

CALLING THE POLICE: THE EVALUATION OF POLICE SERVICE

FRANK F. FURSTENBERG, JR. *University of Pennsylvania*

and

CHARLES F. WELLFORD *Florida State University*

AUTHOR'S NOTE: *This research was supported by the National Institute of Law Enforcement and Criminal Justice.*

INTRODUCTION

In the rapidly expanding literature on the police, one question has been treated with increasing frequency in the past few years: How can the police be encouraged to become more responsive to the publics they serve? Many writers contend that in order for change to take place, it will be necessary to alter those conditions which affect the type of individual who becomes a policeman, either by modifying recruitment procedures, the training process, or the schedule of compensation (President's Commission, 1967a: 20). Others have questioned whether such strategies will work unless they are accompanied by basic social and organizational changes, redefining the role of the police officer and his relationship to the community (President's Commission, 1967a: 149).

This paper proposes one such technique of organizational change, the introduction of a system of citizen evaluation. In the following discussion, we will consider the impact this technique might have on police behavior and report the results of a pilot study, conducted last year, to determine whether this approach toward evaluating the quality of police service is feasible.

Purpose of Citizen Evaluation

Police have traditionally relied on crime statistics as an indication of how well they are carrying out their duties. Preventing crime and apprehending criminals are, of course, primary goals of a police force. But to assess performance by this single standard is to discount many other services which police departments provide, services which help to maintain order but contribute only indirectly to crime control. It is our contention that both the police and the public would benefit if the criteria for evaluating performance were broadened to include these other activities.

However competent the police may be, they can exert only a limited influence on the amount of crime in the community. The type and efficiency of police operations probably have far less impact on the rate of crime in a given location than do, for example, demographic factors (President's Commission, 1967b: 25-27). Similarly, whether or not the police apprehend an offender is often more closely related to the type of offense committed than to the competence of the police investigators. Focusing on crime rates to measure effectiveness — and thereby persuading the public that the police hold the key to solving the crime problem — may have the short-term advantages of obtaining larger budgets and gaining greater public support for strict enforcement. Eventually, however, public confidence in the police is likely to be eroded. People may come to hold the police partly responsible for the level of crime in the community, and the police may, in turn, feel unappreciated and resent the public's loss of faith in them.

Confidence in the police might be enhanced if the police began to measure, and thus make more visible, the other activities they perform, which are now largely hidden from public view. It has been estimated that returning stray pets to their owners, answering sick calls, mediating family quarrels, removing illegally parked cars and other mundane tasks take up as much as three-quarters of police officers' time (Bard, 1970: 1). Such tasks fall to the police because no other agency can or will perform them. And they account, to a great extent, for the positive aspects of the police image. Paradoxically, however, the police have been reluctant to evaluate what they probably do best — dealing with a wide array of these day-to-day crises.

A second consequence of including the order-maintenance tasks in the assessment of police performance might be to make the police more responsive to the communities they serve. As things now stand, when a police officer answers a service call, the gratitude received from those he helps is likely to be his only reward. His standing in the department, and ultimately his promotion, rests largely on his performance in crime-related activities — the number of arrests he makes and the crimes he solves. As a result, the patrolman soon comes to regard service calls as an inevitable but unimportant distraction from his "real work" (President's Commission, 1967a: 13). Carried to an extreme, this attitude can lead to a disregard for the average consumer of police service. We suggest that letting consumers participate in police evaluation might have a powerful impact

on how the police deliver their services, for the patrolman would become more accountable to the citizens he must serve. He is likely to take seriously even the most routine duties if persuaded that his chance for advancement depends on how well they are performed.

The question, then, is how to design this type of evaluational system. Two years ago, the Baltimore Police Department cooperated in a preliminary effort to develop a method of citizen evaluation. The procedure devised proved to be efficient, simple and relatively inexpensive. In this report, we describe the method employed and discuss some of the results of the pilot survey. The research described here represents a preliminary step toward the ultimate goal of assessing the impact of consumer evaluation on departmental organizations. Though this paper is primarily concerned with the methodology of conducting a citizen evaluation, some of our findings on the public's satisfaction with police service proved to be substantively interesting. In the conclusion, we shall return to some of the possibilities and problems of using this approach to assess police behavior.

