached, factors such as the state of the economy, media treatment of crime, aggressive business practices of private security industry, etc., might distort the meaning of such measures. Yet whatever their limitations it is important to go beyond reported crime or clearance rates and to consider citizen feelings and actions.

Over a six month period the ideal and probably rarely reached state of affairs might be that shown in Figure 1 by two asterisks. Here the overall amount of crime known to police declines, reported victimization declines, yet the proportion of crimes known to police (considering victimization surveys) increases. From this one would have reasonable grounds for inferring that crime has actually decreased while citizen confidence in the ability of police to deal with crime and intolerance of crime has increased. The least desirable state of affairs might be the opposite--the amount of crime known to police increases, reported victimization increases and the proportion of crimes known to police decreases. While in reality such a clear pattern might not often emerge, such a table indicates the broader picture possible when both objective and subjective measures of police performance are considered. An even richer, if more complex, measure emerges when other of the safety measures discussed above are considered.

Actual public behavior, rather than expressed attitudes as such can also be considered. Various measures of citizen cooperation or hinderance could be developed and considered at specified time periods, such as the quantity and quality of information about serious crimes from the general public, damage to police property and false alarms, attacks on police, or citizens assisting officers in need of help, the proportion of those sought as witnesses who agree to testify. Other external measures such as the number and nature of complaints and compliments about police made to police themselves, city government, community groups, and the media could be used.

Figure 1

Hypothetical Relationships For Two Time Periods Between Crimes Known to Police and Crime as Revealed by Victimization Surveys

	Time Period I		
	(A)	(B)	(C)
	Crimes Known to Police	Crimes Revealed by Victimization Surveys	Ratio of Crime Known to Police/ Crimes Revealed by Victimization Surveys

	Time Period II		
	(A)	(B)	(C)
	Crimes Known to Police	Crimes Revealed by Victimization Surveys	Ratio of Crime Known to Police/ Crimes Revealed by Victimization Surveys
Some possible changes between Time Periods I and II			
Most desirable pattern of change	A) decreases	decreases	increases
Least desirable pattern of change	B) increases	increases	decreases
	C) same	same	same

Some Composite Measures for Constrasting Departments

Numerous other measures might be developed, particularly to contrast police units with each other, or changes in a given unit over time. Two composite measures worthy of note might be (1) departmental morale and (2) community relatedness.

A composite measure of departmental morale might be systematically taken based on rates of police turnover, requests for transfer, mental and physical health, absenteeism, suicide, divorce, alcoholism, various attitude measures of morale, patrol car accidents and damage to and loss of police property.

In more or less homogeneous areas of a city, a community relatedness measure might be based on knowledge of client population's culture, language, jargon, living in neighborhood or being involved in its voluntary organizations, attitudes of police toward people they serve, extent to which they know the names of people in their area (other than of merchants or criminals) or spend their leisure time with non-police friends, number of civilians and women in the department; minimum physical requirements to join the department (height, weight, glasses), and proportion of force in the field having contact with the public walking or on bicycle, motor-cycles or scooter.
The consequences of a department having a high degree of community relatedness and a high morale for various alternative and traditional measures of performance would have to be empirically demonstrated. But its seems reasonable to predict that departments high on these factors would, on most measures, perform appreciably better. With respect to community relatedness, this was the case for citizen satisfaction and feelings of safety with the 10 city data mentioned earlier. Indeed our beliefs about the advantages of a civil police force are based on such assumptions.

Conclusion

It is of course one thing to point to performance measures that might be used, and quite another to have them adopted. The kinds of measures emphasized are closely related to the goals of the police organization. Until equal, or at least much greater recognition is given to police tasks involving emergency service, conflict management, the quality of inter-action with the public, and the "means" aspects of police work, criteria measuring effectiveness in these areas are not likely to receive the attention they deserve. Even if the emphasis on performance evaluation was broadened, it might mean little until performance evaluation became more closely tied into the rewards system. Even where police leaders wish to see such changes, traditionalism, civil service regulations and employee associations may make implementation difficult. Resistance may be minimized, at least initially, by gathering information about departments as a whole, or various sub-units, over several time periods with data aggregated from individuals, rather than focusing directly on the evaluation of individuals. In the beginning, making new measures part of a general inspectional process, (rather than something that goes into a personal file) would permit familiarity with new measures to develop and some test of their practicality, validity and reliability. It would place emphasis on changes in organizational practice conducive to poor performance rather than on punishing individuals as such.

