PERFORMANCE MEASUREMENT AND THE CRIMINAL JUSTICE SYSTEM:

FOUR CONCEPTUAL APPROACHES

NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
U.S. DEPARTMENT OF JUSTICE

PERFORMANCE MEASUREMENT AND THE CRIMINAL JUSTICE SYSTEM:

FOUR CONCEPTUAL APPROACHES

The four concept papers presented herein were supported by Contract Numbers: 5-1007-J-LEAA, 5-0973-J-LEAA, 5-1008-J-LEAA, and 5-0968-J-LEAA, awarded by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

OCTOBER 1976

NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
U.S. DEPARTMENT OF JUSTICE

FOREWORD

As part of its on-going work in evaluation research, the National Institute's Office of Evaluation in the spring of 1975 commissioned four concept papers dealing generally with the issue of data analysis in evaluating criminal justice programs. Specifically, the papers address the problems of developing system-wide productivity measurements for evaluation.

Four working teams -- headed by Professors Alfred Blumstein at Carnegie-Mellon University, Stuart Deutsch at Georgia Tech, Richard Larson at MIT, and Robert Lind at Cornell -- were asked to provide a research design focusing on productivity evaluation and the attendant problems of data analysis. Each team prepared a draft working paper, which was first reviewed by the Office of Evaluation staff and then presented at a meeting of researchers at the Department of Justice in the fall of 1975. Based on the comments and reviews resulting from that meeting, each team's principal author revised his paper for final submission to the National Institute.

The papers assembled in this volume are working papers designed primarily for the use of the Office of Evaluation staff. However, the National Institute believes the papers are of sufficient interest to the research community to warrant publication on a limited edition basis.

Gerald M. Caplan
Director
National Institute of Law Enforcement
and Criminal Justice

CONTENTS

PERFORMANCE MEASUREMENT AND THE CRIMINAL JUSTICE SYSTEM

by

Jacob Belkin
Alfred Blumstein
Gordon Cassidy*
Jacqueline Cohen

Urban Systems Institute
Carnegie-Mellon University

*Director of Evaluation
Ministry of the Solicitor General, Ottawa, Ontario, Canada

TABLE OF CONTENTS

PERFORMANCE MEASUREMENT AND THE CRIMINAL JUSTICE SYSTEM

by

Jacob Belkin
Alfred Blumstein
Gordon Cassidy*
Jacqueline Cohen

Urban Systems Institute
Carnegie-Mellon University

SUMMARY

In this paper, the measurement of the criminal justice system performance is derived from an initial conceptualization of the "total social cost associated with crime and crime control, and the net costs associated with the service provided by the criminal justice system. The costs of crime are related to victimization risk from one perspective, and "criminality" or crime-committing propensity from another, and are desired to be minimized. The social costs of crime control include conventional resource expenditures by the agencies of the criminal justice system, but also must account for the degree of intervention in the lives of arrestees, witnesses, jurors, and others who encounter the criminal justice system. The considerations here include those of minimizing disruption and assuring equity in the disruption and treatment.

In addition, the criminal justice system provides a variety of public services, including emergency services by the police, conflict resolution by the police and the courts, and a variety of remedial and social services by all parts of the criminal justice system. These represent benefits, not only for themselves, but also for their effect in enhancing people's perceptions of the criminal

*Director of Evaluation, Ministry of the Solicitor General, Ottawa, Ontario, Canada

justice system and their identification with it.

The structuring of the issues in this context leads to the identification of six major projects:

1. Assessment of Trends in Criminality

The basic problem of "criminality" relates to the distinction between the crime-committing propensity of a given demographic group and changes in the demographic mix in the population. The fundamental problem is one of estimating the rate at which members of a particular demographic group commit crimes of a particular type, and examining the trend in that propensity over time. The basic data for such an estimation derives from police arrest records, and these must be calibrated for intensity of police activity.

Some initial analyses of this issue have been revealing, and the issue should be pursued using more extensive data for longer time periods, in different areas, and for different groups.

Because the predominant data source for demographic characteristics of offenders is the arrest data, and because of the confounding in arrests between individual crime-committing propensity and police arresting propensity, it is important to calibrate that bias, and to ascertain whether such bias really exists. One initial step that should be pursued is the use of the offender identification information in the victimization surveys for such calibration. To the extent that the offender demographic distributions in the arrest records and in the victimization survey are consistent, then there would be greater confidence in both sources. If they differ, then a number of potential biases in both sources would have been identified and become a subject for subsequent examination.

2. Cost and Resource Planning and Assessment Using the JUSSIM Model

It would be very useful to develop means for making cross-sectional comparisons of the behavior of the criminal justice system across jurisdictions and for making longitudinal comparisons over time within individual jurisdictions. These comparisons would reflect changing cost intensities and their relation to changing behavior of the CJS, as reflected by probabilities of penetration or drop-out at various CJS stages. An appropriate tool for such a comparative analysis would be the JUSSIM model, in which the criminal justice system is characterized in terms of the flow of people through a sequence of CJS stages, with resources applied to the flow along the path. This general structure is already available and has been implemented with extensive data bases in a number of jurisdictions. The basic structure is compatible with the OBTS system, which can provide many of the parameters as an aggregation of data on individual flows through the CJS. The JUSSIM model, or some variation of it thus represents a well-defined structure for organizing data, is compatible with developing data sources, and so represents a useful tool for undertaking cross-sectional, interjurisdictional comparisons, and for longitudinal tracking of the changing operation of a particular criminal justice system over time. It permits observing the changing parameters as an indicator of time trends in various criminal justice systems and for comparison among the systems. It also represents a means for integrating evaluative information derived from the implementation of various new programs in various jurisdictions in order to assess the impact of identified evaluative changes on the total criminal justice system.

Thus, a major research effort can and should be organized devoted to develop as a pilot project in a number of jurisdictions a JUSSIM model of a reasonably aggregate form (e.g., with about 10-20 stages and about 5-10 resources).

That model, with consistent, compatible, and comparable parameters, would provide a basis for examining longitudinal trends within the jurisdiction and cross-section differences across jurisdictions. Once such a model was developed for a single jurisdiction, it would provide a valuable tool for that jurisdiction to introduce the evaluative information on specific programs to assess the impact of those programs on other components of that jurisdiction's criminal justice system. The evaluative information used need not be derived from that particular district. In fact, an important national need is the provision of evaluative information from jurisdictions in which evaluation is conducted to other jurisdictions so that the latter can more effectively and efficiently assess the impact of particular changes in their own operations.

In this way, a clear national opportunity exists to foster more systematic planning by making effective use of the developing OBTS data, to disseminate evaluative information in a useable form, and to provide a national clearinghouse of compatible system parameters.

3. Measurement of Intervention Rates

One important aspect of the social costs imposed by the criminal justice system is the intervention it makes into the lives of citizens in terms of disrupting their normal activity. These costs are borne by those who violate criminal laws as well as those who are called upon to serve in the crime control function as witnesses or jurors. In general, minimum intervention would be desirable for all groups.

Current arrest rates (over 9 million in 1974) and projected lifetime arrest probabilities (exceeding 50% for males), indicate that the rate of intervention of the criminal justice system is significant and has been increasing. A

longitudinal examination of variables such as arrest rate, lifetime arrest, conviction, and imprisonment probability, and imprisonment rates, are all important indicators of the degree of intervention of the criminal justice system into the lives of those who violate the law. If such indicators are found or projected to be too high, then they might stimulate reconsideration of the appropriate scope of the criminal law.

From a prospective of cooperating citizens, it would be important to estimate the degree of imposition by the criminal justice system on witnesses and jurors, and the effects of frustration with the CJS that would be reflected in failure of witnesses to appear, requests for a dismissal from jury duty. The frustrations would be reflected by the inefficiency in the use of such citizen service (e.g., as shown by the ratio of time effectively serving to total waiting time) and the opportunity costs of the time spent in serving the needs of the criminal justice system.

4. Measurement of Service Effectiveness

The service role of the criminal justice system is both one of the most rapidly expanding and one that has been largely ignored. This service role generates benefits in many ways, including the social benefit of the services themselves, increased willingness to cooperate with the criminal justice system in various ways (e.g., by serving as informants, witnesses, jurors, community treatment volunteers, etc.), and may even have a direct effect on crime-committing propensity.

Measures of the service effectiveness must be developed, including a typology of service activities, an assessment of the time allocation by each CJS agency to the different service functions, and an evaluation of impact of this

time allocation, including the development of appropriate performance measures for service effectiveness. With such a structure, then an appropriate reporting system can be developed.

5. Public Perceptions of Safety and the CJS

Certainly one of the explicit desires of the CJS has been to reduce the objective rate of crime, even though achievement of that goal has been somewhat elusive. In addition, however, the public's perception of personal safety, an important component of the personal quality of life, must also be recognized as an important measure of CJS impact. What is needed is a survey form that probes individuals' perception of their safety, relates these perceptions to the objective measures of victimization risk (controlled for demographic and geographic characteristics), and to objective measures of behavior that might be related to perceived risk (e.g., use of the streets at night). With such instruments, an important additional measure for longitudinal examination will have been provided, along with reported crime rates and surveyed victimization rates to measure output associated with CJS performance. With such instruments, controlled experiments (e.g., street lighting experiments, preventive police patrol experiments) would then have another important dependent variable to be measured.

In addition, such instruments should measure the public's rating of CJS agencies, their personnel, and their attitudes, especially as those ratings result from encounters with police, prosecutors, courts, or corrections. These measures could then be related to more objective measures of citizen interaction with the CJS, such as non-reporting rates of crimes, requests for excuse from jury, and rate of witness no-shows.

6. Measures of CJS Equity

One approach to the determination of equity in the treatment of individuals by CJS agencies is the degree to which "inappropriate" considerations, such as the individual's race or income, influence the decisions made by the CJS. This measure of inequity could be obtained by regressing the various CJS decisions against functional forms that include both appropriate variables (e.g., prior record, seriousness of offense, etc.) as well as inappropriate variables. If a significant weight attaches to the inappropriate variables, that would be an indicator of inequitable treatment on the basis of the variables thus identified. Cross-sectional analyses would indicate any geographic aspects of such inequities. Perspective across different parts of the criminal justice system would indicate whether such inequity was more associated with one or another part of the system. The data for undertaking such analyses would normally be available from any reasonable offender-based transaction statistics system.

PERFORMANCE MEASUREMENT AND THE CRIMINAL JUSTICE SYSTEM

I. INCREASING COMPLEXITY OF CRIMINAL JUSTICE SYSTEM PERFORMANCE

Public systems are increasingly being faced with the need to develop and identify performance measures. This need is driven in part by the pressures fostering greater government accountability, in part by the increasing concern for the quality of performance and service of public agencies, and in part by the growing recognition, (reflected in such movements as the MBO) of the benefits simply of displaying to public managers measures reflecting the quality of their performance as an important stimulus to encourage them to revise their performance to improve those measures.

Any performance measurement of a system should be derived from consideration of the identifiable functions of that system. The criminal justice system, comprising the agencies of police, prosecution, courts and corrections, serves a complex role in American society. Traditionally, its most explicit function is as a social control system associated with enforcement of those laws which prohibit an identified set of legislatively defined crimes. As a consequence of performing law enforcement, the system is put in a position of providing a variety of additional functions, including important roles in conflict resolution and emergency service to the citizenry.

The conflict resolution role is thrust on the criminal justice system, most particularly the police, by being called into a dispute between two or more individuals. In that role, police, with their authority, status, and their legal right to intervene, may simply interpose between the adversaries

while they proceed to settle their conflict peaceably. The CJS might also be called in to exercise a judicial role for actual resolution of the conflict, when the conflict has escalated to the point where third-party adjudication becomes necessary.

Increasingly, attention is being given to the general emergency-service role of the criminal justice system. In that role, the CJS is viewed as providing benefits to the general community, not so much through its intervention or law enforcement roles, but because the system - again, the police especially - represents a resource that is always available and that has access to both special skills and equipment. Such emergency services might include medical aid, animal rescue (the proverbial cats in trees), or opening locked doors.

The increasing complexity of these roles requires that any measurement of performance take account of this rich dimensionality, especially in the aspects that represent a departure from the more conventional crime-control roles. Any longitudinal assessment of resource utilization, for example, must account for the increasing use of CJS resources in these other roles.

A. The CJS as a Crime-Control System

The traditional role of the criminal justice system is as a social control system, principally to protect the middle class citizens from others who would harm them or deny them their rightful property. This continues to be viewed as its primary role, both by most citizens and by most participants in the system. Priority attention must be paid to victim crimes, responsiveness to victims' calls for assistance, and the apprehension of predatory offenders.

The quality of performance on these dimensions is reflected in measures like crime rate. Because of the many social factors outside the control of the CJS that influence crime rate, CJS variables that presumably influence crime rate are much more commonly used. These include measures like responsiveness (e.g., response time), or apprehension capability (e.g., clearance rate). The important uncertainty here is the nature of the relationship between crime rate (which is unquestionably the conceptually appropriate measure) and these variables, which often serve as proxies.

Despite the extensive amount of work on measuring the crime-related performance measures of the criminal justice system, there has been very little effort to relate such performance measures to the policy variables involved in the management of the criminal justice system. Only in recent years has there been a significant attempt through econometric regression models to relate crime rates to various sanction variables and resource commitments, but the effort in these directions continues to be quite primitive. This is certainly a result of data problems but, much more fundamental is a result of the methodological complexity resulting from the variety of other factors intervening between the policy variables and crime consequences, or because similar forces drive both crime rates and the policy decisions, thereby confounding the relationships and making causal inferences extremely difficult.

In view of the extensive literature on the subject of crime rates and crime measurement, the central issues in that literature - index questions and non-reporting problems - are not of central concern in this paper. An extensive body of literature starting with Sellin and Wolfgang's The Measurement

of Delinquency covers the question of crime index. The questions related to non-reporting (the attempts to estimate the "dark figure" of crime) received strong impetus from the work of Reiss and Biderman, and has resulted in the major LEAA/Bureau of Census victimization surveys.

The important needs on the problem of crime relate much more to developing meaningful and validated proxies and for inquiry into the larger question of the criminality in a society and the trends in criminality.

Another large portion of the activity of the criminal justice system is concerned with the class of law violation generally designated as "victimless crimes", where the police intervention is triggered by police initiative rather than by a report or call by some citizen victim. Crimes such as drunkenness, private sexual misconduct, drug abuse, gambling, and prostitution, all involve violation of explicit statutes, but the behavior is voluntary and consensual, and so enforcement of those laws involves a variety of police stratagems to catch the participants in the behavior in order to bring about a legal prosecution and conviction. In many of these cases, the criminal activity simply involves participation in an illicit market process for which a demand exists, and an illicit supply emerges to meet that demand, a characterization that has been articulated very well by Thomas Shelling.

The problem in measuring the volume of these victimless crimes is that the number of such events that come to official attention are much more a function of the intensity of police activity directed at this behavior than of the amount of the behavior. The amount of police resources applied to the problem and their diligence and subtlety in tracking down

and dealing with the activity are very influential in affecting the arrest statistics. Any indicators of that behavior that relied on arrests as an indicator would miss the basic point that volume depends on the intensity of intervention rather than the intensity of the behavior.

Recent trends toward decriminalization by statute, (especially marijuana) or by practice (especially for consensual sex offenses) made this an increasingly important area in which indicators of both the volume of such behavior and the level of intervention are becoming necessary. Naturally, perception of citizens and attitudes toward this set of crimes must play an increasing role in the assessment of their seriousness and development of strategies to deal with them.

B. Conflict Resolution Role

The police and the larger criminal justice system are increasingly assuming a role in which they are being called upon to resolve conflicts. Many of these involve disputes between husbands and wives or between people who otherwise have a prior relationship. The police are called into this role because of their ready availability, and accept it willingly (e.g., Morton Bard's Crisis Intervention Units) because an early successful intervention can prevent the conflict from escalating into what might otherwise become a serious assault or homicide. In part, this expanded role is being thrust upon the criminal justice system because of the gradual disappearance of the extended families and the other cohesive social organizations like clubs or community churches that formerly performed that function.

Also, as more resources have become available, particularly to law enforcement agencies,but also to courts and corrections, it has become a more plausible and possible role for these agencies. The patrolling police officer represents an always available resource. With his authority as a disinterested outsider who can also pose a potential threat of punishment to any participants who refuse to resolve the conflict, he has the capacity to impose a resolution of the conflict if necessary.

Since these conflict-resolution roles are very different from the traditional CJS activities, new strategies have evolved for handling them. These strategies have included diversion of juveniles to programs other than detention homes or juvenile court; special teams of police to handle domestic quarrels; and referral of family problems to the appropriate social agency.

The increasing attention to this conflict-resolution role, and the desire to distinguish it from the crime-control function in order to avoid undesirable labelling will require a separate system of indicators. These would have to describe the types of incidents, the nature of the interventions by the police and other CJS agencies, the effectiveness (particularly in terms of more serious crimes averted) of various forms of intervention, and the resource commitments and consumption by the various agencies.

C. Service Role

The criminal justice system, and the police in particular, have always had an important emergency-service role in the community. In part, that service role has derived from their availability around the clock, and from

their access to special skills and knowledge as well as to special resources that might be needed in an emergency situation. This skill and availability is particularly needed in an urban society, where individuals are increasingly isolated and so are unlikely to call on each other for aid, and where the specialization of the urban society has resulted in less and less self-sufficiency in handling the many complexities that confront an individual in an emergency. Thus, more and more people find themselves dependent upon the police in such a service role to perform such functions as opening their locked house or car doors, rescuing cats from trees, finding missing persons, and handling medical emergencies. Police handle this role with a willingness that ranges from reluctance or refusal (in those departments whose model is purely law enforcement) to encouragement (by those who feel the need for generating closer citizen identification with the police, even if only to enhance citizen cooperation as a source of information and support for other law enforcement operations).

