

MULTI-MODES MEASURES: FROM POTHOLES TO POLICE

by

Elinor Ostrom

Workshop in Political Theory and Policy Analysis
Indiana University

Paper presented at a conference on "Productivity and Program Evaluation: Challenges for the Public Service" organized by the Midwest Intergovernmental Training Committee on June 6, 1975.

Research on multi-mode approaches to the measurement of the output of local public agencies has been supported by the Research Applied to National Needs Division of the National Science Foundation in the form of Grant Number GI 38535.

MULTI-MODE MEASURES: FROM POTHOLES TO POLICE

by

Elinor Ostrom

On Palm Sunday one year ago, I found myself walking down an Indianapolis street, carrying a yardstick and dashing out between passing cars to measure the potholes in the street. Why would any sane person dash out onto a busy street to risk their life to measure some holes in the ground? I must confess that I asked myself that question several times that day and other days while I helped develop our "unobtrusive" measures of road conditions. The answer to that question takes a somewhat long route but will be the focus of this presentation.

Measuring the Output of Police

The story begins in 1970 when we first started a series of studies comparing the levels of service output for small and large police departments serving relatively similar neighborhoods within metropolitan areas.¹ Our substantive interest in police resulted from the repeated assertions that small scale police agencies were less effective than large-scale agencies.²

In our attempt to examine these assertions empirically, we first used a series of measures derived from citizen surveys. The first type of measure was that of citizen experiences with crime and with police. These included:

- Whether anyone in the household had been the victim of criminal activity.
- If a victimization incident had occurred, whether the police were notified.

- If notified, how fast the police had responded.
- If the police responded, what level of activity was undertaken.
- Whether anyone in the household had called upon the police for assistance to deal with a noncriminal, but emergency problem.
- If assistance was requested, how fast did the police respond.
- If assistance was provided, what level of activity was undertaken.
- Whether respondents had been stopped by police serving their neighborhood.
- Whether respondents knew anyone who had been mistreated by police serving their neighborhood.

Secondly, we utilized a series of indicators of citizen evaluations of service levels. These included:

- How fast respondents thought police serving their neighborhood responded to calls for service.
- Whether respondents thought crime in the neighborhood was rising, about the same or decreasing.
- Whether respondents thought that police-community relations were outstanding, good, adequate, inadequate or very poor.
- Whether respondents thought that the job performed by police serving their neighborhood was outstanding, good, adequate, inadequate, or very poor.
- Whether respondents thought that police services were equally available to all in their community.

In later studies we added to this list of measures (derived from random samples of citizens living in matched neighborhoods being served by varying sized departments) by including measures derived from interviews with police and from other agency records.³ Some of these were:

- The evaluation given by officers to the performance of their own department.
- The ratings given by a pool of police officers serving a metropolitan area to the departments in the area.

- o The proportion of reported crimes cleared by arrest.
- o The proportion of warrants issued by the prosecutor to warrants applied for.

Findings

Using such multiple measures of police output, in a series of studies conducted in Indianapolis, Chicago, St. Louis, and then replicated in Grand Rapids and Nashville-Davidson County, we found a consistent pattern across indicators of output.⁴ Larger police agencies serving carefully matched neighborhoods were not more effective. Small departments performed either at equal or higher levels across the range of all indicators used. Having used a series of different indicators, we can make much more confident statements about the relative effectiveness of large and small police departments than if we had used a single measure of output.

Primary Reliance on One Mode of Data Collection

However, many of our individual measures of output were derived from one mode of data collection -- that of individual interviews with respondents selected by a random process. Many scholars and public officials are uneasy about any reliance upon data collected from a survey of citizens about public agency performance. Citizens are thought by some to be uninformed and unable to give reliable perceptions and/or evaluations of service levels. Whether one agrees with this view or not (we obviously do not), reliance upon any single mode of measurement can lead to errors or biases in measuring performance or productivity.

This challenge to a particular mode of data collection about the output of public agencies led us to design a study in which measures of output were obtained from multiple modes of data collection.⁵ We wanted

to compare the similarity of data collected from three modes: citizen interview, physical measurement devices and agency records. The latter two modes of data collection are considered by many as "objective" data while interview data is considered "subjective." With such comparisons we could ask, for example, whether the pattern of service delivery as perceived by citizens was similar to that measured by some form of physical device or by agency records. We sought to determine whether "subjective" measures are related to "objective" measures.

Measuring Road Condition and Street Lighting

While one can talk about "physical" measures of output for local public services, few services exist for which one can "dream up" a way of physically measuring it let alone having a measurement device already available. We selected two service areas -- road condition and street lighting -- where we thought we could develop physical measures, gain information from a random sample of citizens and obtain information from agency records concerning levels of output. The Urban Institute had already undertaken some work in the area of measuring road roughness using a "roughometer".⁶ The idea of measuring street lighting by a light-meter also seemed feasible.