METHODS OF STUDY

For some time prior to this study, the Inspectional Services Unit (ISU) of the Baltimore Police Department had been performing a routine check on the response to calls for service. The procedure followed by the ISU was to conduct a quarterly interview of approximately 200 citizens who had called the department for service during the preceding period. While the sample was not drawn randomly, the ISU made certain that callers from each of the nine police districts were included and that serious and non-serious offenses were represented in equal proportions. Generally, the ISU had managed to obtain completed interviews on 50 to 60 percent of the sample.

The unit had operated without a structured interview schedule; instead, an informal but fairly standardized set of questions was used. The interview focused on the citizen's satisfaction with the response to his call and on whether he thought the police were courteous and efficient in handling his complaint. Since structured schedules were not used, the results of the interviews were difficult to tabulate. However, in reports to the Commissioner, some attempt had been made to summarize the total number of complaints. A complaint, as defined by the ISU, was an expression of extreme dissatisfaction which merited further investigation. In the year preceding

the experiment, less than one percent of the respondents made such complaints.

The Baltimore Police Department's efforts to monitor its performance with a quality control check raised the possibility of standardizing the procedure for use by other departments in the country. After several months of working closely with the ISU, a method was developed that would provide a simple basis for comparison over time or among different departments. There was as little interference as possible with the procedure already in operation; it was changed only to improve its efficiency or replicability. The major task was to construct an interview schedule which included most of the features of the unstructured interview then in use. The interview was kept as brief as possible, focusing on the respondent's description and evaluation of the services rendered by the police department. While the length of the interview varied slightly according to the offense and to the interview situation, it generally could be completed in ten to fifteen minutes.

The sampling procedure followed closely the one employed by the department. A systematic sample of all calls received during a one-month time period was made. As in the police department's survey, the study sample was drawn from all nine district files, but the number taken from each was adjusted in proportion to the load of calls handled by the district office. The sample was stratified to give equal representation to blacks and whites, and to persons reporting serious and miscellaneous incidents. Finally, following the department's practice, certain cases where an interview might prove embarrassing to the caller were excluded. Thus, the study includes no homicides or sex offenses.

In order to estimate the extent of bias that might be introduced by using police interviews, half the sample were interviewed by members of the ISU and half by civilian interviewers recruited and trained by a professional research agency. In all instances, the civilian interviewer was the same race as the respondent, and it was usually possible to match the sex of the two as well. This procedure differed from the one used by the police department, as all the members of the interviewing team were white males.

It was hoped that the interviews would be conducted a short time after the complaint or request for assistance was made. However, unexpected delays occurred both with the police and with the civilian interviewers; consequently, the

field work period was longer than had been anticipated. Because of the relatively long delay (approximately two months) in completing the field work, an additional number of freshly drawn callers was assigned to the civilian interviewers in order to check the effect of delay between the time the call occurred and the interview.

THE FINDINGS

In presenting the data, it is useful to distinguish two types of results. One set of findings bears on a strictly methodological question: What difference did it make whether citizens were interviewed by police or civilians? The second set of findings bears on questions of a more substantive nature: How do the respondents evaluate police performance, and what are the most important determinants of their judgments? As it happens, citizen evaluation was related to the interviewing procedure, so at a certain point the methodological and substantive findings necessarily merge. Nevertheless, if the methodological results are examined first, it will be easier to interpret subsequent findings on evaluation of police services.

Fieldwork Results

When the plan to have civilians conduct some of the interviews was suggested to the police, they were somewhat dubious about the ability of non-police interviewers to gain entry into the respondents' homes. As it turned out, the cooperation given to the professional interviewers was excellent, and nearly as great as the police, who encountered no refusals. Of the 421 interviews attempted by civilians, only ten were not completed because the respondent refused to participate in the study. A refusal rate under three percent, especially when the sample includes a high proportion of individuals with low incomes, is highly satisfactory.

While refusals did not turn out to be an important source of non-completion, problems in locating the respondents were greater than had been anticipated. During the period between their call to the police and the attempted interview, 16 percent of the respondents assigned to the civilians had moved or could not be located at the address recorded on the call sheet. An additional 12 percent were not at home on any of the three occasions at which the interviewer visited the house. Undoubtedly, had more than three attempts been made, a number of these respondents could have been interviewed successfully; it is likely, though, that some of those never contacted were no

longer (or had never been) living at the address indicated on the report form.