II. GENERAL OVERVIEW OF CJS OBJECTIVES AND EFFECTIVENESS MEASURE

The primary goal of the criminal justice system is crime reduction, and so at least some components of the measure of effectiveness must address the issues of crime control. In addition, there should be appropriate accounting for the other benefits derived from the system. One might specify, then, the total measure of effectiveness of the criminal justice system as that of minimizing the sum of 1) the social cost of crime plus 2) the social cost of crime control minus 3) the service benefits. In this section, we seek to explore the various components of this aggregate performance measure as a precursor to the design of explicit projects intended to measure the components.

As additional resources are committed to the criminal justice system (i.e., as the intensity of crime control increases) one would expect to find the following effects:

1. A reduction in crime, and in the associated social cost of crime;
2. An increase in the social cost of crime control, including considerations both of economic expenditures and of the degree of individual liberty.

More generally, we might expect that, as resources are added at low levels of crime control, the total social cost should decline (because the crime-reduction effects dominate) whereas at high levels of crime control, the total social cost should increase (because the crime-control costs dominate). This indicates that if one could compare the different social-cost components here, one would find an optimum level of crime control intensity, at least conceptually. At the low levels of crime control, we suffer too much from crime, and at high levels, we suffer the cost of too much crime control.

In examining the different components in this social-cost format, we seek to highlight and expand on the basic issues in order to identify means of addressing and measuring the various components involved. Clearly, precise measurement will never be possible, but more relevant and appropriate surrogates can be developed for each of the areas.

A. Social Cost of Crime

In exploring the social cost of crime, there are two rather different perspectives that are involved. One is the concept of risk of victimization, or the likelihood that an individual,

particularly of an identifiable demographic and socio-economic group, is likely to become a victim. Here the issues relate primarily to measures of exposure and vulnerability.

The second perspective is the concept of "criminality", a reflection of the degree to which individuals are becoming more crime-prone, and this involves a concern for the weakening of the nation's social fabric. This concern is less personal than the concern over victimization, but it impacts the society itself very deeply.

These two concepts are indistinguishable in terms of the gross crime rate, (i.e., the number of crimes per person in the population) since the gross crime rate is a reflection of both the aggregate victimization risk and of the aggregate level of criminality. Because of the gross character of this highly aggregated crime rate, it fails to provide adequate fine-structure information necessary for dealing in more detail with either the victimization or the criminality considerations. A finer treatment is required.

1. Victimization

The question of victimization has been dealt with extensively in the LEAA/Census Bureau Victimization Surveys, and it is not a primary consideration in this paper. It is important to note, however, that efforts to estimate the victimization rates have resulted in the commitment of approximately $10,000,000 per year in victimization surveys to collect data on only this one aspect of the performance of the criminal justice system in order to provide much better calibration of the true crime rate, and to permit separation of the trends in crime rate from trends in reporting.

It is also interesting to note that the non-reporting of victimization events to the police, which was the prime stimulus for victimization surveys, is strikingly consistent from city to city in the surveys, and not strikingly different among demographic groups. Thus, at some time in the near future, it should be possible to adequately calibrate the non-reporting rate by crime type, to estimate trends in the non-reporting rate, or to conduct victimization surveys at a much lower frequency (say, every five years) than at present.

2. Criminality

The fundamental issue in attempts to measure changing "criminality" in the society is the need to partition changing crime rates between the changing demographic mix in the society and the changing propensity to commit crime by a person of a specific demographic group. Between 1960 and 1970, for example, the number of reported index crimes increased by 178% in the United States. A critical question for those concerned with this rapidly changing crime rate was the degree to which that increase could be attributable primarily to the changing demographic mix (e.g., to the much larger proportion of young people in the population) or to a changing criminality (e.g., to any increase in the likelihood that a 23 year-old, white male would commit a crime). To the extent that the trend could be attributed to the changing demographic mix, then the effect was likely to be reversed in the 1970's because of the end of the "baby boom", and there was little reason for serious concern about a general weakening of the social fabric in the nation. To the extent that some demographic groups were increasing in criminality, that would be an issue for specific focus. If there was a more general increase in criminality across demographic groups that might reflect a more serious decay in the social structure and possibly a matter of concern.

In some initial studies in Pittsburgh (Blumstein and Nagin) these issues have been explored and it was found that arrest trends in Pittsburgh were attributable almost entirely to the changing demographic mix rather than to any increase in criminality. This effect, of course, cannot distinguish between true changes in criminality and changes in police propensity to arrest a person of a specific demographic group. It was found, for example, that the only demographic group with an identifiable increase in criminality was the young white females. Since an offender's group association can only be measured from arrest data (uncleared crimes do not have an identified demographic group for the perpetrator), the trend in estimates of the criminality of young white females precludes distinguishing between a greater propensity to commit such acts (undoubtedly part of the trend towards social equality between the sexes) or the increased willingness of police to arrest a woman for an act for which they might have let her go at an earlier time.

Statistical methods are available for beginning to address the question of criminality more effectively than has been done in the past. This issue, which parallels victimization as an issue of social concern, should be pursued much more diligently, particularly in light of the intensity of the effort on victimization studies. The offender information in the incident reports of the victimization survey could be used to distinguish the components of true demographic criminality and police propensity to arrest. The combination of the arrest and victimization information, together with improved application of statistical methodology for identifying the various time trends and for discriminating the various effects on reported crime rate and arrest rate should reveal some valuable insights about criminality, demographic mix, and their combined effect on crime rate.

B. Social Cost of Crime Control

The effort to control crime represents a significant generator of costs, both economic and social. Fourteen billion dollars were spent in FY 1973 on the maintenance and operation of criminal justice systems at all levels of government.

In addition, there are the many less direct economic costs of reduced production by prisoners, witnesses, and jurors, and welfare costs to maintain prisoners' families. Perhaps even more important are the intangible social costs deriving from the efforts at crime control. These derive from the consequences of erroneous arrests or convictions, from the loss of individual autonomy, and from the social conflict resulting from enforcement of laws deemed to be unjust or outside the proper bounds of the criminal law. More generally, the criminal justice system enforces the law through various forms of intervention in people's lives. In some ideal sense, it would be desirable if the laws could be obeyed while minimizing the degree of such intervention. To the degree that social norms were widely adhered to, for example, without much CJS involvement, that would clearly be preferable - on both counts - to less adherence and more intervention, and the inequities in that intervention.

In this section, we address some of these issues in turn, dealing first with some of the economic costs, and then proceeding to a consideration of the more complex social costs associated with the degree of intervention of the CJS in the lives of citizens.

1. Direct Operating Costs

The combined CJS machinery represents a major public sector expenditure. The budgeting for criminal justice activities involves a number of different

crime control agencies, at different levels of government, which are both functionally and administratively separated from one another. This arrangement has contributed to a very fragmented view of the benefits and drawbacks that result from the monies spent. It is only recently that these costs and benefits have been analyzed from the perspective of the total criminal justice system. Such a system-wide view includes a concern for issues such as the following:

(1) the distribution of total costs among the police, prosecution, court, and correctional subsystems of the CJS;

(2) the distribution of these costs over the various crime types; and

(3) the interaction of the effects of changes in expenditures in one part of the system on resource requirements in the others.

The actual level of costs for the total system and for the individual subsystems, and for individual crime types,are all of interest. Aside from providing a more detailed and complete description of the different factors contributing to costs, they also force the public and the system managers to face up to the budget implications of current or proposed policies. Knowing that the prison costs are $10,000 per man-year, or that the average projected cost per marijuana arrest is $1,000 may stimulate a reconsideration, at least at the margin, of current policies. In any case, these budget implications should enter the choice process, and they rarely do because they are so often unknown.

In addition, detailed cost measures could furnish inputs for cross-sectional cost-effectiveness analyses. Cost-effectiveness measures typically indicate the rate of benefits received per dollar spent. A variety of

possible measures associated with criminal justice activities could be postulated. The number of crimes solved/dollar spent, the number of arrests/dollar spent and the number of convictions/dollar spent are just three. In all instances the measures can reflect total CJS expenditures or the expenditures on selected subsystems (e.g., the police or courts).

These measures could also be computed for individual crime types, thus allowing comparison among the crime types. It is a peculiar feature of criminal justice operations that the most common offenses (like drunkenness) which may represent the most expenditure are also typically of least concern to public safety.

Estimating the individual costs associated with different crimes will provide policy makers with some insights into the trade-offs they must make. Given their estimate of public priorities, they and an informed public must then decide whether the present allocation is optimal and whether the potential gains, in say, robbery arrests, convictions, or imprisonments to be derived from a reallocation, are worth the effects of decreased control of lesser offenses.

Sometimes the objective may be to reduce costs while maintaining the same level of control. This might be achieved by using less expensive resources wherever possible. For example a magistrate's trial may serve just as well as a trial presided over by a judge, particularly for those kinds of offenses for which defendants regularly waive their right to a jury trial, plead guilty, and the typical sentence that is imposed by judges is well within the sentencing prerogatives of magistrates. Also, for some offenders and/or offense types, sentences of fines or probation which require less CJS resources

may be just as effective in crime control as a prison term.

Alternatively, the same control might be achieved by shifting the major responsibility of control to agencies outside the CJS. This would greatly reduce the cost of this control function for the CJS (although the total societal cost may remain the same or even increase). An example of this is the creation of alcohol treatment centers within the mental health system and the diversion of the control of drunkenness to these centers,with the police performing the more limited function of referral agents.

It is also possible, however, that the consideration of costs may uncover instances in which control is being maintained where there is no longer a perceived need or desire for that control. In this case,the costs can be reduced or even eliminated by easing the level of control exercised. This is most likely in victimless crimes where the values and general attitudes of the public have changed,and behavior that was formerly considered criminal is now regarded as acceptable. Homosexuality is one such marginal behavior that is rapidly losing its serious deviant characterization. Truancy is a similar offense that represents a significant burden on some juvenile justice systems, but which is not often regarded as an indicator of criminality. Clearer indication of the expenditures on such crime types will fuel the process of reconsideration of their appropriateness to the criminal justice system.

In summary, then, performance measures that involve the CJS operating costs can be useful. Simple itemization of costs and partitioning of these costs among the various crime control activities and among the different

crime types can illuminate current practice or possible changes. A variety of cost-effectiveness measures can be derived to permit a comparison of the relative effectiveness achieved for the money spent in different jurisdictions, in one unit over time (accounting for inflation, of course), and over different parts of the CJS.

While a cost-effectiveness analysis is a useful tool for comparing the relative benefits derived from various expenditures, the results of such an analysis must be used very carefully in deciding on the allocation of future expenditures. Since the main objective of criminal justice systems is not to maximize monetary profits, operating in accordance with standard economic incentives may often be at odds with the fundamental purpose of crime control. In many decisions, such dimensions as equity and humaneness may be of substantially greater import than the economic measures.

2. Degree of Intervention

In many ways, the ideal criminal justice system is one that never has to be employed, but simply provides symbolic guidelines to shape individual behavior. Because of the infeasibility of the ideal, the criminal justice system is a source of significant intervention into many people's lives in a variety of ways. To the extent that individual autonomy is valued, this intervention represents an important intangible cost of the control function. Some reduction in autonomy, however, represents a price that must be paid for living in society. Furthermore, effective socialization in which the behavioral expectations of society become the personal expectations of individual members can greatly reduce awareness of the external controls and simultaneously moderate any sense of diminished personal freedom.

Intervention is also experienced by those who obey the law scrupulously. Many citizens are called upon to serve the system as a juror or as a witness. There is an immediate economic cost associated with that service reflected in loss in normal productive employment. More to the point, this process generates, in many cases, an even greater social cost associated with the frustration in being used and manipulated by the system or in being the victim of its inefficiencies. For example, a juror who spends most of his time serving on a trial is likely to find the experience interesting, stimulating, and educational. On the other hand, if he is forced to spend the great majority of his service period sitting in a jurors' room waiting for a case, he will undoubtedly leave with a feeling of disillusionment, frustration, and contempt for the inefficiency of the judicial system. Improved methods of scheduling jurors are being developed, and an important indicator of that efficiency would be reflected in a measure like the percentage of juror man-hours spent in actual jury service.

A related but different problem pertains to witnesses. As with jurors, one could simply record the amount of productive time lost by witnesses in the process of cooperating with the criminal justice system. There is, however, an added dimension of frustration associated with aborted hearings at which a witness appears, but the hearing is continued. Measures like the mean number of continuances per case reflect the resulting frustration. The number of times witnesses fail to appear serves as a more direct indicator of the consequences of such frustration.

Aside from those considerations associated with the regular operation of the criminal justice system, the resource demands it imposes, and the problems

it creates for its participants, there are separate considerations reflected in the degree of intervention of the criminal justice system in the lives of American citizens. Even though those arrested might well be violating some law in some form, there is considerable discretion over the intensity with which many criminal laws are enforced.

As the major focus of the attention of the system is directed to an increasing number of minor offenses, this diminishes the importance of its attention to serious offenses. To the extent that there is a general perception that offenses of very different seriousness are being treated similarly, there may be a feeling that the more important offenses are not being handled with sufficient concern. This may precipitate a loss of faith in the CJS as a means of social control. This devaluation of the importance of CJS actions can be avoided by restricting the invocation of at least some responses to those offenses and offenders of greatest public concern.

To some extent the general increase in the numbers of persons subjected to CJS intervention may reflect the extension of the "deviance" label to a wider range of acts and people. In part, this may be an indication that the system is out of synchrony with public attitudes. This situation generates public cynicism about laws in general and the agencies created to enforce them. As behaviors become widespread in acceptance and yet continue to be treated as criminal, this encourages disrespect for other laws and for the CJS.

The CJS also represents the <u>last resort</u> in the system of social control. It is intended to be invoked only when a variety of informal mechanisms fail. Relatively recently, this last resort has been increasingly relied on as the only resort. This greater dependence on the CJS leads to additional problems

for social control function. Not only does it further weaken the potential impact of less formal mechanisms of control, it also reduces the effectiveness of the CJS itself. As its use becomes common, it also becomes cheapened and the moral authority of the CJS as an agent of social control is diminished.

Clearly, then, there are necessarily limits on the extent of intervention by the CJS and it is important to monitor this intervention and establish boundaries. Intervention, both in lives of clients and those who serve the system, represents a unique opportunity to either enhance or degrade the moral authority of the CJS. The impression the CJS leaves on individuals as a result of the character and outcome of their interaction with the system is a significant factor influencing their future relationships with it.

For many citizens, their primary involvement is as one of the 10 million arrestees recorded each year. Thus, an important consideration in the operation of the criminal justice system relates to the simple number of arrests, the relationship between those arrests and crimes committed (e.g., clearance rates) and the mix of "discretionary" arrests to arrests for victim-type crimes (e.g., the index crimes, fraud, embezzlement, etc.), and the probability of arrest some time in an individual's lifetime.

The total arrest rate and its trend are important measures of the degree of intervention. Over the 1960-70 decade, the total number of arrests increased by about 30%, the index crime arrest rate went up by about 85%, drunkenness arrests went down by about 20%, and the arrest rate for gambling (another discretionary crime) went down by about 40%. We see then a situation in which there is considerably more arrest involvement of people with the criminal justice system, but that the <u>discretionary</u> intervention for other than narcotics

has, in fact, declined. This decline may well be a consequence of the adaptation within the CJS as it had to deal with more serious crime, but both indicators are necessary measures of the degree and kind of intervention.

More generally, the degree of involvement of clients of the system should be traced over time at each of the various stages at which they get involved (i.e., as an arrested suspect, as a charged defendant, as a convicted offender, as a prisoner, or as a parolee or probationer under community supervision). A time series that investigates each of these indicators of involvement both in terms of individual variables and in their relationships to crime rate and to each other, particularly for different kinds of crimes, will provide an important indication of the degree of CJS intervention.

These variables and their ratios also could serve to reflect changes in the operation of the various portions of the criminal justice system. It will never be clear whether an increasing ratio (say in the convictions per reported crime) is desirable or undesirable, since there will always be at least two conflicting interpretations of such trends. On one hand, upward trends may reflect an increase in the effectiveness of the control system; on the other hand, downward trends might reflect an increase in the delivery of due process, justice, and more equitable trial. Such effects are possible to locate in the system and it is worthwhile finding them, but the interpretation of goodness or badness will inherently continue to be elusive.

3. Equity of Treatment

Clearly, one of the objectives of a justice system is that justice be delivered in equal measure to all who come before it. This is not to say that two individuals charged with the same offense, but with significantly

different prior criminal records should not be treated differently. Indeed, a central concept in our punishment system involves one of accumulation of vulnerability to punishment, so that the individual with a long criminal record can expect to experience more severe punishment than a first-time offender. Rather, the concern is over disparity in treatment between people with the same kind of offense and record.

With the accumulation of computer-based information about suspects and offenders, there is a growing trend toward using this information to predict future criminality. These predictors use factors such as length of prior criminal career, number of serious convictions, involvement in drug use, sibling involvement in crime, etc. While there may be some valid legal debate over whether such predictions are even legally tenable in considering an individual's punishment (an issue that has been raised by the "just desert" school, including von Hirsch[1] and Morris[2]) it is clear that they do have some predictive capability. Furthermore, regardless of how they are used, there is widespread agreement that the use of certain variables in punishment decisions is at least inappropriate (race, for example), and should be irrelevant (for example, income level).

One could explore this question of equity by examining the difference in the treatments, say, of blacks and whites. But that comparison should reflect the fact that blacks tend to commit different kinds of crimes than do whites, and, if such crimes are more worthy of severe punishment, these effects might be responsible for any differences in aggregate punishment rates for blacks and whites. Much more careful analysis is required to isolate these separate factors. Conceptually, one can imagine regressing

---

[1] vonHirsch, Andrew, "Prediction of Criminal Conduct and Preventive Confinement of Convicted Persons", Buffalo Law Review.