Types of Measures Used

For social scientists, this project took us far astray into the realm of physics, optics and mechanics. We learned that developing valid and reliable physical measures of output - even for such relatively simple services as road condition and street lighting - is not easy.

For **street** lighting, **we** developed **a** method of using a precision light-**meter** to record the level of night lighting on sidewalks and streets facing **a** particular blockface. For street condition, **we** developed **a** mechanical device, called the Residential Street Roughness Indicator, to measure the roughness of residential streets.⁷ We also developed an observation form and procedure which can be used by trained observers to record specific data about various aspects of street condition. The procedure includes, among other items, measuring all potholes on a blockface with a "yardstick pothole measurer." It was that observation procedure that I was testing on Palm Sunday one year ago.

A citizen survey represented our second mode of data collection. This survey was pretested by interviewing 326 respondents living in seven neighborhoods in Indianapolis during the spring of 1974. Citizens were asked **a** series of questions to elicit their perceptions of street lighting and road condition, their evaluations of these services and their preferences for different levels of these services.

Agency records were to be our third mode of data collection. Unfortunately, we found this mode of data collection to be the most difficult of all. Agency records were so fragmentary that few consistent indicators could be developed. For road conditions we were only able to code the frequency of complaints directed to the Indianapolis Department of Transportation concerning the roads facing the respondents in our survey. For street lighting, we were not able to code much more than the frequency and pattern of street lights shown on power company maps.⁸

Preliminary Findings

Preliminary data analysis has been initiated. One of the first

questions we have addressed is "What is the relationship between citizen perceptions of service levels and our physical measures of service levels?" This is an important question since many policy analysts are hesitant to rely upon citizen reported evaluation of output due to the assumed inaccuracy of citizen perceptions of service levels. Early analysis does not provide a uniform picture of accurate perception across all indicators. However, the more specific and concrete the referent to which our questions were addressed, the more likely a high level of association exists between physical measures and citizen perceptions of service levels.

For example, citizens were quite accurate in their perceptions of specific aspects of road condition on their block. Citizens accurately reported the type of street surface, the presence or absence of curbs, the condition of their curbs, the presence of surface disintegration, and the presence of potholes.⁹

Sue Carroll of the Workshop staff developed a roughness scale composed of individual items derived from our observation procedures.¹⁰ The scale was developed for each quadrant of a blockface and for an entire blockface. The scale included the observer's coding of the amount of surface disintegration, the number and size of potholes, the presence or absence of cracks, the presence or absence of bumps and the presence or absence of utility cuts. Each observer was also asked to rate each blockface as being "very rough," "fairly rough," "fairly smooth," and "very smooth." These observer ratings were strongly related to the "roughness scale" for both a quadrant ($\gamma = .94$) and for the blockface as a whole ($\gamma = .97$).

When respondents' perceptions of the roughness of the street on their block were then associated with the quadrant and blockface roughness scale, the measure of association between them is fairly strong ($\gamma = .76$ for both scales). Some variation occurred across various control variables. Those persons with more than a high school diploma, those over 45, those who have lived on a block more than five years, and those living on medium to short blocks tended to be more "accurate" in their perceptions of road roughness. Initial data analysis with scores produced by the Residential Street Roughness Indicator device are consistent with these findings. A high association exists among all these individual modes of data collection concerning the level of road roughness. The perceptions of citizens, the output from a mechanical device for measuring road roughness and the coded observations of trained field-workers are strongly associated.

Although the levels of association are not, in general, as high as in the case of street condition, statistically significant correlations between citizen perceived streetlight brightness levels on their blockface and data from a precision photoelectric meter were found.¹¹ Further, a distinct pattern emerged between the strength of association and the distance on either side of a respondent's house over which light-meter readings were averaged. Correlations reach a maximum when meter readings are averaged over intervals relatively proximate to a respondent's home and decline as the meter readings are averaged over widening intervals.

The lowest correlation is between citizen perceptions of streetlight brightness and light-meter readings averaged for an entire blockface. For some subsets of the sample, this correlation was not statistically significant. Citizens, thus, appear to show a pronounced tendency

to perceive blockface streetlight brightness conditions in terms of the brightness levels relatively proximate to their own homes. When asked specifically about conditions near to their homes, citizens are more accurate. We found that citizens who had lived on their block for more than 10 years, who had a high school or better education, or who lived on relatively short blocks showed a higher than average degree of accuracy.¹²

Significance of Findings

Ascertaining the Least Expensive, Valid Measure of Output

The finding that measures derived from interviews with citizens and measures derived from physical devices or field observation forms are positively related for two service areas has considerable importance for those interested in measuring the productivity of local public service agencies. In the first place, given the close association between the diverse modes of data collection regarding road conditions, public officials or public interest groups concerned in ascertaining the relative productivity of agencies engaged in road repair activities can select the mode of data collection which is least expensive. The results obtained should correlate highly with results obtained from more expensive modes of data collection. Thus, we can increase the productivity of measuring productivity! Of the three modes of data collection compared -- the Residential Street Roughness Indicator, the field observation form and the citizen survey -- the field observation form is the least expensive mode of data collection.