The police encountered as much difficulty as the civilian interviewers in locating respondents. In fact, their overall completion rate was appreciably lower than the civilian rate (56 percent as compared to 69 percent), probably in part because time did not permit the police to make as many call-backs. To determine if additional visits would allow the police to complete more interviews they were asked to make three attempts on a designated sub-sample of the respondents assigned to them. While the overall completion rate for this "special" sample was slightly higher (59 percent), it still did not reach the level achieved by the civilian interviews (Table 1).

One reason for the higher completion rate of the civilian sample is that it included respondents who were interviewed a week or two (instead of a month or more) after they had called the police. Quite clearly, the number of completed interviews could be increased significantly when the time between the call and the interview is reduced (Table 1). However, even when the number of attempted interviews and the timing of the interview are discounted as influential factors, a disparity remains between the two samples.

When the respondent's race is introduced as a factor one reason for the residual difference emerges: the police were less successful in interviewing black callers. There are two possible explanations. First, perhaps policemen encounter more difficulty interviewing black citizens. Alternatively, the problem may have arisen because the police were all white. In any case, the police appeared to recognize the difficulty and preferred to interview white callers. On the occasion when researchers went with them while they conducted the inter-

TABLE 1: PERCENT OF INTERVIEWS COMPLETED, BY INTERVIEW TYPE AND RACE*

	Civilian			Police		
	Total	Regular	Recent	Total	Regular	Special
White	72 (213)	68 (146)	79 (67)	60 (196)	59 (158)	66 (38)
Black	66 (208)	65 (134)	68 (74)	53 (198)	51 (136)	55 (62)
TOTAL	69	67	73	56	55	59
N=	(421)	(280)	(141)	(398)	(294)	(100)

* Cells of the table indicate completion rate and the number of attempted interviews.

views, the policemen seemed less comfortable in the black areas of the city, possibly because they perceived that they were less welcome there. Even for the special sample, where the police were asked to make several call-backs, they had greater difficulty completing interviews with blacks. Whether black police interviewers would have had as much success as black civilians is a matter for further investigation.

Satisfaction with Police Service

The interview contained several measures of the respondent's satisfaction with the service provided by the police. While the level of satisfaction varied slightly with the measure used, the same finding emerged: the majority of the respondents were pleased with the service received. Most respondents ranked the police high on courtesy, understanding and capability, and concern about their problems. When these four items were combined into an index of performance, three-fourths of the respondents gave the police the highest possible rating. Most respondents (86 percent) thought that the police had done everything they could to handle their complaints, and nearly everyone said he would call the police again if a similar problem arose. Finally, asked to report overall satisfaction with the way the police had performed their duties, 75 percent felt very satisfied, and only 12 percent indicated a low level of satisfaction. Though comparative data are unavailable, it hardly seems likely that other municipal agencies would fare much better if similar evaluations were made of their services.

Racial Difference in Satisfaction

The figures presented above are based on the total sample, comprised of respondents interviewed by both police and civilians. When the sample is subdivided, it appears that citizens interviewed by the police tend to be much less critical of the service they received than the civilian-interviewed respondents. This finding applies especially to the blacks in the sample assigned to the police.¹ Blacks interviewed by civilians were much less likely to report themselves very satisfied on the summary measure of satisfaction than were blacks interviewed by police (55 percent as compared to 76 percent, Table 2), while among whites, there was only a 10 percent difference.² Thus, when police conduct the interviews, it appears that the overall degree of satisfaction expressed, especially by blacks, is probably somewhat inflated. Again, the data at hand do not reveal which specific factors — having interviewers who

were white, police, or white policemen — affected the responses of callers contacted by the police.

TABLE 2: PERCENT OF RESPONDENTS SATISFIED, BY INTERVIEW TYPE AND RACE

Satisfaction with Service	Interview Type			
	Civilian		Police	
	White	Black	White	Black
Very	74	55	84	76
Fairly	17	33	11	21
Dissatisfied	9	12	5	3
N=	(138)	(112)	(113)	(100)
Performance Index*				
High (12)	75	54	87	83
Medium (10-11)	16	20	8	7
Low (4-9)	9	26	5	10
N=	(139)	(113)	(114)	(100)

* Weighting of responses: Good = 3, Fair = 2, Poor = 1, and don't know = 2.