[2] Morris, Norval, The Future of Imprisonment, University of Chicago Press, 1974.

the CJS decisions against a group of "appropriate" variables and a separate group of "inappropriate" variables (like race or income status). An equitable system should find all the weight on the "appropriate" variables, and should have no significant coefficients on the "inappropriate" ones. To the extent that this is not the case, i.e., that there is a significant weight on the inappropriate variables, then this should be further investigated to see if there is some evidence of discriminatory treatment. Corrective measures could then be taken. The magnitude of the weights on the inappropriate variables would serve as an important equity indicator reflecting the kind of discrimination (which inappropriate variables have significant weights) and the magnitude of discriminatory inequity (the size of their coefficients).

A separate dimension of the equity question relates to the disparity in treatment, even though there is no consistent bias towards one or another group. Thus, in that same concept of a regression relationship, two different jurisdictions may well both have insignificant coefficients associated with the race variable, but the residual variation could be much greater in one than in the other. This greater variation could result from a variety of factors, but certainly important considerations would include differences among judges, or differences in the urban composition of the areas. In such an analysis, for example, one could identify a reasonable set of "appropriate" variables that should explain most of the variation in criminal justice decisions. To the extent that the residual variation is large, then that suggests considerable disparity (and perhaps arbitrariness, albeit without discrimination) in the CJS decision making or local conditions. Thus, an important indicator would

be the residual unexplained variance (i.e., $1-R^2$ for the regression equation that includes all the"appropriate" variables).

4. Correctness of Decisions

Ideally, one would hope that the decisions made within the criminal justice system are correct in that they reflect a proper and careful weighing of the relevant facts and take into account the appropriate considerations of due-process.

If one had an independent standard of comparison for each such decision, then one might generate indicators of the relative frequency of such "correct" decisions. Since there are no external independent standards for the correctness of decisions, however, there is inherently no good basis for challenging or validating any such decisions. Thus, we argue that any attempt to develop an indicator of the correctness of decisions is probably doomed to frustration.

C. Public Perceptions of the Criminal Justice System

In addition to the objective measures of performance which have been considered above, it is particularly important to deal with public perceptions which may or may not have a strong relationship to such objective measures. In fact, much behavior is shaped by people's perceptions of issues related to crime, crime control, and personal safety and by their perception of the integrity of the criminal justice system. For example, effective performance of the criminal justice system is certainly influenced by the willingness of citizens to cooperate with CJS agencies, through reporting crimes to the police; through providing information about likely suspects; through cooperation with the courts by acting as witnesses or jurors; and through

cooperating with corrections agencies either through volunteer treatment groups or simply as community residents willing to tolerate community-based treatment programs. This readiness to participate in itself reflects these perceptions, and has the advantage that it is linked directly to public behavior rather than the more vague variant of opinion and attitude measurement.

In addition, the public's fear of crime, whether objectively based or not, can well lead to behavior that is itself crime-generating, or that reflects negatively on other components of the quality of life. Frightened citizens, for example, are more likely to purchase weapons, legally or illegally, and these weapons could be misused either in an accident or in some criminal act. Citizen fear could lead to individuals barricading themselves within their own homes, thereby restricting their purchases at commercial establishments or their use of recreational opportunities in a city, and so decreasing the benefit to the public and increasing their cost by abandoning the streets and recreational facilities to those who would wreck havoc. Simply reducing the level of fear would increase the patronage and traffic and may thereby reduce the objective measure of crime in those areas, presumably reducing any prior feeling of insecurity. Extreme cases of feelings of insecurity could trigger residential movement from the areas of perceived high crime into more suburban or rural areas, or to other metropolitan areas where there would be less of a feeling of insecurity. These effects on migration could also be measured and used as an indicator of public insecurity in certain areas if linked to relevant attitudinal measurements.

More narrowly, even the budget performance of criminal justice agencies or LEAA could be influenced by citizens' perception of their effectiveness and of their integrity in dealing with many problems related to crime and its control.

The kinds of indicators of citizen perception of security and of the criminal justice system would include 1) indicators of citizen initiative in providing personal security; 2) attitudes towards the criminal justice agencies; 3) attitudes about the "crime problem"; 4) the willingness to participate with the criminal justice system; and 5) attitudes towards the individual criminal justice subsystems. The data for such indicators would be derived both from the objective measures of certain attitude-related behavior and from explicit attitude surveys designed to elicit citizens' views and perceptions.

Measures that reflect citizen anxieties and concerns would include sales of various forms of security systems including burglar alarms or weapons as positive actions (perceived to be serving self-protection functions), and presumably therefore reflecting a decreased confidence in the governmental mechanisms for providing such protection.

Attitudes towards criminal justice agencies could be probed by inquiring into people's willingness to support an increase in taxes, and by their attitudes toward using new tax revenues for such purposes as increasing police salaries, hiring more policemen, prosecutors, judges, probation officers, or prison guards, or for building additional prisons.

The attitude towards the crime problem could be probed by inquiring into citizens' views of how the criminal justice system can become more effective

both in the direction of enhancing justice and individual liberty as well as increasing the effectiveness in crime control. This might be done by obtaining perceived trends in ratings in the quality of justice and of crime-control effectiveness.

The attitudes toward the various criminal justice subsystems could be probed by inquiring into the details and the reactions to encounters citizens might have had with the police or with courts, by their perception of the effectiveness of the correctional programs, by ranking city police, state police, district attorneys, judges, and correctional officials on such items as professionalism, corruptability, friendliness, commitment to their jobs, concern for public safety, concern for justice, and other indicators associated with each of the subsystems.

In general, then, the surveys would be directed at absolute rankings of the participants in the various subsystems. Comparison with their perceived ranking five years ago, with their perceived ranking one year ago, and with a projection of how they would rank next year could also be undertaken. Time series of these rankings should reflect changing public attitudes toward the subsystems and should be correlated with various other measures of cooperation and participation in the functioning of the criminal justice system.

In particular, the issue of personal feeling of security is a crucial factor influencing the perceived performance of criminal justice systems and should have continual monitoring, just as the objective indicators of the FBI crime index and the victimization survey. Correlation of these subjective measures with the appropriate objective indicators should be undertaken to ascertain the degree to which the subjective indicators are in fact driven by

the objective measures, and if so, whether they are driven by the published reports, or to identify other extraneous factors which drive them.

Furthermore, some key policy variables when correlated with the subjective indicators,might be found to have a very strong influence on their values. For example, police patrol methods and allocation procedures, or police response time, or police visibility in high density areas,might all be found to have a significant impact on individuals' feelings of personal security. To the extent that that happens and is shown to be the case, then that might well suggest various tactics for enhancing the feeling of security, an important function of a criminal justice system. Similarly, with regard to the courts, one might relate such features as architecture of the building or of the courtroom, and the treatment of witnesses and jurors by the clerical staff of the court,with performance measures such as the percentage of prospective jurors seeking to omit jury service,or witness no-show rates. All of these factors should influence court management policies.

This, of course, raises the much larger issue of how such measures of perception and values can be used to influence the direction and priorities for the CJS and its individual subsystems. Needless to say, there must be continual interaction between the attitude measurement and the policy directions, both for political responsiveness and for dealing with citizen concerns.

D. Conflict Resolution

Within the operation of the criminal justice system, there is a need to measure the performance in the two roles of conflict resolution and service. The first of these roles permeates the total criminal justice process including

the police, particularly in domestic quarrel situations; the family court, where informality and counseling are major parts of the operations in domestic disputes; and the supervisory part of the system such as probation, parole, and community supervision where services are offered to assist the offender in reintegrating into the community.

In measuring the effectiveness of the conflict resolution activity for the total justice process, it is necessary to address it first as an independent function. Because of the increasing involvement of criminal justice agencies in this role, and the significant volume of resources allocated to it, there is a need first to identify the volume of such activities and the agencies which perform them. From this, trends can be estimated of the demand for such service and the various forms of response to that demand. This will require independent information on the development of the demand. This will provide decision makers and the public some measure of the present performance of the criminal justice system in this role and of its potential need in the future. The effectiveness of the CJS performance in this role can only be measured by determining the success of the interventions or the initiatives taken by the system in each of these circumstances. This would include follow-up of the conflicts presumably resolved, and determination of "success" indicators in each incident.

An important rationale for the involvement of the criminal justice system in conflict resolution is its implications for crime control. The number of aborted incidents which might have escalated into criminal violence is of direct interest as a specific but indirect crime control measure. Thus, it would be

particularly important in attempting to measure the "success" of conflict-resolution activities to estimate the number of potential criminal incidents averted.

A further need in measuring the role of the criminal justice system in this area is to identify new and innovative initiatives which might be taken or are being taken on an experimental basis. Thus, a simple inventory of types of intervention associated with measures and data on their success or failure would be a useful first step.

One problem in obtaining records on such informal activity at the periphery of the CJS is that recording and monitoring may be dysfunctional to the activity itself. However, sampling of such activity, measurement on a random and unobtrusive basis, and later recording by the CJS participants could provide the necessary information without interfering with the roles.

E. Service

The role of the criminal justice system in providing service to the community has been largely ignored in most attempts to measure present levels of activity.

Actually, in trying to quantify present activities in this service role, there is a need first to identify the types of agencies involved and their initiatives. A primary agency is the police. Other community service roles are undertaken by the majority of agencies involved in the administration of justice, and those oriented towards the prevention of further offenses (especially in the juvenile area) are important to include. The volume and distribution of activities by CJS functionaries in such service roles should

be taken into account more formally so that the nature and extent of these activities can be accounted for and properly attributed to the CJS. In such measurement, the time and resource allocations to each such activity should be identified and the outcome assessed.

III. PROPOSED RESEARCH PROJECTS

In this section, we build on the previous discussion of the problems and issues involved in devising improved indicators of performance for the criminal justice system, and we identify a number of specific projects which could be undertaken to develop, test, and validate criminal justice system indicators. These projects are discussed in a form such that they might stimulate the development of an RFP for the actual performance of the task.

A. Assessment of Trends in Criminality

The basic problem of "criminality" relates to the distinction between the crime-committing propensity of a given demographic group and changes in the demographic mix in the population. Gross crime rates (i.e., crimes divided by population) confound these two separate effects. In general, the problem may be stated in terms of the following relationships:

$$N_{it} = \sum_j p_{ijt} M_{jt}$$

where N_{it} is the number of crimes of type i in year t

p_{ij} is the rate of committing crimes of type i by persons in demographic group j during year t

M_{jt} is the number of people in a demographic (e.g., partitioned by age, race, and sex combination) group j in year t.

The fundamental issue here is estimating p_{ijt} and examining its trend over time. The basic data for such an estimation derives from arrest records and these must be calibrated for the intensity of police activity.

As described earlier, these are the issues that have been addressed in the paper "Analysis of Arrest Rates for Trends in Criminality" by Blumstein and Nagin.* [In that paper, it was found that even though there was an upward trend in crime rate, that that trend could all be explained by the changing demographic composition and that the within-group crime rate (p_{ij}) was constant for all demographic groups except young white females, among whom there was an apparent increase in "criminality".] Further examination of this phenomena and the validity of the hypothesis for longer time periods, different areas and different groups should be examined.

It is inherently extremely difficult when using arrest rate information to distinguish between trends in inherent criminality and trends in police propensity to arrest people from a particular demographic group, so that those two effects will continue to be confounded. One important attempt to test for such police bias (which could well result from differential vulnerability to arrest by the different groups - e.g., the inexperience of young offenders compared to older ones - as well as conscious police discrimination) would be through comparison of the demographic distribution of arrestees with that of offenders reported by the victims in the victimization survey. To the extent that those distributions are consistent, then there would be greater confidence in both sources. If they differ, then a number of potential biases in both sources would be focused on.

B. Cost and Resource Planning Assessment Through the JUSSIM Model

It would be very useful to obtain a basis for making cross-sectional comparisons of the behavior of the criminal justice system across states or cities, and for making longitudinal comparisons over time of individual criminal justice systems. These comparisons would reflect changing cost

*Blumstein, Alfred and Daniel Nagin, "Analysis of Arrest Rates for Trends in Criminality", Socio-Economic Planning Sciences, Vol. 9, pp. 221-227.

intensities and their relation to changing behavior of the CJS, as reflected by probabilities of penetration or probabilities of drop-out. An appropriate tool for such a comparative analysis would be the JUSSIM model, in which the criminal justice system is characterized in terms of the flow of people through a sequence of stages, starting from "crimes committed" through arrest, prosecution, adjudication and corrections. In that flow process, resources are applied to the units of flow (i.e., the suspects, defendants, offenders, and prisoners) with a specified workload (i.e., the amount of resource time consumed at each stage by each unit of flow) at an associated unit cost. The flow between stages is characterized by "branching ratios" which depict the proportion of flow from each stage to each other stage (e.g., the proportion found guilty at a jury trial, and the proportion found not guilty).

This general structure (with parameters of branching ratios for each crime type, workload for each crime type, unit costs for each resource) is already available and has been implemented with extensive data bases in a number of jurisdictions. The basic structure is entirely compatible with the Offender Based Transaction Statistics system, which can provide many of the JUSSIM parameters, especially the branching ratios, as an aggregation of data on individual flows through the CJS. Thus, it is to be expected that JUSSIM parameters will become increasingly available as the OBTS is implemented more widely.

The JUSSIM model was developed as an interactive computer program in order to facilitate its use as a planning tool for dealing with the total criminal justice system. As such, it is expected to see increasingly intensive use

as the data for using it becomes more widely available. At the same time, it represents a well defined structure for organizing data, and so represents a tool for cross-sectional, inter-jurisdictional comparisons, and for longitudinal tracking of the changing operation of a particular criminal justice system over time. Thus, it represents a device for

- observing the changing parameters (i.e., branching ratios, unit costs, workloads, all by stage or by crime type) as an indicator of time trends in various criminal justice systems and for comparison among them;
- integrating the evaluative information derived from the implementation of various new programs in order to assess the impact of identified evaluative changes in the criminal justice system;
- testing the relation between system resources and system behavior. Recent econometric studies have postulated certain functional relations but more in-depth and comprehensive examination is required to determine exactly the relation between system performance and cost. This in itself is a rare opportunity to examine the CJS as a large social institution characterized by a well-defined and comparable set of parameters.

Thus, a major research effort can and should be organized devoted to develop as a pilot project in a number of jurisdictions a JUSSIM model of a reasonably aggregate form (e.g., with about ten to twenty stages and about five to ten resources). That model, with the consistent, compatible, and comparable parameters, would provide a basis for examining the longitudinal trends within a jurisdiction and the cross-sectional differences across

jurisdictions. Once such a model is available for a single jurisdiction, it would provide a valuable tool for that jurisdiction to introduce the evaluative information on specific programs to assess the impact of those programs on other components of that jurisdiction's criminal justice system. It also provides an appropriate tool for the introduction of test information on system changes experienced in other jurisdictions.

The evaluative information used need not be derived directly from that particular district. In fact, an important national role would be the provision of information from jurisdictions in which evaluation is conducted to other jurisdictions so that they can more effectively and efficiently assess the impact of particular changes in their own operations.

One particular set of issues and potential program changes that should be evaluated through this process are those relating to officially promulgated standards such as those presented by the National Commission(s) on Criminal Justice Standards and Goals. Any jurisdiction with such a model could make a rapid assessment of the impact on that jurisdiction of adopting any one or combination of such standards and goals expressible as model parameters.

The basic data for such models is being readily and increasingly made available through the Offender Based Transaction Statistics program which provides flow information about individual cases as they are processed through the criminal justice system. Aggregation across individuals provides the detailed data on branching ratios. Cost information can be derived fairly directly from budget documents. Workload information is relatively more difficult to obtain but could be obtained from various jurisdictions and spread more widely. Thus, a clear national opportunity exists to foster more systematic planning, to make effective use of the developing OBTS data for planning, to disseminate evaluative information in a useful form, and to provide a national clearinghouse of compatible system parameters.

C. Measurement of Intervention Rates

Clearly, measuring public protection by the criminal justice system as in the previous two projects, is viewed as of immediate and direct concern to criminal justice agencies. However, the problems of individual freedom, particularly in an era of increasing concern with civil liberties, is as important to consider in measuring criminal justice system activity.

From this point of view there is a real need to examine the increasing intervention which the criminal justice system seems to have taken in the daily activities of the American public. This implies a better examination of the trends of arrest rates particularly for the different serious types of crimes (including violent and property crimes) and for the victimless crimes (such as drug abuse and consensual sex offenses). By examining these different types of criminal activity or deviance,it is then possible to define life-time arrest probabilities for these and to compare them to previous periods in time. Earlier examinations [Christensen, Belkin et al] point out that for a male who is born now in the United States, the probability of being arrested during his lifetime for offenses more serious than a traffic offense is .6.[1] The comparison of this probability with previous time periods, particularly within the different crime types, is an important area for further examination. It is important to say that such an examination is needed because of the fundamental inter-relationship of the police resources and arrest rates mentioned earlier. That is, both have been increasing, although it is not clear to what extent each is the cause of the other. Thus, although violent

[1] Such a study has been done for all crime types. See Belkin, Blumstein and Glass "Recidivism as a Feedback Process: An Analytical Model and Empirical Validation", Journal of Criminal Justice, Vol. 1, #1, March, 1973.

crime may be different, there is a real question as to whether arrest and reporting rates,conviction rates,and incarceration rates for other types of crimes have too often been a function of the agency resources rather than public values. Thus, an examination needs to be carried out in depth, particularly for violent crime types of the lifetime arrest probabilities.

The increasing resources which have been applied to all parts of the criminal justice system must also come under increased scrutiny, partly because of their implications for degree of intervention to keep those resources occupied and to identify changing intervention threshholds, and to plan future resource allocation.

Some initial attempts by economists to relate system activities and resources have dealt only with the institutional response not with the actual crime rates. The extensive and valuable victimization information now available provides an estimate of unreported crime. There is a question of how resource allocations have related to this and how they may relate to it in the future, i.e., if only thirty percent of crime is being reported, there is a large supply that could be handled more formally at some time in the future.