The Workshop in Political Theory and Policy Analysis is currently preparing a "Portfolio of Professional Papers" on "Multiple Measures of

Municipal Output." The Portfolio will contain a copy of the Road Condition Observation Form and an instruction booklet on the procedures for using this form. We will include in this Portfolio copies of a citizen survey related to road condition and streetlighting and drawings of the Residential Street Roughness Indicator for those who wish to utilize the other modes of data collection or to further experiment with simultaneous collection of output measures with more than a single mode of collection. Field procedures for utilizing a precision light-meter, descriptions of the two instruments we utilized in our citizen survey - the Street Lighting Simulator and the Street Lighting Photograph Display - will also be included in this portfolio.¹³

Citizen Perceptions of Specific Service Attributes Relatively Accurate

The findings from these studies are important also for those areas where no physical measures of output are possible. Given that citizens seem to be fairly accurate in their perception of specific and clearcut attributes of road condition and street lighting, one can have somewhat more confidence that citizens will be fairly accurate in their perception of specific attributes of other services. To the extent that questions can be phrased about specific aspects of services for which there are no physical measures of output, it would appear that one can obtain fairly accurate perceptions from citizens.

This means that considerable attention must be paid to the construction of questions on citizen surveys. We found, for example, in our early pretesting of our police-services instrument that asking the question: How fast did the police arrive? (to a person who had been a victim of crime and had called upon the police) was not specific enough.

Many respondents indicated only "very fast" or "slow." We then found that some people thought "slow" was 10 minutes while others thought "slow" was 30 minutes. Given this information from our pretests, we obtained data about response time in terms of minutes elapsed between call and response.

Citizen Preferences May Differ from Those of Officials

In addition to asking citizens about their perceptions of service conditions, we also asked about their preferences concerning the purposes of street lighting and road repair. For example, we asked respondents:

Which of the following purposes do you personally consider to be the most important purpose street lighting should serve on your block: To discourage crime? To assist pedestrians or to help prevent traffic accidents?

Sixty percent of our respondents indicated that discouraging crime was the most important purpose, 23 percent indicated that assisting pedestrians was most important while only 14 percent indicated that preventing traffic accidents was most important (three percent responded that they did not know).

In talking with public officials in the City of Indianapolis, we were told that the purpose of street lighting in the City was to prevent traffic accidents. Most of Indianapolis has light fixtures at only intersections with the direction of the lights pointed toward the street rather than the surrounding sidewalks and front yards.

Without obtaining systematic citizen feedback, public officials have little opportunity to learn that citizen preferences may vary significantly from the purposes for which public officials are working.

Testing Ways to Improve Productivity

Many assertions are made about ways of improving the productivity of local public agencies. However, most of the presumptions about the factors likely to increase productivity of public agencies have not been subjected to empirical test. They are accepted as part of the conventional wisdom. Many of the findings in our current research about comparative levels of output, however, run contrary to conventional wisdom.¹⁴ A surprising number of studies indicate either no significant economies or substantial diseconomies to be associated with larger size of public service agencies.¹⁵ Yet these findings are often ignored; as though we should not be confused by the facts. When presumptions which are assumed to be true are not supported by empirical evidence, we are confronted with the serious task of rethinking our presumptions and of developing alternative explanations for why we get the results we do.¹⁶ The crisis of confidence that pervades much of American public life calls for more than measures of output. But output measures -- derived from multiple modes of data collection -- are key elements in efforts to explain why we often get such poor results with such lavish efforts. A Palm Sunday measuring potholes may help in the process of improving the delivery of local services in urban areas --at least I hope so.