The lower level of satisfaction among black respondents cannot be interpreted merely as methodological artifact, resulting from the interview situation. Regardless of who conducted the interview, blacks are more critical of the service they received. For example, 84 percent of the police-interviewed whites and 74 percent of the civilian-interviewed whites were very satisfied with the service rendered by the police. Corresponding figures for blacks are 76 percent and 55 percent. Similarly, 26 percent of the civilian-interviewed blacks rated the police low on the performance index as compared to 9 percent of the whites. Police-interviewed blacks, while much less severe in their judgments of performance, were still more critical than police-interviewed whites.

Sources of Racial Differences

Do blacks actually receive inferior service, or can their more unfavorable evaluation be interpreted as a reflection of general disenchantment with the police? Blacks may have more negative feelings toward police because they have experienced racial discrimination in the past, suffered higher rates of crime in their neighborhoods, or witnessed abuses of police authority. These factors may predispose them to be more critical of the service they receive even though it is similar to that provided to white callers. On the other hand, it is equally plausible that blacks do not receive the same treatment as whites when they request police assistance. Al-

though these alternative suppositions cannot be tested conclusively in this study, it is possible to provide a tentative answer from the data at hand.

In addition to the various measures of satisfaction, the interview contained a number of items which asked respondents to describe, rather than evaluate, police response to their calls. Because they call for no explicit evaluation, these measures might be considered as more "objective" indications of police service — although it is obvious that they do not completely eliminate the influence of the respondents' general feelings toward the police. Nevertheless, comparing the reports of black and white callers on these items provides an understanding of the quality of police service that is somewhat more detached from respondents' general sentiments toward the police.

One such measure was based on the time it takes a police car to respond to a call for assistance, an aspect of their performance which the police monitor closely. The respondents in the survey were asked to estimate how long it took the police to arrive after they were contacted. The respondents' self-reports indicate that blacks experienced a longer response time. However, this disparity appears only among the respondents who were interviewed by civilians. Within this sample, twice as many blacks as whites had to wait at least fifteen minutes for the police to arrive (23 percent compared to 12 percent). There was a noticeable drop in overall satisfaction as reported response time increased. Respondents, both black and white, who were kept waiting were generally less pleased with overall performance of the police. While the differential in response time accounts partially for the difference in satisfaction between blacks and whites, it is not a full explanation; regardless of response time, blacks reported being less satisfied.

Other than the difference in response time discussed above, no consistent or sizable differences in assessment of the quality of police service were evident for the racial groups. A slightly higher proportion of black respondents reported that the police did not make a home visit, but blacks also said they spent more time with the police when such visits did occur. Blacks were no less likely than whites to indicate that the police had followed up their complaints after the initial contact was made.

While several of these factors proved to be related to the

overall satisfaction of the respondent, none could completely account for the racial difference reported earlier. Specifically, it was found that when the police took the time to explain what they were doing or what they would do to handle the complaint, respondents were generally more satisfied with the job the police did. Similarly, when the police followed up the complaint by some further action (either by a second call or an investigation), respondents were more pleased with the service they received. But, regardless of the action taken by the police, black respondents remained, in each instance, more dissatisfied with police performance than whites.

This persistent racial difference occurs primarily among those respondents who found fault with the service they received. Specifically, blacks who report poor service (a delay in response time or no follow-up) are much more likely to react negatively than whites who had similar complaints. Conversely, whites are somewhat more inclined than blacks to discount these instances of poor service in their overall estimate of satisfaction with police performance (Table 3). Thus, it is not the quality of service as such that accounts for the racial difference — blacks and whites generally received the same treatment — but the way that service is defined by the racial groups.