The exponential growth of the resources of CJS agencies must be examined, both in relation to the crime problem in the past and to its projection in the future. The increasing size of these agencies in recent years is likely to be attributable to the expanding crime problem, and may not even have kept pace with the growth in serious crime. As one projects a decrease in crime in the future, largely as a result of the aging of the population, either the recources will be reduced, they will go idle, or they will lower the threshhold for intervention. Which of these happens is an extremely important public issue, and indicators can and should be developed to identify the directions as soon as they become visible.

Clearly, the intervention probabilities are quite different for different socio-economic and demographic groups. For certain groups, such as those with substantial criminal activity, this is to be expected and in the interest of public protection, probably to be encouraged. However, for ethnic minorities or the disadvantaged, this should be examined in order to identify any inappropriate intervention or formal processing. For example, the comparison of offenses and subsequent dispositions across population groups would reflect how different groups are handled for similar offenses. The JUSSIM branching ratios for each group would provide such information.

On the other hand, there needs to be more complete examination of the intervention suffered by those persons who serve the criminal justice system as witnesses or jurors. Some indicators should be developed here to measure:

1. the absolute numbers of such participants;
2. the cost to the CJS as well as the cost to the participants, especially in terms of their opportunity cost of income foregone;
3. identification of other factors in the disruption of their normal activities;
4. indicators of the inefficiency of the CJS in using them.

From these basic statistics an exploration could be made of alternative means for their participation in the system or methods of reimbursement that are more equitable and that would encourage more efficient use of them.

The question of opportunity cost must also be raised for those who become clients of the criminal justice system, particularly those who are eventually found not guilty. The substantial opportunity cost to a person charged and tried but acquitted needs to be examined over different offense types and

population groups. The expected earnings denied to those incarcerated must also be examined, taking account of their employability. Not only must expected earnings be examined, but also factors such as direct losses to families and longer term costs of future handicaps resulting from the criminal record.

D. Measurement of Conflict Resolution Effectiveness

Measurement of the conflict resolution role of the criminal justice system is difficult, since this function is normally handled in an informal way. In these cases, there is a real and justified reluctance of the criminal justice agencies or social agencies to report or make known formally individual offender-or offense-based statistics (often because of the labelling implications of such statistics).

However, if such conflict resolution activities are successful, then there should be a real attempt (in line with the rational allocation of resources by government) to try to measure the serious or public crimes aborted by these activities. This could use as a base measurement control groups for whom there was no intervention, or have not been affected by such activities.

The qualitative measurement on successful and formal diversion of offenders may be obtained on an aggregate level from social agencies themselves (directly or indirectly concerned with justice systems) compared to similar groups handled formally. This has implications for data structure of criminal justice system agencies and other social agencies involved in such activities and raises questions of:

1. How to keep track of offenders and interventions when informality is a prime requirement;

2. The question of follow-up of the diverted offenders;

3. The question of follow-up of the incidents in which there has been intervention as well as to those where there has been no intervention.

This, then, has real implications for measurement of the effectiveness of different types of information as well as providing indicators to the public on the evaluation of different types of intervention.

For this reason, an in-depth study of data and information structures for such agencies within the above constraints and the development of relevant indicators could be pursued as a possible project.

E. Measurement of Service Effectiveness

The service function of the criminal justice system, especially of the police, has been ignored in the past both because of the image of the police as a law enforcement body and because this was not viewed as the primary and most important role of such agencies.

However, recent studies have shown that particularly the police, have dedicated up to ninety percent of their time in such service roles. For this reason, there is clearly a need to first identify these activities and develop a relevant typology of activities. From this it would be possible to examine the time allocation which such agencies make to different service functions and to evaluate this division for better allocations in the future. The question of future allocations raises the problem of developing relevant criteria for the involvement of police and other CJS agencies in such service roles. Having developed a set of criteria, performance measures could be related to these to determine present service effectiveness and the evaluation of future allocation to this area.

There is then a need to develop in-depth reporting systems, again within the constraints of informality, for reporting these calls and follow-up to them. This can then be compared with either the absence of the activity or other similar activities and their impact.

From this there can then be developed detailed data structures, aggregation, and indicators for service effectiveness, both the input (i.e., the number and types of events), and the output (i.e., follow-up comparison with other possible activities of the same inputs and comparison with similar activities of ther agencies.

F. Public Perceptions of the CJS

Efforts to determine the public's concern for their personal safety and to establish along a variety of dimensions the public's regard for agencies of the CJS has been spotty at best. For example, no attempt has yet been made to establish longitudinal data of public perceptions of the CJS. Little effort has been made either to link these perceptions to objective measures of performance or to determine how these perceptions may be influenced by CJS policy alternatives. Because these perceptions can shape behavior in a number of ways that can significantly influence CJS performance it is important that survey instruments be developed that can be used to acquire baseline data. Since these perceptions, particularly those related to fear of victimization, significantly shape those facets of the public's "reality" for which the CJS is generally held responsible, such perceptions warrant continual monitoring in much the same way that such objective measures as reported crimes are monitored. Like the results of victimization studies, these data should be analyzed to determine the difference if any between geographic areas and demographic groups, and so appropriate sampling schemes must be devised.

In addition, these perception data should be correlated with objective performance measures to determine if indeed these perceptions do reflect objective measures, and to determine the nature of the discrepancies. Correlation of the perception measures with CJS activity indicators will also point to means by which CJS policies can shape these perceptions. For example, fear of victimization may be only poorly related to actual victimization rates, especially for some demographic groups, like the elderly or females. These fears may be at least partially allayed by such measures as increased police visibility or improved street lighting. Thus, controlled experiments could then be designed and executed to determine how police patrolling methods (e.g., foot patrol vs. patrol car) and police allocations (e.g., to such areas as housing projects for the elderly and to hospitals and parking lots in which women are likely to be alone), influence not only such subjective measures as reported feelings of safety, but also objective behavioral indicators of such attitudes, such as transit ridership, patronage of commercial activities, or pedestrian traffic.

Data relating to the public's rating of CJS agencies can also be subjected to analysis. For example, survey instruments should be developed to determine under what circumstances individuals observe police (e.g., in passing patrol cars, in parked vehicles, directing traffic, as interrogators), and how they rate these encounters on a positive-negative scale. The respondents' rating of police in such terms as professionalism, corruptability, and commitment to their jobs, would then be correlated with the kind, frequency, and ratings of these encounters. These data would be collected for both municipal and state police. The analyses of these data would provide some

indication of how police style, demeanor, uniform, and training can be matched to particular police functions so as to enhance the image of the police as effective protectors of personal security and community decorum, and warrant support and respect. Similar measures of citizen attitudes toward the other CJS agencies can be correlated with such behavioral measures as the ratio between reported crimes and victimizations, proportion of prospective jurors seeking excusal from service, rate of witness no-shows, and other measures of support and commitment to the criminal justice system.

A CONCEPTUAL BASIS FOR EFFECTIVENESS MEASUREMENT OF LAW ENFORCEMENT ACTIVITIES

prepared by

Stuart Jay Deutsch
Associate Professor
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia
30332

for

Department of Justice
NILECJ/LEAA
Office of Evaluation
September 1975

Acknowledgments

The author is indebted to Mr. Bruce Brownlee, Ph. D. candidate, Georgia Institute of Technology, for his invaluable assistance in preparation of this text and to Richard L. Linster and Richard Laymon of the NILECJ/LEAA for the opportunity to participate in this effort.

Executive Summary of

"A Conceptual Basis For Effectiveness Measurement of Law Enforcement Activities"

The general framework for measurement of organizational performance that is proposed here does not depend on use of any single performance measure such as effectiveness, efficiency, responsiveness or equity. Instead, the procedures conceptualized here can be used with little modification for measuring other qualities of performance. The majority of previous efforts involving performance measurement for public service has been oriented toward specific applications. Little has been done toward consideration of the larger issues---the development of a theoretical framework for such measurements, a framework that is needed to ensure consistent evaluations and a meaningful learning experience with respect to field experimentation. It is the contention of this author that the underlying motivational issues which have given rise to the exploration of questions pertaining to the evaluation of the performance of the Criminal Justice System would be ill served if resulting efforts reflect the direction of previous efforts but only differing in the magnitude of effort.

The purpose of this text is to document the need of developing the theoretical constructs of a uniform measurement philosophy along with necessary methodological vehicles for its attainment. As such, repeated application in an evaluative framework will yield consistent inferences which in the short horizon will lead to "empirical truths." It is important

to distinguish between results which are empirical, or "empirical truths," or causal. The former best describes a majority of evaluation efforts which are not encompassed in a uniform measurement philosophy, whereas the latter is considered to be realizable in a longer horizon than that described in the text material. However, before causal hypotheses can be tested, there must emerge "empirical truths." The text attempts to demonstrate a logical and systematic approach for determining a methodology or family of methodologies for the measurement of law enforcement effectiveness and to suggest which topic areas should be considered for future development in order to have evaluative processes yield "empirical truths."

The descriptive material couches the entire Criminal Justice System in an organizational format and operates from the following hypothesis...: An organization engages in certain activities dictated by its chosen objectives. Its objectives are chosen in regard to its stated goals and its perception of available technology. Because existing technology and existing activities influence the selection of objectives, conflicts between objectives and activities may arise. As "effectiveness" depends on the extent to which objectives are met, effectiveness measurement is meaningless in the context of conflicting objectives. Considering the behavior of an organization, there are a number of different types of performance measures that can be utilized to describe the stimulus-response behavior. Each measure concerns itself with a different aspect of performance or reports the nature or magnitude of performance in different terms.

The text is written in eight succinct sections:

1. Performance Measurement in Criminal Justice System,
2. Elements of Organizational Behavior,
3. Characterizing Organizational Behavior
4. Developing Effectiveness Measurement Approaches for the CJS,
5. Identification and Classification of CJS Goals,
6. Identification and Classification Measures,
7. Selection, Application and Validation of Approach,
8. Research Needs in Effectiveness Measurement for the CJS

In each section, material is presented to describe fundamental background justification while substantive quantitative, hypothetical examples in the CJS environment are developed for illustration. An extensive bibliography of pertinent reference material is also presented.

It should be noted that a major question concerning performance measurement of law enforcement activities has typically revolved around the problem of choosing the right source of data concerning an organization's performance (e.g., FBI Uniform Crime Report data vs. Victimization data). The text describes various documented arguments on both behalves. However, it is the contention that this issue has obscured the real issue. First a uniform measurement philosophy must be developed before additional time and money is spent refining the "quality" of data which is perhaps already suitable or of inconsistent form for substantive evaluation efforts. It

is expected that such a development would give specificity of the types and quality of data needs.

Throughout the text, therefore, particular emphasis and continuity is achieved via description and need specifications of the measurement process. The overall development strives toward finding:

1) performance measures appropriate to the behavior to be studied
2) measurement strategies appropriate to the performance measure being applied and to the utility of the information supplied through measurement, and
3) specific structures or types of measurement processes that best support the chosen performance measures and measurement strategies while meeting the purpose of the evaluation.

TABLE OF CONTENTS

Section 1. PERFORMANCE MEASUREMENT IN THE CRIMINAL JUSTICE SYSTEM

The popularity of performance measurement in Criminal Justice System (CJS) applications is growing steadily, yet there is no definitive and comprehensive conceptual basis for such measurements. The various efforts have been largely application-oriented, computing different measures of performance chosen largely on intuition. Indeed, the major issue to this point seems to be concern over what types of data to use in analyses, rather than determining how the measurement process ought be designed, or what types of measurement approaches can be developed which can consistently yield representative results at low cost. In fact, the role performance measurement should play in the CJS has become a nebulous issue, chiefly because each effort by an individual researcher or group tends to center on a small part of the measurement process, rather than on establishing a uniform approach to the entire problem.

Current Measurement Issues

The major question concerning performance measurement of law enforcement activities revolves around the problem of choosing the right source of data concerning an organization's performance. The emphasis on choosing between the FBI's Uniform Crime Report data and data from victimization surveys has obscured a real issue - whether or not crime rate data of any type is suitable for evaluating a law enforcement agency. If an adequate basis for performance measurement had been developed before attempting to get meaningful evaluations, the real questions about performance measurement

would have been answered. As this has not been the case, some effort must be made to identify the real issues involved.

UCR Statistics as Measures of Effectiveness

The effectiveness of the CJS in dealing with a seemingly rising crime rate has become a critical national issue. Increasing scrutiny of the CJS, its components, and their programs and policies by the media, Congress, and other organizations has succeeded in identifying many of the salient deficiencies of the CJS. The problems associated with operating the CJS have been widely publicized during a period when recognized indicators of crime levels show both reported crime and possibly actual victimization are on a general increase. It is only natural that the public has associated CJS shortcomings with the rise in volume and seriousness of crime as perceived by the public.

By the process mentioned above, the FBI's Uniform Crime Report (UCR) and other such official crime reports have become de facto measures of CJS effectiveness. The reason for this use of the UCR and other crime-rate type indicators as measures of effectiveness is not difficult to understand. Citizens require some information to use in evaluating their government. Legislators and administrators need simple, low-cost performance indicators. Crime-rate type indicators are readily available and are seemingly simple to interpret. Therein lies the difficulty. The use of crime-rate indices as performance indicators presumes that,

1. crime rate data accurately reflects true victimization, and that
2. changes in CJS effectiveness account for nearly all changes in in victimization.

It can be shown that available research does not support these assumptions.

The Validity of Official Crime Statistics

The first assumption, that official crime rates accurately reflect true victimization, is subject to great controversy among social scientists and other researchers. Several researchers, such as Ostrom [54], suggest that the FBI Crime Index is widely regarded as being extremely unreliable, and that the data maintained by police agencies is not the most useful data for measurement purposes. Ostrom reports that even the President's Commission on Law Enforcement and Administration of Justice considered victimization to be best determined by citizen surveys. Cho [13] acknowledges that the UCR data is deficient, but uses the data anyway as it was the only available data. Tittle [20] uses the UCR data in his study of certainty of arrest and crime rates, but mentions work by Beattie [59] that suggests the UCR data is faulty. In support of the validity of the UCR data, Wellford [9] states that:

1. The UCR is the only major data source currently available.
2. Official data and victimization data, although different in magnitude, correlate quite highly with each other.
3. Alternatives to official data correlate to the independent variables (used in crime models) to the same degree and the same direction as official data.

Berk [50] insists that:

1. It is impossible to judge the absolute quality of data.
2. Data need only be as fine-grained as the perception of those who are using it.
3. All types of measurement have some noise and error.
4. Most current analysis techniques are insensitive to absolute error, since most operate on relative measures.
5. New techniques are available to remove or randomize error.

Shogan [40] offers a complete paper on the validity of official crime statistics. Shogan admits that official crime statistics are not on a

1 to 1 basis with actual crime, but cites several reasons for using official crime data. Shogan supports use of the UCR as a measure of crime because:

1. Official statistics are at least moderately correlated with victimization survey results.
2. The UCR is the only source of data that has been compiled over a 40-year period.

Additionally, Shogan questions the validity of some victimization surveys which rely upon respondents to classify the crime. Shogan cites examples showing that not all known victims of crime report that fact in victimization surveys.

In a paper concerned with developing a planning-oriented measure of crime and delinquency, Narms [39] states the UCR is a useful measure of crime because the UCR

1. has existed for 40 years,
2. is familiar to police and the public, and
3. offers a broad national overview of crime.

On the disadvantage side, the UCR

1. is too highly aggregated,
2. does not discriminate between various levels of victimization, and
3. suffers from different reporting procedures used in different jurisdictions.

To summarize, certain weaknesses have been identified in the accuracy of the FBI's UCR. However, the fact that UCR data has been collected for many years and it has shown to behave in the same way as several other indicators of actual victimization such as victimization surveys support use of UCR data as a good estimate of real crime. The chief problems involved with UCR data are more likely to result from the high degree of aggregation in the statistics and the differences in methods of reporting crime across the nation, rather than in the accuracy of the statistics.

Therefore, it is likely that UCR data would be useful for a number of research and practical applications, and that the UCR does indicate, to at least a moderate extent, the true rate of victimization.

CJS Impact on Crime

The second major assumption made by those who would use official crime statistics as a measure of CJS effectiveness is that changes in CJS effectiveness account for most changes in victimization. This is the weaker of the two assumptions. It is the weaker of the two assumptions because studies have repeatedly shown that law enforcement activities explain very little of the variation in crime levels. Indeed, many researchers find that crime does not seem to be a constant flow of criminal activities that can be "turned-off" by increasing CJS effectiveness. Crime has become more precisely identified as a feedback process by such researchers as Blumstein and Belkin [5], Blumstein and Larson [12], AVI-ITZHAK [9], and Rardin and Gray [7]. In these feedback processes, recidivism is usually considered to be the primary feedback signal.

AVI-ITZHAK [6] constructs a crime control model that computes the probability of recidivism on

1. the seriousness of the previous crime and
2. the number of previous arrests.

To justify this recidivism probability estimator, previous work by Wolfgang and Sellin [60] is cited. It is not surprising that, using the fraction of cases cleared by conviction as a measure of law enforcement effectiveness, AVI-ITZHAK's model found that

1. the crime level is fairly insensitive to increases in effectiveness unless the average incarceration is quite long, and that
2. when effectiveness is low, little reduction in crime level can be achieved by increasing penalties at sentencing.

Other studies have demonstrated the relative lack of impact of law enforcement activities on crime rates. In a study by Tittle [20], it was shown that a deterrent effect can be achieved only after a certain "probability of arrest" has been reached. Logan [19] also found only a moderate negative curvilinear relationship between certainty of imprisonment and crime rate. Wellford [9] used a multiple correlation analysis of crime rate aggregated at the city level using social, economic, and police activity indicators as independent variables to show that police activity and money expenditures for crime control have very little effect on variations in the crime rate. In fact, there was a low correlation between police effort and expenditures and clearance rates.

Jones [17] describes results of a study of the effect of changes in police manpower and funding on crime rates. It was shown that for the 155 cities in the study, year to year budget and manpower changes generally have almost no effect on crime rates. Swimmer [10] conducted a non-linear regression analysis of the relationship between police and crime in 119 cities, from which it was concluded that

1. the data did not support the conclusion that increased police expenditure reduced crime, and
2. unemployment rate and schooling rate were not significant predictors of crime in the model used.