There is obviously no simple answer to reforms involving police and various trade-offs are present. One approach which could be called "the kinds of people approach" is concerned with getting better educated police, more minorities and those psychologically suited in exercising authority.

Another approach focuses on altering the police role--restricting it to law enforcement, dropping concern with vice crimes, or developing specialized for greater decentralization or regionalization. Still other approaches focus on altering authority patterns by increasing the power of the Mayor, city council, or citizen groups against that of the Chief or patrolmen's association and by changing civil service requirements. Another approach is more technical and involves computers, helicopters, and sophisticated crime labs. One large West Coast department is even rumored to have a sub-marine.

Yet still another aspect of police change has to do with the reward structure. Without conjuring up wooden images of a reward seeking, punishment avoiding man, if one wishes the restrained use of force, greater police conformity to law, better community relations, and more effective police behavior in conflict and helping situations. It is important to structure the job to measure and reward such behavior. Those concerned with social change and understanding inequality in American life could usefully focus more attention on micro-level issues involving selection, training, performance evaluation, civil service, public employee associations, promotion, sanctioning, client-bureaucrat inter-action, and policy formation within police, as well as other public service bureaucracies.

Selected References

Bard, M., Training Police as Specialists in Family Crisis Intervention, Washington, D.C., GPO, 1970.

Bittner, E., The Functions of Police in Modern Society, Washington, D.C., GPO, 1971.

Baehr, M., Furcon, J., and Froemen, E., Psychological Assessment of Patrolman Qualifications in Relation to Field Performance, Washington, D.C., GPO, 1969.

Bardua, D., and Tufts, L., Citizen Interviews, Organizational Feedback and Police Community Relation Decisions, Law & Society Review, November, 1971.

Cohen, B., and Chaiken, J., Police Background Characteristics and Performance, New York, Rand, 1972.

Cumming, E., et al, Policeman as Philosopher, Guide and Friend, Social Problems, Winter, 1965.

Eastman, G., Municipal Police Administration, Washington, D.C., International City Management Association, 1971.

Epstein, G., and Laymon, R., Guidelines for Police Performance Appraisal, Promotion and Placement Procedures, LEAA, 1973.

Furcon, J., et al, A Longitudinal Study of Psychological Test Predictors and Assessment of Patrolman Field Performance, June, 1971, LEAA.

Furstenberg, F., and Wellford, C., Calling the Police: The Evaluation of Police Service, Law and Society Review.

Goldstein, Administrative Problems in Controlling the Exercise of Police Discretion, JCLCPS, 1967.

Hatry, H., Measuring the Effectiveness of Non-defense Public Programs, Operations Research, Sept.-Oct., 1970.

Institutional Arrangements and Measurement of Policy Consequences in Urban Areas, Urban Affairs Quarterly, 1971.

Hoffman, R. B., Performance Measurements in Crime Control, SUNY at Buffalo, School of Management, working paper series, No. 112.

Iannone, N. F., Supervision of Police Personnel, Prentice Hall, Englewood Cliffs, N. J., 1970.

IACP, Guide to Performance Evaluation.

Landry, F. and Fari, J., Police Performance Appraisal, Dept. of Psychology, Penn. State Univ., 1973.

Lewis, J., Evaluation of Experiments in Policing, Washington, D.C., Police Foundation, 1973.

Kent, D., and Eisenberg, T., The Selection and Promotion of Police Officers: A Selected Review of Recent Literature, The Police Chief, February, 1973.

Malty, M., Evaluation of Crime Control Programs, LEAA, 1972.

Price, J., The Study of Organizational Effectiveness, The Sociological Quarterly, Winter, 1972.

Reiss, A., Police and Public, Yale University Press, New Haven, 1971.

Skolnik, J., Justice Without Trial, Wiley, New York, 1964, Chapter 8.

Webb, E., et al, Unobstructive Measures: Non-reactive Research in the Social Sciences, New York, Rand McNally, 1969.

END