TABLE 3: PERCENT VERY SATISFIED, BY PERFORMANCE MEASURES, RACE, AND INTERVIEW TYPE (N)

Performance Measure	Interview Type			
	Civilian		Police	
	Black	White	Black	White
A. Response Time				
5 min.	77 (30)	83 (54)	95 (37)	93 (40)
6-10	54 (24)	81 (36)	73 (33)	88 (33)
11-15	48 (27)	62 (26)	57 (19)	84 (19)
16+	33 (24)	38 (13)	*	55 (11)
B. Police Explain Actions				
No	36 (44)	63 (46)	65 (26)	77 (30)
Yes	54 (35)	80 (66)	75 (28)	80 (41)

* Denotes number of cases is less than 10.

This finding indicates the possibility, raised earlier, that blacks apply stricter standards than whites when asked to judge police performance. That blacks generally do, in fact, have a lower opinion of the police was convincingly established in 1965 by the President's Commission on Law Enforcement and the Administration of Justice (1967). On the basis of evi-

dence collected in a national survey, the Crime Commission reported "striking" racial differences in attitudes toward the police. Blacks were much more likely to doubt the honesty of police than were whites and thought more often than whites that the police did a poor job of enforcing the law. In explaining these differences, the Commission observed that "too many policemen do misunderstand and are indifferent to minority-group aspirations, attitudes, and customs" and that "incidents involving physical or verbal mistreatment of minority-group citizens do occur and do contribute to the resentment that some minority-group members feel" (p. 257).

More recently, a Harris Poll conducted in Baltimore confirmed this racial pattern of antipathy toward the police. A much greater proportion of blacks (46 percent compared to 27 percent of the whites) gave the police a negative rating (*i.e.*, rated overall performance as either only fair or poor). The same question, used by Harris to measure citizens' overall rating of the police, was repeated in this study. Although the findings indicate that this sample of citizen callers has a more favorable view of the police, the racial differential again emerges. Indeed, among the civilian-interviewed respondents, the difference is even greater (30 percent) than it was in the Harris survey.³

Assuming that this general rating represents an underlying attitude toward the police, it is possible to determine whether these sentiments predisposed a citizen to be more or less critical of the service rendered by police officers who responded to his call. This apparently is the case. Whether black or white, police- or civilian-interviewed, a respondent's specific rating of the quality of police service is strongly related to his overall opinions of the Baltimore police. Although less than 20 percent of the entire sample had negative general opinions of the police, 60 percent of the callers who were most dissatisfied with the service they had received held generally negative attitudes. Since blacks are much more likely to hold negative opinions of the police, it follows that they would be more critical of the service they received. Indeed, the differences in satisfaction between blacks and whites disappears if we control for the respondents' general attitudes toward the police. Undoubtedly, too, the specific experience probably caused some respondents to modify their overall opinions of the police. However, judging from the reports, such instances were relatively infrequent. Only 10 percent said that their experi-

ences with the police caused them to change their general opinions.

TABLE 4: PERCENT OF RESPONDENTS VERY SATISFIED, BY GENERAL EVALUATION, RACE, AND INTERVIEW TYPE (N)

Evaluation of Police	Interview Type			
	Civilian		Police	
	Black	White	Black	White
Excellent	92 (12)	92 (60)	94 (34)	94 (50)
Good	66 (47)	69 (54)	73 (52)	82 (54)
Fair/Poor	33 (42)	33 (18)	36 (11)	*

* Denotes number of cases is less than 10.

This pattern of attitudinal consistency leads to an understanding of why black respondents may have been more critical in their specific ratings of the police response to their complaints. But it probably would be a mistake to dismiss their more negative evaluations entirely. Even when comparable service is provided by the police, blacks may perceive differences in the treatment they receive. For example, the use of a first name, a not uncommon practice among police, may be interpreted as a gesture of friendship by a white caller and a sign of condescension by a black, especially if the police officer is white. Another difference which may lead blacks to be more critical is the under-representation of blacks on the police force. The fact that when police assistance is required, blacks will usually be visited by a white officer may be an important source of resentment.⁴ Thus, there may be differences unrelated to the performance of the police that convince blacks that they are being treated unfairly.