This second conclusion is probably an artifact of the model used, for there is a wealth of literature showing a significant relationship between socio-economic factors and crime.

Socio-Economic Conditions and Crime

In a typical study of social and economic factors and criminal activity, Vatey and Phillips [55] write about a study of 18 and 19 year old males that showed economic opportunity to be such a key factor in generating

youthful crime, that

> "...properly weighted, participation rates [in crime] may be a better measure of economic opportunity than unemployment rates."

Several other studies have also identified the significant role economics plays in determining crime levels. Articles by Becker [26], Harris [23], Ireland [24], Stiglu [61] and Erlich [25] tend to support an economic basis for crime. Most of these papers describe economic models for determining the optimal level of law enforcement. Most consider socio-economic conditions as independent variables and crime as the dependent variable.

For example, Ireland [24] writes that fear of retaliation from victims is a much more significant deterrent to crime than fear of retaliation from government agencies. Ireland states that because of increasing wealth, victims are less willing to defend parcels of their property, so that increasing wealth contributes to crime by lessening deterrence, regardless of government law enforcement efforts. This is the sort of evidence that confirms the suspicion of social scientists that disparity of income and certain social factors are the real controlling factors in crime.

Indeed, Erlich [25] applied a regression analysis to variations in the rate of index crimes across the nation and concluded from his results that the rate of specific felonies is positively related to the estimated relative gains available through illegitimate activity. Erlich found a particularly strong correlation between income differentials and property crimes. Erlich did find that the rate of specific felonies is negatively related to estimates of the costs associated with commission of the felony (i.e. fines, penalties), and feels there may be some potential means of deterrence wrapped up in the result. However, Erlich does show

that evidence exists for dividing offenders into risk avoiders and risk acceptors.' He suggests that this accounts for the fact that a distinct proportion of criminals exhibit irrational behavior. That is, behavior that is not geared to maximizing wealth or financial security.

In a thoughtful paper on crime control management, Repetto [56] also points out that current crime control strategy is to deter, detect, and apprehend. He shows that this strategy is considerably hampered by the irrational nature of criminal behavior. Many current quantitative models of the CJS and criminal activities presume that the criminal is a rational decision-maker. However, Repetto points out that rationality is not the hallmark of a criminal. Many criminals are young, and most all criminals have little formal education. In addition, many serious crimes such as murder, rape, assault, and some drug related crimes are considered to be "crimes of passion," instances where the target of opportunity is attacked without lengthy premeditation and rationalization. Repetto even shows the existence of irrational behavior among crimes that are to some extent "planned," such as robberies, burglaries, and thefts, by citing studies that show criminals engaged in these activities often block out fear of consequences entirely, or considerably underestimate the probability of arrest.

To summarize briefly, it is unsafe to assume that effective law enforcement is primary determinant of victimization rates for two basic reasons:

1. Law enforcement activities do not generally seem to account for changes in crime rates.

2. Social and economic indicators are strongly related to crime rates, and therefore seem to identify causes of crime.

Criminal activities and the operation of the CJS can be identified as a single feedback process with recidivism as the feedback signal. Models of this type have been used to show that crime is fairly insensitive to law enforcement effectiveness increases. Police manpower and budget increases have been shown to have little effect on year to year crime rate changes, and deterrence effects related to certainty of arrest, which is a measure of effectiveness in itself, have been shown to take effect only after a critical certainty has been reached.

Social and economic indicators, particularly employment and income disparity have been shown to explain variation in crime rates by several researchers. In addition, studies have identified a distinct type of criminal behavior, risk acceptance behavior, which defies the popular deterrence-oriented crime control strategies. Criminals exhibiting this irrational behavior are probably not deterred by increases in law enforcement effectiveness, whatever we define that effectiveness to be.

The Need for a Uniform Measurement Philosophy

The majority of effort done involving performance measurement for public services has been oriented toward specific applications. Little has been done toward consideration of the larger issue, that is, the development of a theoretical framework for such measurements, even though the need has been widely recognized by such authors as Poland [32], and Holzer [8]. This paper is directed toward defining some of the larger considerations involved in measurement, and toward identifying some specific areas where further development work is required.

The Need for Measurement

From a managerial standpoint, there are clear advantages to be had from a certain knowledge of the performance level of a law enforcement

agency. In general, there are a number of important purposes toward which performance information can be directed. Specifically, the optimal allocation of tax dollars among competing programs or agencies requires a knowledge of their performance. When possible, resources will be spent for those activities that are deemed effective, and ineffective programs are terminated. Implicit in this optimal allocation is the determination of an optimal level of effort (expenditure) for each program, and for all programs considered together.

Another action taken on the basis of measurement information is the restructuring of public agencies. Taken in terms of the optimal response of a law enforcement agency to variations in crime levels and other social and economic conditions, changing the composition of an agency can be best effected when there is reasonably accurate performance information available.

Purpose of the Paper

Having cost sufficient suspicion upon the use of crime rate statistics as measures of law enforcement effectiveness, it is logical to determine what means can be used to evaluate the performance of the CJS, its member organizations, and their programs. It is the purpose of this paper to demonstrate a logical and systematic approach for determining a methodology or family of methodologies for measurement of law enforcement effectiveness, and to suggest which topic areas should be considered for further development work.

In particular, it will be shown that an organization engages in certain activities dictated by its chosen objectives. Its objectives are chosen in regard to its stated goals and its perception of available technology. Because existing technology and existing activities influence the selection

of objectives, conflicts between objectives and activities may arise. As effectiveness depends on the extent to which objectives are met, effectiveness measurement is meaningless in the context of conflicting objectives. Considering the behavior of an organization, there are a number of different types of performance measures that can be utilized to describe the stimulus-response behavior. Each measure concerns itself with a different aspect of performance, or reports the nature or magnitude of performance in different terms.

To evaluate an organization, a general outline for selecting a measurement approach is proposed. The outline describes what information about the organization must be known, and how the measure, measurement strategy, and measurement process define a general measurement approach. A few examples will be given to demonstrate how the general outline for selecting a measurement approach operates. Both the CJS and component organization perspectives are discussed, and the concept of an effectiveness function for an organization and for the CJS is to be put forth. In addition, the difference between the "top-down" and "bottom-up" philosophies of measurement will be emphasized. It will be shown that a careful researcher can develop a set of "induced objectives" which account for all the observable effects of an organization's behavior to a greater degree than the formal objectives. Finally, several areas of research that need be explored to develop a secure basis for performance measurement will be presented. These new areas are relatively young. It seems that several may give rise to significant research efforts into the very foundations of the Criminal Justice System itself.

Throughout the paper, effectiveness is the primary aspect of organizational behavior considered. However, other performance measures are not

disregarded. Indeed, the value of using multiple complementary measures is demonstrated, and nearly all the arguments and developments offered apply to these other performance measures as well. The general framework for measurement of organizational performance that is proposed here does not depend on use of any single performance measure such as effectiveness, efficiency, responsiveness, or equity. Instead, the procedures conceptualized here can be used with little modification for measuring other qualities of performance.

Each of the sections to follow serve the purpose of the paper by attempting to develop a logical rationale for performance measurement and to discuss conceptual basis for selecting measurement approaches. Section 2 serves to recognize the basic elements of organizational behavior, and Section 3 discusses some of the general aspects of measuring organizational behavior. Within Section 4, the concepts presented in Sections 2 and 3 are used to generate a process for finding effectiveness measures, measurement strategies, and the measurement process. Collectively, the measure, strategy, and process define a measurement approach.

Sections 5, 6, and 7 suggest how the process for finding measurement approaches can be applied to CJS applications. The applications demonstrated are included not to show the particular application, but to demonstrate the process for finding measurement approaches. Section 8 contains an analysis of current social service measurement efforts. The analysis points out some of the critical deficiencies of the way in which measurement approaches are being selected and applied, and suggests that many of the problems result from the lack of a uniform, coherent rationale for performance measurement in the social services area. Also contained in the section is an outline of the major topic areas that need be investigated in order to build a uniform and coherent measurement rationale.

Section 2. ELEMENTS OF ORGANIZATIONAL BEHAVIOR

A logical approach to examining the performance of an organization which underlies the philosophy of measurement is through behavioral analysis. This type of analysis is simple and its theoretical background is fully developed. Its major advantage is that it only deals with observable effects.

Objective-Oriented Behavior

An organization is characterized by its behavior. Its behavior consists of the observable activities engaged in by its components. At least nominally, the organization seeks to bring about an ideal target condition, a goal, by engaging in certain activities to reach one or more objectives. The attainment or the set of objectives is perceived by the organization to be the best means for producing the target condition, that is, the goal. In certain cases, the activities of an organization do not serve the stated objectives of the organization. Analysis of the effects of the activities of the organization in these cases can yield an induced objective, which can be used to ascertain a more accurate estimate of the value of the services provided by the organization.

Organizational Behavior

An organization is known to its environment by its behavior. Behavior can be defined as "an observable change." For the organization, the observable changes are the activities of the organization, as those are the "observable changes" associated with its existence. For organizations composed of people, equipment, and procedures, the activities of the organization are the activities of its personnel.

Behavior is usually analyzed in terms of stimuli and responses. In other words, a given stimulus evokes a certain response from the organ-

ization. Using simple notation as used by Keller [81], this stimulus-response pair is denoted

$$S \rightarrow R.$$

From a systems standpoint, the stimuli presented to an organization are the inputs to that organization, and the response of the organization are its outputs. It can be seen that the nature and magnitude of the response to a given stimulus are determined by the goals and objectives of the organization. It is the observable response of an organization which is considered its behavior.

Goals and Objectives

Having described goals and objectives as determinants of an organization's behavior, it is necessary to discuss goals and objectives. A goal of an organization is best defined as the formal reason for the existence of an organization. A goal is the overall mission of an organization. It is a target condition to be achieved, usually by reaching a set of objectives. Whereas a goal is normally an ideal state or accomplishment specified without reference to how it is to be achieved, objectives are the operational expression of that ideal state. The objective is an event that the occurrence of which is perceived to contribute to the overall target condition, the goal. The activities of an organization are its observable behavior. These activities are directed toward accomplishing one or more objectives.

As an illustration, consider Figure 1, which shows the conceptual relationship between goals, objectives, and activities. The conceptualization is based on work by Hirsch [16], and uses an urban police department as a model organization.

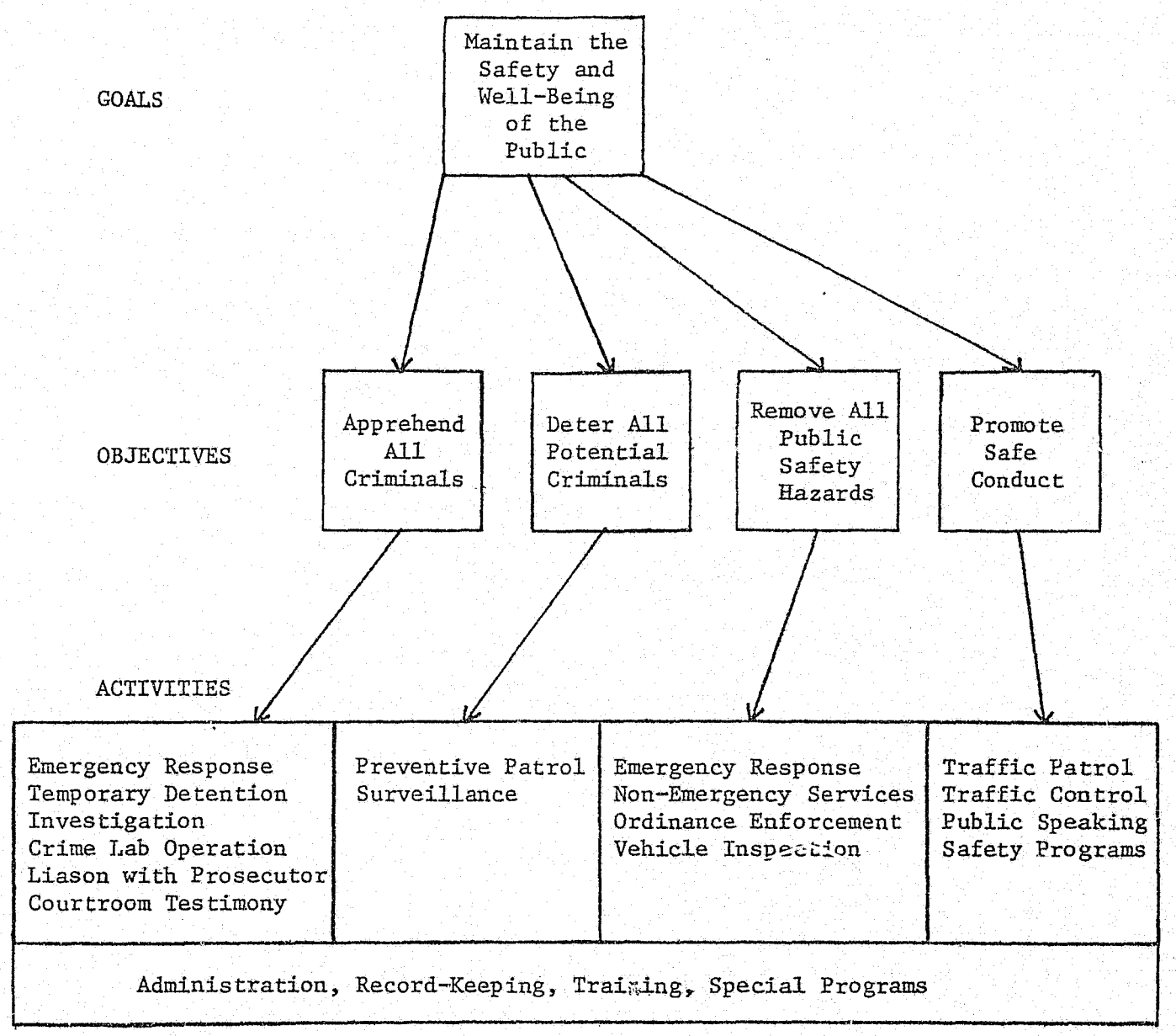


Figure 1
The Relationship Between
Goals, Objectives, and Activities
of an Urban Police Department

The figure demonstrates that several activities may serve multiple objectives. This causes difficulty in properly assigning expended resources to observable outcomes. The challenge is to separate the resources expended during some activity and fairly charge them against each of the objectives served by the activity to the extent which that each objective is served by the activity. This question has been rather thoroughly described by Ostrom [35]. It is shown that there are three factors involved in determining the degree of difficulty experienced in allocating the responsibility for expended resources to various consumers. (For the purposes of this paper, it is suggested that this allocation to consumers is identical to allocating responsibility for resources expended by a group of activities to the objectives they serve.)

The three factors mentioned above are;

1) exclusion,
2) choice, and
3) divisibility.

Exclusion refers to the costs of excluding individual consumers from the benefits of an activity. For the situation at hand, exclusion costs are the costs for preventing an activity from serving an objective. Obviously, if it is very expensive or difficult to prevent any single activity from serving any objective, it will be difficult to decide how to allocate the responsibility for resources expended in any activity to the proper objective. In a consumer application, choice refers to the cost to the consumer of enforcing the decision whether to consume or not. There is no direct significance for this in the situation at hand. Hence the term will not be used in this context. Divisibility, in the consumer sense,

refers to the extent to which benefits can be "used up" by any single consumer. In the activity-objective application, divisibility refers to the extent to which the effects of one activity are "used up" by a single objective. Higher divisibility leads to easier allocation of expended resources. For example, the police patrol provides a number of fairly indivisible benefits, serving a number of objectives. The operation of a bicycle safety program, however, would provide benefits mostly to a public safety objective.

Technology and the Perception of Objectives

In many cases, it can be shown that the goal of an organization is completely or partially determined by factors external to the organization. For example, a corporation reports to its stockholders on its goal, profit-making. At least theoretically, society enfranchises a law enforcement agency with a single goal, public safety, as its reason for existence. Even the objectives developed by those who must enforce the law are subject to public scrutiny, particularly when stated objectives lead to activities not in accordance with the public perception of the role of law enforcement agencies.

However, to a great extent the objectives chosen as appropriate for reaching a goal are determined by the existing technology of the organization and historical experiences of the organization. For example, it might be said that prior to the late 1960's, major U.S. corporations did not place great importance on pollution control. This was because

1. there was no historical precedent within the organization, and
2. the technology for pollution control was not available.

Therefore, those at the operational and policy-making levels within the corporations did not perceive pollution control as an objective. Today, major corporations pursue pollution control because legislation has

served in the place of historical precedent, and because the technology of pollution control has emerged as a unified technology available to the corporations.

This goes to demonstrate that objectives are the organization's own perception of the best means for attaining its goals, and that these perceptions do not necessarily lead to the best objectives. The lesson to be learned from this is simple. Objective-oriented measures of performance will always be limited in their meaning by the technology of the organization. For example, consider the objectives set by a public health organization prior to the introduction of polio vaccines. These objectives might reflect a policy of containing the disease. Any leveling of the rate of spread of the diseases might then be considered to be an indication of effective performance. However, after the advent of a new technology, that is, polio vaccine, the objectives are changed to reflect a new policy of eradication of the disease. It can be seen that although the organization has maintained a uniform level of effectiveness based on its objectives at any given time, there is a real difference between its effectiveness prior to the introduction of new technology and afterward. This demonstrates the limitation technology places on the meaning of a performance measure.

The Induced Objective

Given that an organization chooses objectives that seem to best serve its goals, it is logical to assume that the organization also engages in those activities that the organization perceives as the best means for reaching its objectives. Again the organization's restricted perception of its role is brought to bear, so that there is reason to suspect not all activities serve organizational objectives. Therefore,

it may be useful to carefully examine organizational activities, their intended impact, and any secondary effects of the activities to determine what actual objective is being served by each activity.