FOOTNOTES

See Elinor Ostrom, William H. Baugh, Richard Guarasci, Roger B. Parks, and Gordon P. Whitaker, Community Organization and the Provision of Police Services, Sage Professional Papers in Administrative and Policy Studies 03-001 (Beverly Hills, California: Sage Publications, 1973); Elinor Ostrom, Roger B. Parks and Gordon P. Whitaker, "Do We Really Want to Consolidate Urban Police Forces? A Reappraisal of Some Old Assertions," Public Administration Review 33 (September/October 1973), 423-433; Elinor Ostrom and Gordon P. Whitaker, "Community Control and Governmental Responsiveness: The Case of Police in Black Communities," in David Rogers and Willis Hawley, eds., Improving the Quality of Urban Management, Vol. 8, Urban Affairs Annual Reviews (Beverly Hills, California: Sage Publications, 1974), pp. 303-334; Bruce D. Rogers and C. McCurdy Lipsey, "Metropolitan Reform: Citizen Evaluations of Performances in Nashville-Davidson County, Tennessee," Publius 4 (Fall 1974), 19-34; and S.T. Ishak, "Consumers' Perception of Police Performance: Consolidation vs. Deconcentration; The Case of Grand Rapids, Michigan Metropolitan Area" (Ph.D. Dissertation, Indiana University, 1972); Elinor Ostrom, "Size and Performance in a Federal System" Publius 5 (forthcoming); Roger B. Parks, "Complementary Measures of Police Performance," in Kenneth M. Dolbeare, ed., Evaluative Research (Beverly Hills: Sage Publications, 1976).

2

National Advisory Commission on Criminal Justice Standards and Goals, Report on Police (Washington, D.C.: Government Printing Office, 1973).

* Scholars at the Urban Institute have also been concerned with the problems of developing measures of police productivity. See Harry P. Hatry, "Wrestling with Police Crime Control Productivity Measurement," in Joan L. Wifley and John F. Heaphy (eds) Readings on Productivity in Policing (Washington, D.C.: The Police Foundation, 1975), pp. 86-128; Peter B. Bloch, Equality of Distribution of Police Services - A Case Study of Washington, D.C. (Washington, D.C.: The Urban Institute, 1974).

See citations in footnote 1.

⁵For a more detailed description of this project see Vernon Greene, Elinor Ostrom, Roger Parks and Richard Rich, "The Measures Project - A Theoretical and Methodological Overview," (Bloomington, Indiana: Indiana University, Department of Political Science, Workshop in Political Theory and Policy Analysis, Research Report Number 1) and Elinor Ostrom, "The Need for Multiple Indicators in Measuring the Output of Public Agencies," Policy Studies Journal 2 (Winter 1973), pp. 87-91.

6

See Andrew J. Boots, Grace Dawson, William Silverman, Harry P. Hatry, Inequality in Local Government Services: A Case Study of Neighborhood Roads (Washington, D.C.: The Urban Institute, no date).

⁷See Richard Rich, "The Development of a Technique for the Physical Measurement of Residential Street Lighting," (Bloomington, Indiana: Indiana University, Department of Political Science, Workshop in Political Theory and Policy Analysis, Research Report Number 5).

⁸The power company has available to it, computer programs which can calculate expected light intensity on a street given light type, spacing, etc. These programs are not used, however, in any attempt to evaluate output.

⁹See Sue Carroll, "An Analysis of the Relationship Between Citizen Perceptions and Unobtrusive Measures of Street Conditions," (Bloomington, Indiana: Indiana University, Department of Political Science, Workshop in Political Theory and Policy Analysis, Research Report Number 10).

¹⁰Ibid.

¹¹See Vernon Greene, "An Analysis of the Relationship Between Citizen Perceptions and Physical Measures of Street Lighting," (Bloomington, Indiana: Indiana University, Department of Political Science, Workshop in Political Theory and Policy Analysis, Research Report Number 7).

¹²Ibid.

¹³These portfolios will be available by writing to: Publication's Secretary, Workshop in Political Theory and Policy Analysis, Indiana University, Morgan Hall 121, Bloomington, Indiana 47401.

¹⁴See citations in footnote 5 and Elinor Ostrom and Roger B. Parks "Suburban Police Departments: Too Many and Too Small?" in Louis H. Masotti and Jeffrey K. Hadden, eds., The Urbanization of the Suburbs (Beverly Hills, California: Sage Publications, 1973), pp. 367-402; Dennis Smith and Elinor Ostrom, "The Effects of Training and Education on Police Attitudes and Performance," in Herbert Jacob, ed. The Potential for Reform of Criminal Justice (Beverly Hills: Sage Publications, 1974), pp. 45-82; James C. McDavid, "Interjurisdictional Cooperation Among Police Departments in the St. Louis Metropolitan Area," Publius 4 (Fall 1974), pp. 35-58; Elinor Ostrom and Dennis Smith, "Are the Lilliputs in Metropolitan Policing Failures?", paper presented at the American Society for Public Administration meetings in Chicago, April 2, 1975.

¹⁵For a review see Robert L. Bish and Vincent Ostrom, Understanding Urban Government: Metropolitan Reform Reconsidered. (Washington, D.C.: American Enterprise Institute, 1973).

¹⁶See Vincent Ostrom, The Intellectual Crisis in American Public Administration (University, Alabama: University of Alabama Press, 1973).