SUMMARY AND CONCLUSIONS

In the beginning of this paper, we suggested that altering the evaluation system of police to allow for consumer feedback might have a powerful impact on how police deliver their service and how the public perceives the police. The possibility of police departments using consumer surveys to assess their performance promises several benefits: *First*, by conducting periodic surveys, the police would be able to detect trends and shifts in citizen reaction to the police. It should be possible also to compare these trends to other localities where similar evaluative measures were employed. Thus, the police would have a continuous baseline for gauging the success of new policies and programs. *Secondly*, consumer surveys would reveal variations within communities, enabling the police to concentrate their efforts on areas of the city where public

satisfaction is low. In this way, police might be encouraged to become more attentive to communities which have been neglected in the past or are presently not well served. *Finally*, if consumer satisfaction were to become a criterion for evaluating the performance of departments, and individuals within the departments, the general effect should be to increase police responsiveness, particularly in their day-to-day contacts with the public.

The study described in this paper represents a first step in developing a procedure of citizen evaluation. We conducted a survey of citizens who had requested service from the police to determine the feasibility of this technique and to obtain some idea of what the police might expect if they were to employ such a method. Our findings were:

1. Evaluative research on the quality of police service is highly feasible. Cooperation from respondents in this study was generally excellent, and respondents seemed more than willing to volunteer their impressions.
2. While the effect of using police as interviewers is not great if rigorous sampling conditions are observed, it would seem preferable to use trained civilian interviewers. Black respondents in particular are unwilling to criticize police performance unless interviewed by civilians. Whites are less hesitant in this regard, perhaps only because they are generally less critical of the police.
3. The majority of the citizens who call police for service are satisfied with the quality of service provided by the police. Most respondents could find little fault with the performance of the police on any of the measures included in the questionnaire.
4. Blacks were consistently more critical than whites of the quality of police service. While there is some evidence that blacks experienced a slightly longer delay in waiting for the police to arrive, in most other respects the service they received was comparable to that provided to whites.
5. The dissatisfaction of blacks seems to derive, at least in part, from their generally lower opinion of the police. However, the observation that general satisfaction can only be minimally influenced by satisfaction with specific acts of service offers little direction for change.⁵ We suspect that a more rigorous and responsive attempt

at evaluation of police service might produce as positive an effect on community perception of police as any changes that are likely to occur in the near future. It is important to note though (as we alluded to earlier) that this technique does not really tap the sources of dissatisfaction, particularly among the most dissatisfied — those who do not call the police. We suggest that evaluation by this group must also be included, most likely through community-wide service evaluation.

FOOTNOTES

- ¹ This difference among the whites virtually disappears if the police sample is adjusted to include only those callers in the Special Police Sample (i.e., where the police made at least three attempts to interview the respondent). This corrective procedure also reduces the difference in satisfaction among blacks, but a sizable disparity (16 percent) remains.
- ² We found no significant zero-order relationships between other independent variables (type of report, age, and sex) and level of satisfaction, so they were excluded from the subsequent analysis.
- ³ There are two important differences between our survey and the one conducted by Harris which may help to explain why our results are more favorable toward the police. First, our study pertains only to citizen initiated contacts with the police and excludes any references to situations that may provoke the greatest dissatisfaction (e.g., stopping for interrogation, corruption, etc.). Furthermore, our population consists only of those who called for service; the Crime Commission noted that the highest levels of dissatisfaction occurred among those who did not contact the police when victimized. Thus, our sample probably overrepresents citizens who have a positive opinion of the police.
- ⁴ There are no available figures on the proportion of black policemen in Baltimore, but an impression from several weeks of observation is that the percentage is a good deal lower than the percentage of blacks living in Baltimore City. Future evaluational studies no doubt will be able to test the effect of racial similarities and differences on citizen satisfaction with police service.
- ⁵ Block (1969) reports similar findings on the relationship between quality of service and the general evaluation of police.

REFERENCES

- BARD, Morton (1970) *Training Police as Specialists in Family Crisis Intervention*. Washington, D.C.: U.S. Government Printing Office.
- BLOCK, H. Richard (1969) "Public Attitudes Toward the Police," Paper presented at the American Sociological Association meetings, August, 1969.
- PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF JUSTICE (1967a) *Task Force Report on the Police*. Washington, D.C.: U.S. Government Printing Office.
- PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF JUSTICE (1967b) *Task Force Report on the Assessment of Crime*. Washington, D.C.: U.S. Government Printing Office.