In accord with the concept of organizational behavior, this examination of activities and impacts would necessarily hinge on identifying observable effects relating to the activities of the organization. It is not necessary for there to be a single common denominator that expresses both organizational behavior and its impacts. There may be different measures which may be quantitative, qualitative, objective, or subjective. For example, a crime prevention program could be described in terms of the labor hours, citizen contacts made, and so forth. The impacts might be specified in terms of police visibility, number of houses secured, and citizen approval of the program. As a second example, consider the prostitution situation found in major urban centers. Following the desires of elected representatives of the urban constituency, police administrators put forward as a formal objective the elimination of prostitution. However, an objective observer reviewing police activities and conditions "on the street" might surmise that the actual objective toward which police were laboring was:

1) reducing the visibility of prostitutes' activities in certain neighborhoods,

2) confining pick-up areas to a small downtown area, and

3) eliminating crimes secondary to prostitution (such as robbery).

This second objective is the "induced objective". When the induced objective is different from the stated objective, one of three causes can usually be identified. A variance is caused when either

a) the stated objective is infeasible,

b) the stated objective can be achieved but the chosen activities do not serve the objective, or

c) the stated objective is served by the activity, but significant secondary effects result.

As presented, the concept of induced objectives is important because, as defined, effectiveness is an objective-oriented measure of validity. If there is a dual set of objectives at play, then those seeking an evaluation of a law enforcement organization must first decide which set of objectives should be utilized in the evaluation. This is a serious theoretical and practical concern. There is no doubt that a case can be made for using either set of objectives in evaluation. If measurements are based on formal objectives, then the results are more likely to reflect the effects of objective-objective and objective-activity conflicts, and also the effects of striving toward an unattainable objective. If measurements are based on induced objectives which are developed from observations of the organization's activities, there will be less likelihood of the effects of objective inconsistencies and infeasible objectives appearing in the results. The question then is choosing between evaluating an organization on the basis of objectives it is supposed to be working toward, or on the basis of the objectives it actually seems to be working toward.

Activity: The Expenditure of Resources

The activities of a law enforcement agency are the observable day-to-day operations, the actual expenditure of resources to achieve an objective. Those activities of a law enforcement agency directly affecting in a useful or beneficial manner individuals or organizations external to the law enforcement agency are considered to be "services". These activities represent observable behavior on the part of the organization, they can to some extent be described and quantified. Activities of an organization

can be quantified in terms of resources expended. This is best understood by considering the basic laws of thermodynamics that demand the expenditure of resources for all observable behavior. This suggests that a useful alternative definition of "activity" is "the expenditure of resources".

"Service" is a more difficult concept to evaluate. First it must be determined that the activity in question is beneficial. Then the second problem faced is in quantifying the value of the service. Hopefully, the service afforded by an activity is of greater net value than the resources expended in providing the service. Given this assumption, the value of resources expended in providing a service becomes the lower bound on the value of the service provided by an activity. There is however, no guarantee that the value of a service provided is of greater value than the resources expended to produce the service. In many cases, the value of the services is not known in economic terms, giving rise to the popular use of the cost/benefit ratio for analyzing the ability of the organization to provide the service in question. The rationale is that the use of the ratio does not require economic interpretation of the value of a service, only some quantification of the magnitude of the service provided, be it a subjective or objective quantification.

Indeed, the entire concept of determining the value of a service is a study unto itself. This paper deals with determining the value of a service in that some types of measurement techniques for police effectiveness may be based on the process of determining the value of services provided by law enforcement activities and then relating the value to a desired level of service.

Service Levels Over Time

The time element can be a significant factor in an organization's response to its environment. First, there is the question of consistency of behavior. Obviously, consistency of performance of an organization is a desirable behavior trait. Any observer of corporate behavior recognizes this when he notes the great efforts taken by most corporations to pay regular dividends. Second, there seems to be somewhat more value in receiving benefits immediately, rather than at some future date. This principle is widely recognized in economics, where the benefits are monetary, but there is little existing work toward use of any type of weighting factors to non-monetary benefits being delivered over a continuum of time. Certainly, however, there is an added value to receiving more immediate benefits, be they monetary or not. For example, which is better, a permanent reduction of crime rates by 50% beginning this year, or a 60% permanent reduction not beginning until next year? This is the nature of the question which might be addressed by this sort of analysis.

Section 3. CHARACTERIZING ORGANIZATIONAL BEHAVIOR

There are several different elements of an organization's behavior that can be evaluated through use of a measure of performance with some measurement strategy. A measurement strategy is a single set of procedures derived from the general measurement process. The appropriateness of a measurement strategy depends on the procedures that compose it, the measure of performance it is applying, and the behavior it evaluates.

Strategies and Measures

A measure or performance is an indicator that expresses the magnitude or direction of an organization's activities, or the degree to which the behavior of the organization is perceived as desirable. Many quantities can be put forth as performance measures. The basis for any measure depends on the basic process,

$$S \rightarrow R,$$

since this process describes the observable behavior of an organization, and because what is not observed is not measured. In this sense, measurement is equivalent to observation.

Measurement strategies are the means by which the quantitative values of performance measures are determined. A measurement strategy is a procedure designed to estimate the value of a performance measure based on observations of behavior of the organization or the effects of that behavior on its enviornment. In short, measurement strategies are methodologies for applying performance measures, which in turn quantify some aspect of behavior or the impacts of behavior.

The Nature of Measures

Examination of the concept of performance measurement as it applies to an organization reveals that the different measures of performance can

be characterized by the specific features of the organization's performance measured and by the terms in which the results of the measurement are expressed. All measures can be classified as either being absolute or relative. The absolute performance measures quantify the organization's performance in terms of the absolute value of the measured feature. For example, the number of arrests made by a police patrol would be an absolute performance indicator. Many absolute measures are of the "volume" variety. Relative measures compare some aspect of the organization's performance to a standard or ideal performance, or to the same aspect of another organization's performance. The key distinction of a relative measure is that it is given in terms of a unit-free ratio.

A second characteristic of a performance measure is in its directness to the performance being measured. Direct measures evaluate performance as expressly as possible. For example, a typist's performance might be evaluated in terms of pages typed. Indirect measures are those performance indicators that are separate and distinct from the measured feature of the organization's performance, but which is related to that aspect of performance. Surrogate measures are useful in situations where the element of performance to be evaluated cannot be measured directly because of the difficulty or expense of gathering the necessary information. For example, the protection afforded a neighborhood by a police patrol might be measured in terms of annual patrol miles. The development of surrogate measures depends on locating certain elements of performance or the environments response to that performance that can be shown to be correlated to the aspect of performance to be evaluated. The desirability of developing surrogate measures lies in the need for inexpensive and practical measures of performance. In some cases, surrogate measures can be used to assess intangibles,

such as the effectiveness of a crime deterrence program. Although there is no way to specify exactly how many crimes did not occur because of the crime deterrence program, any decrease in reported crime would be viewed as an indication of the program's success.

A third characteristic of a measure of performance takes rise from the distinction between the behavior of the organization and the response of the organization's environment to that behavior. The behavior of the organization in relation to the stimuli it receives from its environment is modeled as

$$S^{*}_{\text{environment}} \xrightarrow{\underline{\text{org}}} R_{\text{organization}}$$

The behavior of the environment specific to the organization's response can be shown as

$$R_{\text{organization}} \xrightarrow{\underline{\text{org}}} S_{\text{organization}} \xrightarrow{\underline{\text{env}}} R_{\text{environment}}$$

The recognition of the process, ($\xrightarrow{\underline{\text{org}}}$), brings added detail to the meaning of the simple behavioral model, because it accounts for the nature and structure of the organization, its resources, and its technology. Certainly these factors must be recognized in the determination of an organization's responses to stimuli. These factors are discussed in Section 5. They include

1) The goals, objectives, and activities of the organization
2) The relationship to other CJS agencies
3) The technology and environment of the organization
4) The internal structure
5) The internal workflow
6) The internal flow of resources

All these factors are included in the "process" portion of the model represented by

$$\xrightarrow{\underline{\text{org}}}$$

*S environment refers to a stimulus produced by or received from the environment.

CONTINUED

1 OF 5

used between the R organization, (the responses provided by the organization), and the S organization (the stimuli provided to the environment by the organization.)

If the process, (org), by which the organization produces a response to a stimulus is recognized, it can be shown that performances measures fall into three categories. These are:

1) Process-oriented measures
2) Response-oriented measures
3) Impact-oriented measures

Process-oriented measures are likely to be quantifications of some aspect of the internal workings of an organization. Response-oriented measures deal with evaluating the external response of an organization, that is, the activities of the organization that interface directly with the environment. Impact-oriented measures are those measures of the environment's response to the activities or the organization.

A fourth quality of a performance measure is its objectivity. For example, the number of investigations completed by an investigative agency is an objective measure of its level of effort. It might be possible to conduct a citizen-survey to ask whether or not the agency appeared to be in operation at a proper level of effort. The results of the survey would represent a subjective measure of level of effort. It is known that certain measures of performance are meant to evaluate quantitative features of performance, such as number of arrests, man-hours expended, and so on. Other performance measures are quality-oriented, and describe the precision or the level of excellence corresponding to the measured behavior.

Finally, measures of performance can be divided in two other categories: those performance measures which are primarily concerned with the expenditure of resources, and those measures concerned primarily with the achievement

of objectives. The first measure is resource-oriented and the second is objective-oriented. The two types of measures are not necessarily both applicable in the same situation, but when used together, each measure tends to enhance the meaning and significance of the other.

To summarize briefly, performance measures can be typified in at least six ways. Specifically, performance measures may be

1) absolute or relative,
2) direct or indirect,
3) process, response, or impact-oriented,
4) objective or subjective,
5) quantitative or qualitative,
6) resource-oriented or objective-oriented.

Any given measure may embody several of these characteristics. In addition, it is often useful to apply two or more measures together in order to increase the significance of the results obtained.

Multiple Complementary Measures

Multiple complementary measures are performance measures of different orientation that provide a better definition of organizational performance when applied simultaneously than when used individually. Orientation relates to the specific type of measure applied. As has been suggested, there are resource-oriented measures, objective oriented, and so forth. Multiple complementary measures are currently used in many social service measurement efforts such as those reported by Mantel et.al. [78] and Ostrom [30]. Sometimes these measures are considered together as a set of criteria for use in evaluation. These criterion are then tied into a general performance equation that is developed in a way similar to the method by which an overall effectiveness measure is determined, as given in Section 4.

That multiple complementary measures can give a more accurate representation of behavior is not surprising, particularly if the examined behavior of the organization elicits multiple responses from the organization's environment, or if the organization provides multiple responses for each stimulus it receives. Ostrom [14] writes that the difficulty in evaluating police performance is that police provide a multiplicity of services. In behavioral terms this means that police provide multiple responses (outputs) to a given stimulus (input), and that these outputs elicit multiple responses from the environment. This leads to use of multiple measures.

As an example of the use of multiple complementary measures, suppose that a prosecutor's office sets an objective of clearing the 20 most serious cases on its schedule during a one year period. Furthermore, suppose that an expenditure of $3,500 in salaries and supplies is considered the standard cost of such a case. Under these performance criterion, we can analyze two interesting situations, where efficiency is defined as the percentage of excess expenditure to the standard. First, consider the outcome where 18 of the 20 cases were cleared, but an average $7000 was expended in the pursuit of each case. Were effectiveness the only criterion, a 90% rating (18/20) would result, even though only a 0% efficiency $(1 - \frac{7000-3500}{3500})$ rating had been achieved. This shows good effectiveness, low efficiency. Second, the expenditures for the year may have been held to an average of $4000 per case, but with only 10 cases being cleared. Judging by efficiency alone, this would correspond to 86% efficiency $(1 - \frac{4000-3500}{3500})$, but only 50% effectiveness. This clearly demonstrates how misleading conclusions may be drawn if only a single aspect of performance is surveyed. Conversely, it shows that there is more meaningfulness in using multiple complementary measures.

Elements of a Measurement Strategy

A measurement strategy is a policy that defines which data is to be gathered, when it will be gathered, and how much will be collected. The data consists of information about the behavior being evaluated. It is also apparent that performance measure selected partially contributes to the determination of they type of data that is gathered. Conceptually, a particular sampling strategy for collecting data can be represented as a set of points in a three-dimensional vector space. For example, consider Figure 2. The figure shown how the strategy of gathering a single type of data at regular intervals, but with every second collection being only a small "check-up" on the larger quantities of data gathered at the end of every other time epach.

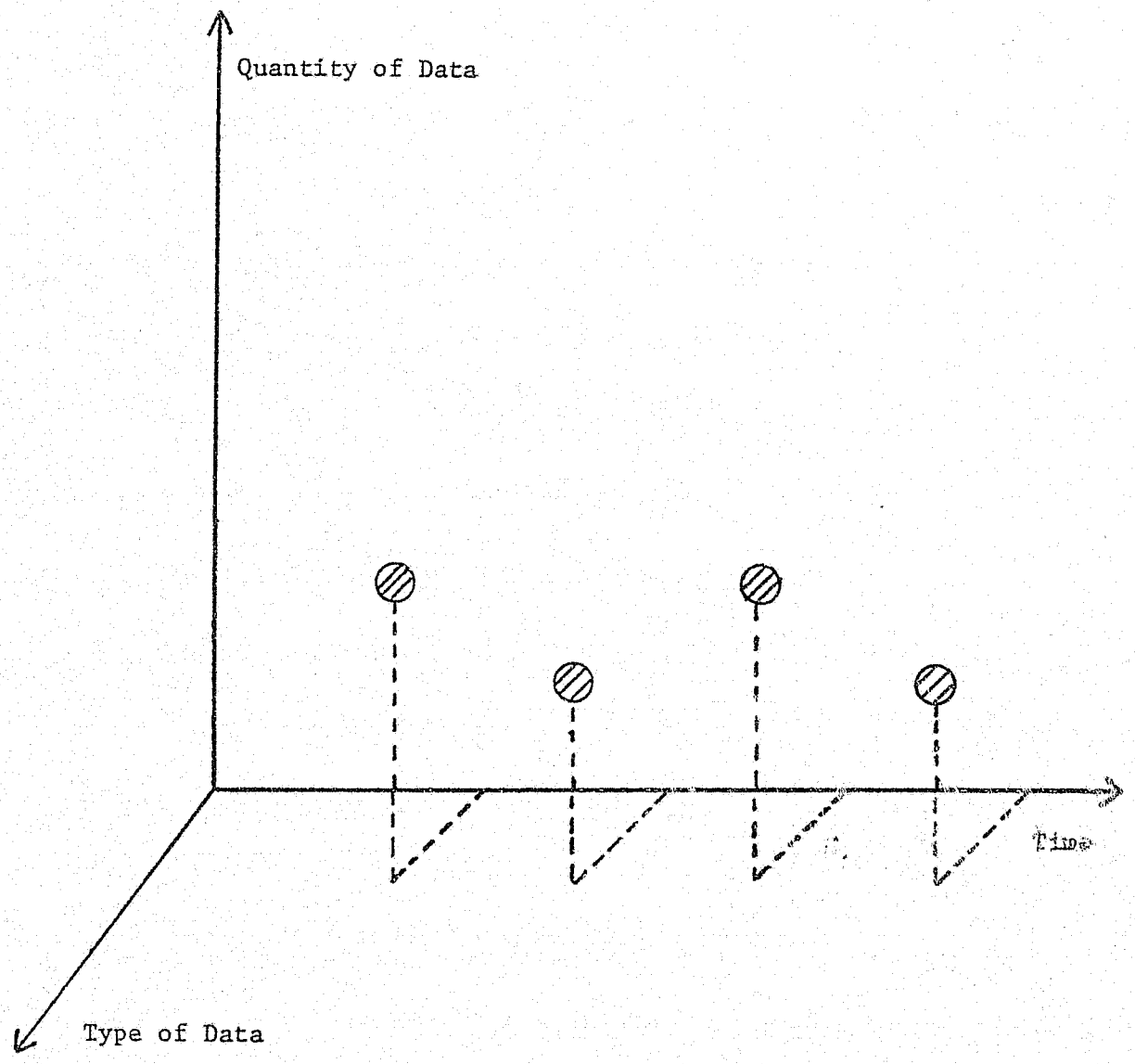


Figure 2
Graphic Representation of
a Measurement Strategy

It is important to be able to select an appropriate measurement strategy. Even if a meaningful and reliable performance measure is chosen, and a highly efficient measurement process is designed, failure of the measurement strategy can lead to inconsistent or unreliable results. For example, suppose it is desired to estimate the number of cases being handled at any time by a group of investigators within an investigative agency. Further suppose that the measurement strategy chosen was: select an investigator at random, calculate the average case load handled by the investigator over the past year, and multiply the figure by the number of investigators in the group. The results thus obtained are entirely dependent on the individual selected. If the individual worked on a great number of trivial cases, the measurement would be biased toward the high side. If the investigator handled only a few major cases annually, the measurement would be biased toward the low side. Obviously, as much care must be taken in selection of a measurement strategy as in choosing the measure of performance.

Determination of the measurement strategy is a challenging problem when the organization is a component organization of the CJS. The single element of the measurement strategy that has thus far received the greatest attention is the type of data. As has been mentioned, there is great controversey as to whether UCR data is the best indicator of criminal activity or victimization surveys are best. This type of question seems to have distracted researchers from developing better understanding of the other two elements within the measurement strategy, when the data will be gathered, and how much data will be gathered.

There is good cause for attending to the logistics of data gathering rather than concentrating on the single question of choosing a source of

data. There are currently be developed massive data collection and storage centers for handling criminal data. Most of these programs under development are taxpayer-funded. Logically, it is reasonable to expect government agencies to attempt to provide maximum services at minimum cost. In this context, it is apalling that very little consideration has been given to development of measurement strategies that do not require all possible data to be gathered. The theory of sampling is well developed and is currently applied in many areas by private industry and certain government agencies.

From reports of current research, it also appears that costs involved in data gathering are non-existent. Very little attention has been given to performance measurement within the larger context of an ongoing periodic audit of government performance, particularly with regard to costs. Most reported efforts deal only with selecting or applying a performance measure that is chosen without regard to the way in which it will be used in the long run. Typically, performance measures turn into quotas, and evaluative programs are so poorly designed as to be seen as obstacles by many involved in the CJS such as Conrad [53]. To summarize, there is a dearth of research into determining the quantity of data that should be gathered, and into choosing the best time to take data and the best data interval. Current work neglects the long run logistics of performance measurement. However, there are a few researchers who have attended to the other elements of the measurement process, usually because they seek to apply a new measurement approach that requires it.

One researcher, Ross [18], based a study on the use of an interrupted time series design, and produced rather significant results concerning the effects of a social reform. The work was productive because more

emphasis was placed on the realization that the sampling interval and the quantity sampled must be considered along with the type of data gathered. This would not have been possible without realizing the relation between measurement and sampling.

Measurement vs. Sampling

Measurement is the process by which dimensions are determined. Sampling is the process by which individual elements of a population of elements are examined. Sampling techniques are used for making an estimate on some feature of the population. When data is collected concerning the behavior of an organization such as a law enforcement agency, it is usually produced by compiling information on the agencies' behavior during a certain time epoch called the study period. In the sense that not all information about the behavior studied is gathered, that is, only data acquired during the study period is available, the information acquired constitutes a sample. This suggests that the process of data collection is really sampling. Because of this, the measurements obtained by applying a measure of performance with a measurement strategy are really estimates obtained by sampling. Knowing this, full advantage can be taken of statistical theory when developing a measurement strategy.

The Measurement Process

The means by which a particular measurement strategy is executed is but a single phase of the complete measurement process. The measurement process is a set of procedures and flows of information that describe the interaction between the evaluators, the evaluated organization, the organization's environment, and the measurements and measurement strategies used. The measurement process for organizational behavior consists of

1) identification and classification of the behavior to be evaluated,
2) analysis of the organization and its environment,
3) selection of a measure and measurement strategy,
4) execution of the strategy,
5) validation and analysis of the results, and
6) improvement of the measure and/or measurement strategy.

The measurement process is graphically depicted in Figure 3.

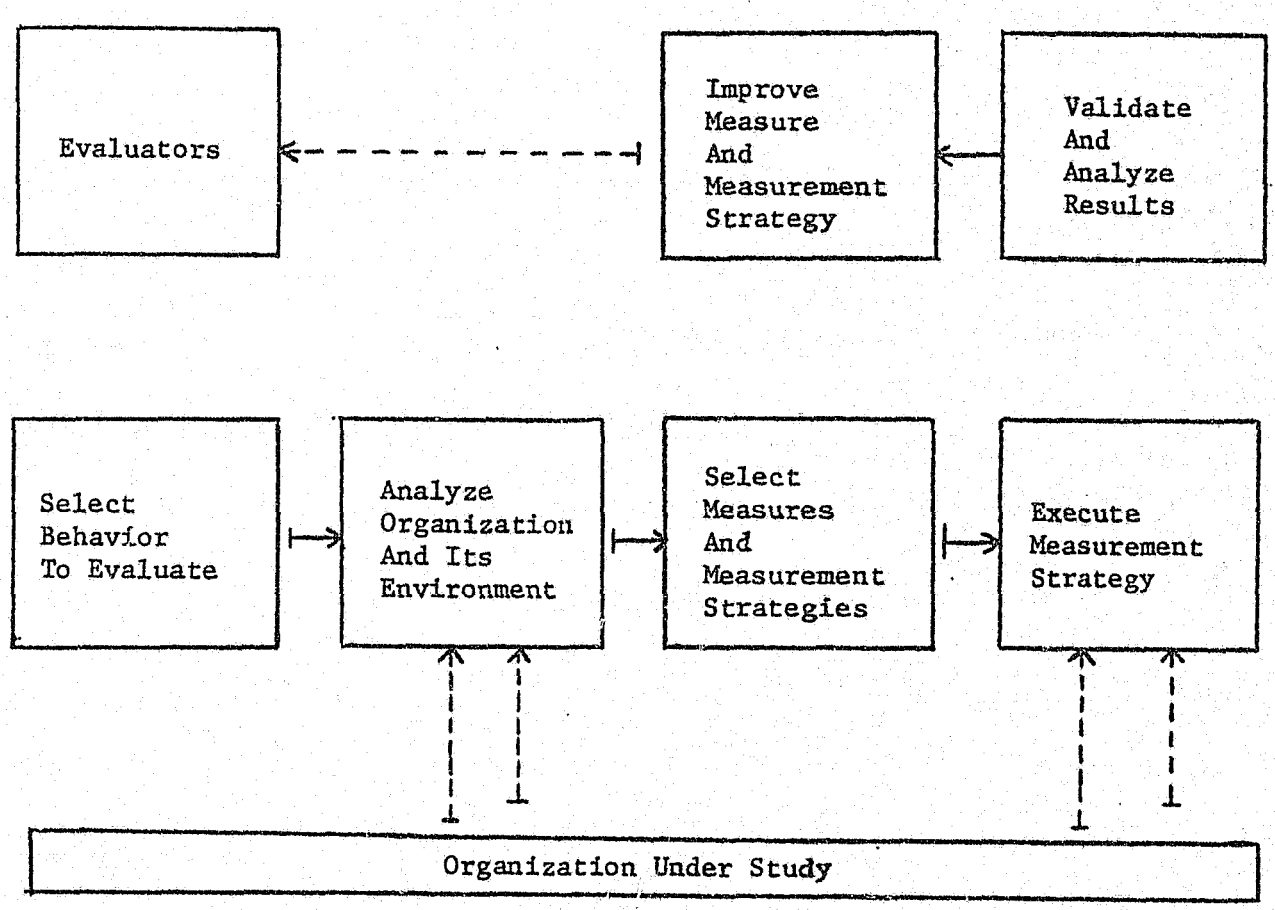


Figure 3.
The Measurement Process

The diagram shows a highly compressed version of the true state of affairs. For example, the information flows shown leaving the organization could be divided into two components, one representing flow of information about the structure of the organization, and the other component representing flow of information about the behavior of the organization. In addition, the entire process of selecting measures and measurement strategies, the question this paper discusses, is a major process in itself.

The measurement process has been studied rather incompletely as it applies to evaluating the CJS. Instead, much effort has been expended on considering only single questions within one of the steps of the measurement process. Typically, reported works dwell on describing or developing sources of data. Although this is important, source of data is only one characteristic of the overall measurement process. For example, a paper by Hirsch and Riccio [16] proposes a variety of separate measures of police effectiveness and efficiency, but nothing is said about

1) when and how the data is to be gathered,
2) what it will cost to gather the data,
3) how the measures will be updated,
4) who will evaluate the measurement results.

In addition, the measures proposed are given in certain knowledge that many of the quantities involved relate to behavior police cannot control; the measures were chosen through intuition rather than observation. No attention at all is given to the more important considerations of determining how the overall measurement process will fit in with police activities. This well-meaning effort, however, is typical of all first attempts at applying the relatively new management sciences to public service applications. There is usually a considerable time lag between the first applications of management science techniques and the serious, well conceived efforts required for the long run. Other work, such as that done by Larson [2],

Avi-Itzhak and Shinnar [6], Hirsch [16], Zacker [44], Reppetto [56], and others goes as far as to thrust forward particular "measures of effectiveness" and cite applications or models developed on the basis of these arbitrarily selected measures. Yet each work has failed to recognize the importance of developing and testing a complete measurement process.

Without a completely developed and tested measurement process, it is doubtful that consistent, accurate, and useful information about the performance of an organization can be obtained. Consistent results can only be produced when a careful attempt is made to define data gathering and analysis activities. There must be a systematic approach to continuing measurements, rather than a single massive effort which may be completely flawed by data or poor methods. As technological, social and economic changes develop, an effort must be made to adapt the measurement process to the new conditions so that consistency of results is maintained. To cite a previous example, the introduction of a polio vaccine for use by a health organization would so drastically change the data that the measurement process would need adapt itself to the new technolgoy by changing the objectives and data used in evaluation.

Accurate data is required for accurate results. Unusual effects not under control of the agency evaluated must also be removed to a degree determined within the measurement process itself. Also, data collection methods must be well-defined to avoid introducing artifacts of the collection method into the data. Finally, useful results cannot be consistently obtained unless there is an ongoing review within the measurement process to determine which information is relevant. Data gathering efforts, just as any other operational activities, tend to develop a life of their own, independent of the purpose toward which they were initially directed. This can result in mounds of useless data at enormous cost.

One author, Caputo [21], does describe a model for urban public policy evaluation developed for policy evaluation in Gary, Indiana. The model was based on a systematic application of subjective evaluations. An interesting feature of the model was the use of more than one body of evaluators. Different groups of evaluators worked with the process in order to legitamize the evaluations in the face of political realities.

What should be done by those seeking to evaluate social services such as law enforcement, is to develope the concept of a measurement process for certain measurement strategies using quantitative measures of performance. The overall development would necessarily need to strive toward finding

1) performance measures appropriate to the behavior to be studied,

2) measurement strategies appropriate to the performance measure being applied and to the utility of the information supplied through measurement, and

3) specific structures or types of measurement processes that best support the chosen performance measures and measurement strategies, while meeting the purposes of the evaluation.

This 3-step procedure defines a logical approach to the determination of a measurement approach.

In the development of the measures, measurement strategies, and the measurement process for a particular organization, there are an immense number of considerations that enter into the problem, but these can be dealt with by dividing them into a new categories. In other words, there are a number of factors affecting the nature of the "best" measures, measurement strategies, and overall structure of the measurement process. These factors depend on the organization to be evaluated and its environment.

Some of the factors affecting choice of a measure include;

1) the nature of the activity,
2) the features of the activity to be measured, and
3) the data gathering capability of the evaluators.

The activity may be economic in nature, which would call for a resource-oriented measure. However, if there are certain non-economic features, an objective-oriented measure may be appropriate. Finally, if the evaluators cannot afford direct measurement, surrogate measures may be chosen. Factors involved in choice of a measurement strategy include;

1) the nature of the measure chosen,
2) the nature of the data or observed behavior, and
3) the costs for data gathering.

Some measures require massive amounts of data, other depend on simple questionnaires. In some cases, the observed behavior occurs on a regular schedule, such as a work schedule, while in others the behavioral events occur more or less randomly. Each type of behavior suggests different measurement strategies. Finally, the quantity of data gathered will likely be inversely proportional to measurement costs.

The overall measurement process is designed in accordance with;

1) the chosen measure,
2) the chosen measurement strategy,
3) the purposes of the measurement, and
4) the identity of the evaluators.

Obviously, the process must complement the measure and strategy selected. What value is there to having daily meetings of the evaluators if the measurement strategy allows for data gathering only once a year?

The process need be designed with its purposes in mind. For example, if the measurements are made for internal evaluations of the agency, it could suffice to utilize agency personnel as evaluators. However, if the measurements are part of an audit of the agency with results given to the public, prudence dictates a group of outsiders must participate on the evaluating committee. Conversely, the identity of the evaluators will determine the nature of the process. Managers are interested in more

frequent appraisals than outside auditors. The identity of the evaluators is quite closely associated with the purposes of the evaluation, as is suggested by the similar relationships to the measurement process chosen.

To briefly summarize, there are a number of factors at play in determining a best measure, strategy, and process. The measures, strategy, and process are chosen so as to best accomodate the purposes of the measurement within the restrictions imposed by these factors. The procedure for identifying these factors is discussed in Section 4.

Value of Measurements

When a measurement strategy is used to complement a particular performance measure, an important consideration in developing that strategy is the value of the information provided by operating the chosen performance measure over the sample data. Obviously, it is desirable for the value of the information provided to exceed the costs incurred during the entire measurement process. Determining the value of the information provided by a measurement is a difficult practical problem, because it is difficult to determine how much closer to its objectives a law enforcement organization will operate given that it receives an additional increment of information from a measurement. Despite the difficulty inherent in this determination, it should be attempted; an estimate of the value of measurement information as a function of the type and quantity of measured data should be developed to prevent excessive expenditures and to determine best sample sizes. There are well developed statistical decision methods available for assisting in determining the best quantity of data to gather. There is no excuse for not developing this decision methodology to CJS applications.

SECTION 4. DEVELOPING EFFECTIVENESS MEASUREMENT APPROACHES FOR THE CJS

This section is directed at applying the principles involved in characterizing organizational behavior, as discussed in Section 3, to the problem of developing performance measurement approaches for the CJS. Particular emphasis is placed on developing approaches suitable for effectiveness measurement. The first effort made in this section is to consider several important performance measures that are useful for CJS applications. Thereafter, the problems involved in determining the overall performance of an organization, including recognition of inconsistent objectives, are discussed. The use of functional and empirical models for determination of the effectiveness function is described.

Defining Performance Measures

This paper is primarily concerned with laying the guidelines and ground work for effectiveness measurement in CJS. Other performance measures, however, are not ignored. Indeed, the informational value of applying multiple complementary measures has been clearly demonstrated in Section 3. There have been a number of attempts at applying performance measures to CJS applications; several efforts have already been cited. Not all these efforts are uniformly successful, but the authors cited have recognized some of the more important performance measures. Four of these are efficiency, equity, responsiveness, and most importantly effectiveness.

The 3-part procedure for designing a measurement approach described in Section 3 begins with the selection of a performance measure. The choice being made according to the factors previously mentioned. The selection therefore depends on certain information about the organization gathered according to a process mentioned, later in this section. When the information is gathered and analyzed, the evaluator has a variety of

performance measures to select from. Effectiveness and efficiency are two of the most useful performance indicators.

Effectiveness and Efficiency

Effectiveness can be defined as validity: the degree to which objectives are met. Effectiveness is a relative, objective-oriented performance measure. In terms of stimuli and responses, the degree to which objectives are met can depend not only on impact, which is essentially a measure of the environment's response to the activities of the organization, but to some extent depend on the actual response of the organization to stimuli. In short, depending upon what the objectives have been defined to be, the effectiveness of an organization is determined by its response to stimuli and the environment reaction to that response. For example, a police patrol force may have two objectives:

1. emergency response within two minutes in all parts of the city, and
2. deter all street crime.

The first of the objectives depends almost entirely on the patrol force, so a measure of effectiveness based on that objective would be response-oriented. A measure of effectiveness based on the second objective would be impact-oriented, as the objective calls for a specific behavior from the environment.

This is a critical issue that has been raised in relation to effectiveness measurement. There is a great difference between evaluating an agency's performance in terms of that performance's relation to an expected performance and evaluating the performance of the organization in terms of its environments response to the organization's activities. For example, consider the hypothetical situation where a new patrol program has a total deterrent effect on neighborhood crime. Were we to evaluate the patrol on arrests, there would result a low effectiveness rating. How-

ever, if the changes in crime were considered to be indicative of effectiveness, total effectiveness would be the finding. This clearly demonstrates the difference between the measurement of the organization's behavior and the response of the environment to that behavior, and points out the type of pitfall encountered when using a single measure of effectiveness. Relative to the behavioral notation, given $S_{environmental}$ as the stimulus presented to the organization, the first effectiveness measure is made with respect to some standard set as a function of $S_{environmental}$, say $R^*(S_{environmental})$. This measure of effectiveness would then be

$$R_{organization}/R^*(S_{environmental}).$$

The second effectiveness measure is defined in terms of the response from its environment that the activity of the organization elicits and some function of the stimulus presented to the organization by its environment. This would then be

$$R_{environmental}/f^*(S_{environmental}).$$

As can be seen, the second measure of effectiveness depends on events that are further away from the control of the organization than those on which the first effectiveness measure depends.

When it is difficult to determine whether or not the objectives set for an organization are feasible, the performance of the organization can be compared to that of another organization giving relative effectiveness. When it is difficult to assess the impact of the organization upon its environment in monetary terms, the cost effectiveness index, as discussed by Grant and Ireson [74] and Fisk [37] can sometimes be used as the measure of effectiveness. The index is defined as the ratio of the quantification of some useful consequence to the resources expended. An example of a cost-effectiveness ratio might be the number of prisoners detained

in a prison per dollar expenditure in the prison budget. It can be seen that the cost-effectiveness ratio is a resource-oriented measure of effectiveness. However, in many cases, the cost effectiveness index is not valid for use as a measure of effectiveness. The difficulty lies in the attempt to correlate expenditure of resources with benefits produced. This tends to overly compress available information into a single ratio. Therefore, the single ratio does not always provide the evaluators with a broad enough basis for accurate measurement. This problem is recognized by a few researchers such as Hatry [79]. Further discussion of this problem will follow in this section and in Section 8.

Efficiency is the output of resources from the organization divided by the input of resources to the organization, as normally defined. Efficiency is also a resource-oriented measure of effectiveness. The term efficiency is most frequently applied to such questions as the efficiency of the police patrol, where efficiency is computed as the ratio of net total man hours on the street in actual patrol duties to the total number of man hours assigned to patrol duty. Efficiency can also be applied to other allocation of resources questions, such as calculating the efficiency of manpower allocation as the total number of individuals assigned direct patrol duties divided by the total number of employees.

Responsiveness and Equity

Responsiveness and equity are social indicators of performance. As discussed by Whitaker [49], responsiveness is the degree to which the public is served by the activities of a law enforcement agency or other organization under study. Equity is an indicator of how uniformly the benefits of the law enforcement activities are distributed. Both these indicators serve as useful complements to the performance measures of effect-

iveness and efficiency, and would naturally be considered at any time effectiveness is being measured.

Multidimensional Measures

In order to fully express each of the important aspects of an organization's behavior, it is beneficial to develop a multidemensional measure of performance as discussed in Section 3, different measures that complement each other can be applied to good advantage. Effectiveness, efficiency, responsiveness, and equity measures are typical of these types of complementary measures. These measures can be included in a multidimensional measure of overall performance in order to provide a more complete picture of organizational behavior then could be had through use of a simple cost effectiveness index or other single performance measure.

As an example, consider how a multidimensional measure of performance might be defined for an organization under study. Assume that the organization pursues j objectives, that k measures of efficiency are included, and that a single measure is used for quantifying responsiveness and equity of distribution. One might then characterize the performance of the organization with a multidimensional performance measure P defined as

$$P = (e_1, e_2, \ldots, e_j, n_1, n_2, \ldots, n_k, r, d).$$

In this situation, if absolute measures are not available for responsiveness and equity, subjective ratings can be substituted. The main point is, however, that there is no overriding need to attempt to compress all aspects of organizational performance into a single performance index, and that any multidimensional measure developed can be conveniently represented with simple vector notation.

Determination of Organizational Effectiveness

The overall effectiveness of an organization depends on how well the

activities serve the objectives, and, in turn, the relationship between the objectives. From a practical standpoint, an observer of an organization with given objectives must rely on the effectiveness achieved with regard to each objective and the relationships between objectives when attempting to determine an overall measure of effectiveness. In some cases, however, it may not be desirable to compute a single value for the overall effectiveness of an organization. In these cases a multi-dimensional measure can be utilized. In most situations, the relationship between the effectiveness associated with each objective and overall effectiveness is not an exact one, so that an interval estimate may be appropriate.

The Effectiveness Function

The overall effectiveness of an organization is a function of the degree to which it meets each of its objectives. For an organization with n objectives, an effectiveness vector of the form

$$\bar{e} = (e_1, e_2, \ldots, e_n)$$

can be constructed. The variable e_i represents the degree to which the ith objective is met. The overall organizational effectiveness might then be expressed as

$$E = f(\bar{e}),$$

where f represents the function that transforms the effectiveness vector $\bar{e}$ into a scalar value E. The challenge to the evaluators is in selecting the objectives to be used in the evaluation, and in determining the effectiveness function, f. The question of selecting a set of objectives was discussed in Section 2, where it was suggested that it is possible to use either formal or induced objectives in the evaluation, and that the interpretation of the effectiveness calculated in the evaluation depended

on which set of objectives was used.

Finding the effectiveness function, f, is a separate question. In the trivial case where n = 1, the overall effectiveness E equals the effectiveness of the single objective e_1. However, it is difficult to find single objective organizations, particularly in the CJS. For the more general multi-objective case, there are two alternatives available. The first alternative is to merely use the effectiveness vector e as a multidemensional measure of effectiveness, without any attempt to compute an overall measure of effectiveness. This is the approach suggested by Hatry [79] and others. The justification is that the specification of the effectiveness function is a political decision, not rightfully the task of the evaluators. A second justification is that in some cases, there may be a mixture of economic and non-economic objectives, and that weighting the value of the non-economic objectives may call for attaching some dollar value to the achievement of these objectives. Attempting to express non-economic values in monetary terms for the convenience of the evaluator is a dubious endeavor. This practice is all too common among social research efforts. The weakness of this compression of subjective values into economic terms is that there is no simple tradeoff between money and the satisfaction of a non-economic objective that does not depend on the degree to which other objectives are satisfied. That is, the tradeoff between money and a subjective value depends on the degree to which other objectives are satisfied.

For example, citizens might be willing to say more to improve the appearance of their local police force by purchasing new uniforms and equipment if the police are able to arrest most offenders quickly rather than if the police are able to make few arrests. Therefore, in the exam-

ple, the monetary value of appearance depends on apprehension effectiveness. This shows that dollars and services are both commodities, and that there is no simple value relationship between any single service by itself and dollar expenditures. However, as was suggested at the end of Section 2, there may be some merit in setting upper and lower bounds on the economic value of a service to estimate its monetary value.

The second alternative available for calculating overall effectiveness involves determining or estimating the effectiveness function, f. This approach is more applicable when there is relative homogeneity among the objectives and is applied when the evaluators determine that an overall measure of effectiveness is needed, and desire to express the results of measurements as a scalar value. There are a number of methods available for determining the effectiveness function, but each approach corresponds to a particular model of organizational effectiveness.

Empirical Models of Organizational Effectiveness

The most commonly used model for overall organizational efffectiveness is an additive model using weighting factors to adjust for the difference in value of achieving the various objectives. For this model, a weighting vector $\bar{w}$ is defined such that for an n-objective organization,

$$\bar{w} = (w_1, w_2, \ldots, w_n),$$

where w_i is a positive weighting factor for the ith objective, chosen such that

$$\sum_{i=1}^{n} w_i = 1.$$

In the weighted model, the effectiveness achieved is in regard to that any objective contributes to the overall effectiveness. The overall effectiveness is computed as

$$E = \sum_{i=1}^{n} w_i e_i,$$

or in simpler vector notation,

$$E = \bar{\bar{w}}\ \bar{\bar{e}}'.$$

With this particular model, the weights can be assigned through a number of techniques, several of which will be discussed later in this section. This weighting factor model is a simple linear first order model, that is, there are no effects of interaction between objectives accounted for in the effectiveness function. The advantage of the model is its simplicity. Among its more current applications is the effort by Mantel et. al. [78], which is discussed in Section 8.

A second type of model is the second and higher order model, which allow for interaction between the various effectiveness scores. As an example of this sort of effectiveness model, consider an organization with two objectives, with e_1 representing the degree to which the first objective is satisfied, and e_2 the degree to which the second is satisfied. Again using coefficients serving as weighting factors, the overall effectiveness of the organization is expressed as

$$E = w_1e_1 + w_2e_2 + w_3e_1^2 + w_4e_1e_2 + w_se_2^2,$$

which can be expressed as

$$E = [w_1w_2]\underline{e} + \underline{e}\begin{bmatrix} w_3 & w_{4/2} \\ w_{4/2} & w_5 \end{bmatrix}\underline{e}.$$

This type of model has the advantage of being able to account for interaction between the achievement of different objectives. The disadvantage is not great, simply that the model is somewhat more complicated and has only an empirical basis. That is, there is not necessarily any specific rationale in nature for selecting a particular set of coefficients for use in such a model.

Model Coefficients

The coefficients used in either the simple linear model or the higher order empirical models cannot be determined exactly except in the simplest of cases. What may be done however, is to choose a set of coefficients that produce results that are both intuitively reasonable and consistent over a number of applications. There is a general method for finding a set of reasonable coefficients which relies upon agreement between evaluators on the overall effectiveness of an organization. The general idea is to apply some correlative analysis for relating the achievement of various objectives to the evaluator's estimates of the overall effectiveness of the organization. After an empirical relationship is established, the empirical model is compared to the judgements of the evaluators for purposes of testing and adjustment, and for setting limits on the application of the model.

There are several techniques available for performing the correlative analysis. There are simple linear regression methods and polynomial curve fitting methods, to mention just two. The more interesting aspect of the search for weighting coefficients is a process by which the judgements by the evaluators of the effectiveness of the organization are combined and analyzed. The most promising such technique is known as the Delphi Method, which combines informed opinion, directed questioning, and feedback of responses to produce a convergence of opinion. An application to finding weighting coefficients is described by Mantel et. al. [78].

With an approach using correlative analysis and a method for analyzing and compiling the estimates of organizational effectiveness made by informed evaluators, it is generally possible to develop a model for determining the effectiveness function, f, even if it is decided not to

attempt to estimate overall effectiveness directly, but to bracket it in a confidence interval. This might result from the diversity of opinion of the evaluators in the original estimation of organizational effectiveness performed prior to correlation with the e_i values related to each estimation. For example, a group of evaluators assesses the effectiveness of an organization three times in an 18-month period, each time calculating the individual e_i values and also applying the Delphi Method in a series of iterations which produce an interval estimate of the organization's effectiveness each time. A typical interval estimate might be specified by stating that the overall effectiveness lay between 70% and 80% with a probability of .95. After the 18-month period terminated, there would be an $\bar{e}$ vector for each of the three evaluations, and a corresponding interval estimate of overall effectiveness. After adapting a correlative analysis technique to handle interval inputs rather than point estimates, and applying the analysis to the interval estimates and associated $\bar{e}$ vectors, a $\bar{w}$ vector could be specified. Testing of the $\bar{w}$ vector would essentially involve testing the predictive power of the effectiveness function given by

$$E = f(\bar{e}) = w_1 e_1 + w_2 e_2 + \ldots + w_{n+1} e_1 e_2 + \ldots,$$

that is, comparing the overall effectiveness calculated using the effectiveness function to the estimate produced by the evaluators.

One comment about the nature of the effectiveness function should be made here. The order of the polynomial describing f is an arbitrary integer that is set by the evaluators at a level giving the best fit, that is, the order is chosen such that the predictive power of the function is maximized. There must however, be some restrictions. First, there is no basis in nature for using an effectiveness function of order greater than

the number of objectives, as only n objectives may interact. Secondly, once the effectiveness function is determined, it should be examined to identify any higher order terms that can be dropped without significantly reducing the power of the function to give accurate estimates of organizational effectiveness. Indeed, one would intuitively suspect that a great deal of the weighting will be on the linear and second order terms, showing near additivity of the individual effectiveness ratings with some interaction between pairs of ratings.

Deterministic Models of Organizational Effectiveness

A second general class of models of organizational effectiveness is the deterministic models. This type of model is founded on the assumption that the effectiveness function, f, can be rather precisely specified through observing the organization, developing cause and effect relationships between each of the e_i, and thus subsequently reducing the dimensionality of the $\bar{e}$ vector by eliminating redundant measures of effectiveness. For example, in evaluating a local police force, similar objectives such as,

1. reducing property crimes, and
2. reducing auto thefts

may be shown to be so highly related as to eliminate the need for the second objective in calculating overall effectiveness. Thus an analysis of effectiveness need only center on the first objective.

Deterministic models rely upon the ability to find a single common denominator of each of the selected performance measures, or the ability to compress a group of objectives into a single objective. Recall that in the discussion of the effectiveness function, it was suggested that determining the effectiveness function is a simpler task if there exists relative homogeneity of the objectives. For deterministic models of or-

ganizational effectiveness, absolute homogenity must be shown. That is, there must be a single element common to all objectives or there must be a way to precisely compute all the e_i from a small grouping of precisely determined information about the organization. An an example, consider a small police force in an urban area. Although there may be several activities involved and multiple objectives, it may be possible to apply analyses of the flow of resources, information, and casework to produce a general effectiveness function. These types of analysis are discussed and exemplified in the next sections. In general, the "common denominator" into which the objectives are reduced is the unit of flow, such as the flow of casework or the flow of dollars.

This then represents the general idea of applying functional models as a means of developing the general effectiveness function. Each of the objectives is expressed in flow units or a similar expression, and then the contribution made to overall effectiveness made by achievement of each individual objective is tested by determining how the various flows affect each other. The nice property of this type of model is that it is based on physical transactions, rather than on empirical relationships, so that the relationships determined have, in fact, some basis in nature.

Determination of CJS Effectiveness: Two Methods

The question of determining the overall effectiveness of the CJS is identical to the situation described concerning determining the overall effectiveness of a component organization. This is because, just as the component organization is a multi-objective multi-department organization, the CJS is a multi-objective multi-component entity. As in the component organization, there is the possibility of conflicts between activities and objectives, or between objectives. The ultimate difference is that

even though its component organizations have formal communications and assist one another, there is no real chain of command in the CJS. As such, coordination and optimization of activities is a matter of cooperation rather than authority and responsibility. It is possible that citizens are frustrated by this lack of a nationwide chain of command and the lack of a repository of responsibility, particularly when each subsequent UCR indicates crime is on the general increase.

Regardless of the lack of a chain of command, there are available two basic approaches to determining an overall measure of CJS effectiveness. The first is to relate crime rates or some other observable feature to CJS effectiveness. This would indicate that the CJS is completely effective when reported crime or victimizations fall below a certain "noise" level. The drawback to this approach is this: the approach is the one in current use, and not yielding a clear picture of law enforcement effectiveness. It is not truly indicative of the limitations on what way law enforcement system, operating within the bounds of the Constitution, can do toward reducing crime. The use of the single measure on such a large scale without any major validation is questionable. Besides these arguments, there is the implicit, and perhaps faulty, assumption that police can significantly deter all types of crime because there is a cause-effect relationship at work. On the contrary, as has been mentioned, property crimes appear to be generated by income differentials, and violent crimes against persons occuring off the street are only very slightly deterred by police activities.

A final damnation of the UCR as a measure of CJS effectiveness lies in the fact that UCR statistics as such do not indicate what specific role many police, judicial, and detention organizations play in determin-

ing and contributing to overall effectiveness. The citizen perceives only that crime in his city has risen or declined. There is no information given about the role poor corrections centers on slow adjudication played in an increase of reported crime. This leaves police to accept responsibility for events outside their control. For example, suppose that a police force in an urban area is very effective in making quick arrests, maintains high public visibility, promotes citizen involvement, and makes efficient use of its scarce resources, labor and capital. Assume that the local courts, however, do not sentence offenders to incarceration of any appreciable duration. Under these circumstances, Avi-Itzhak and Shinnar [6] have shown that reported crime rates will be insensitive to changes in the effectiveness of police. Crime rates can then rise without regard to police effectiveness.

A second means for estimating the overall effectiveness of the CJS can be found in the approaches offered for the component organizations. In short, the idea is to develop a conceptual model of the contributions each component organization makes toward the overall system effectiveness. The model may be either empirical or of the functional variety, as discussed previously for component organizations. The flow of resources and the flow of criminals would be logical candidates for quantifying any functional models for estimating the effectiveness relationships between each component organization and the CJS overall. Quantitative modeling offers the specific advantage of being able to pinpoint weak links in the CJS. That is, instead of using a single nationwide measure, all of the operations are considered and analyzed. In addition, quantitative models have usually met popular use in planning, as it is much easier and less costly to manipulate a model than to experiment in the field.

To outline this approach for determining a general quantitative relationship, be they empirically or functionally based, between agency effectiveness and CJS effectiveness, it is suggested that it would be best to,

a) make a quantitative evaluation of the effectiveness* of each component agency,

b) study the relationships between the component organizations to determine the influence of each agency upon the behavior of the CJS, and

c) use the knowledge gained in a) and b) to derive a general relationship between the effectiveness of each component organization and overall CJS effectiveness.

In part c, some mathematical expression such as was discussed previously in this section might be developed from the data to yield the overall effectiveness figure. The major modification is that conflicts of objectives and activities can be accounted for in the expression in order to make the effectiveness estimate realistic. The major effort is to develop the set of relationships mentioned in part b. Several authors such as Hatry [79] have come to recognize the importance of interaction between component organizations.

In summary, it is recommended that an overall measure of CJS effectiveness be developed as a function of the individual effectiveness achieved by each component organization and of the relationships between organizations. It is expected that this will provide the specificity necessary for identifying difficulties within the system, and that it will not be subject to unproven assumptions about the cause-effect nature or crime such as is the practice of selecting some single indicator of overall performance, such as UCR trends.

---

* or the group of measured effectivenesses if it is determined that multiple objectives prevent combining individual measures.

Inconsistency of Objectives and Activities

As has previously been discussed in this section, real organizations can often be found to have either,

1) inconsistent objectives, or

2) activities supportive of some objectives but inconsistent with other objectives.

These concepts are represented in Figure 4 and Figure 5, respectively. Recognizing these inconsistencies is of importance when relating the overall organizational effectiveness to the individual objective-effectivenesses. This inconsistency concept suggests that in some cases, more effort expended in an individual activity to serve a specific objective may reduce the effectiveness associated with another objective. This is a serious and obviously inefficient situation for an organization, and should be identified before developing an overall organizational effectiveness model, lest there be produced spurious results.

Indeed, a distinction must be made between the goals, objectives, and activities (GOA) of a single component organization and any GOA as might be assigned to the CJS itself. When seeking any measure of effectiveness, ther must be stipulated a set of uniform and consistent objectives which serve as logical stepping stones to a given set of goals. The problem lies in determining which set of objectives to utilize. If the set of objectives espoused by a component organization of the CJS is selected to be used in some measure of effectiveness, the resultant measurements can be entirely meaningless if, in fact, the component organization's objectives do not complement CJS objectives. For example, a 1972 report by the New York City - Rand Institute [51] cited work by S. J. Press that involved increasing the number of police officers on patrol in a certain precinct in New York City. It was found that although significant reductions in

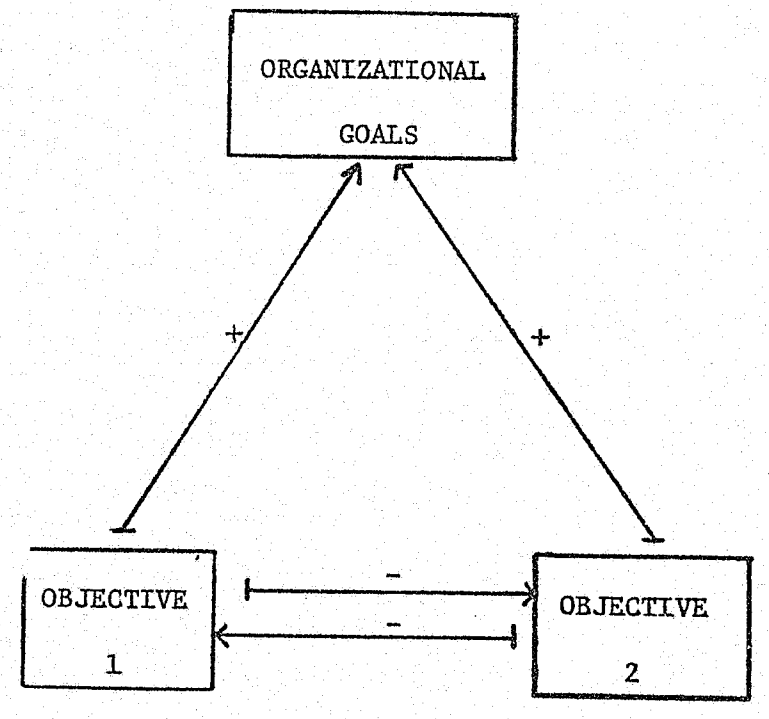


Figure 4

The Interaction Between

Conflicting Objectives

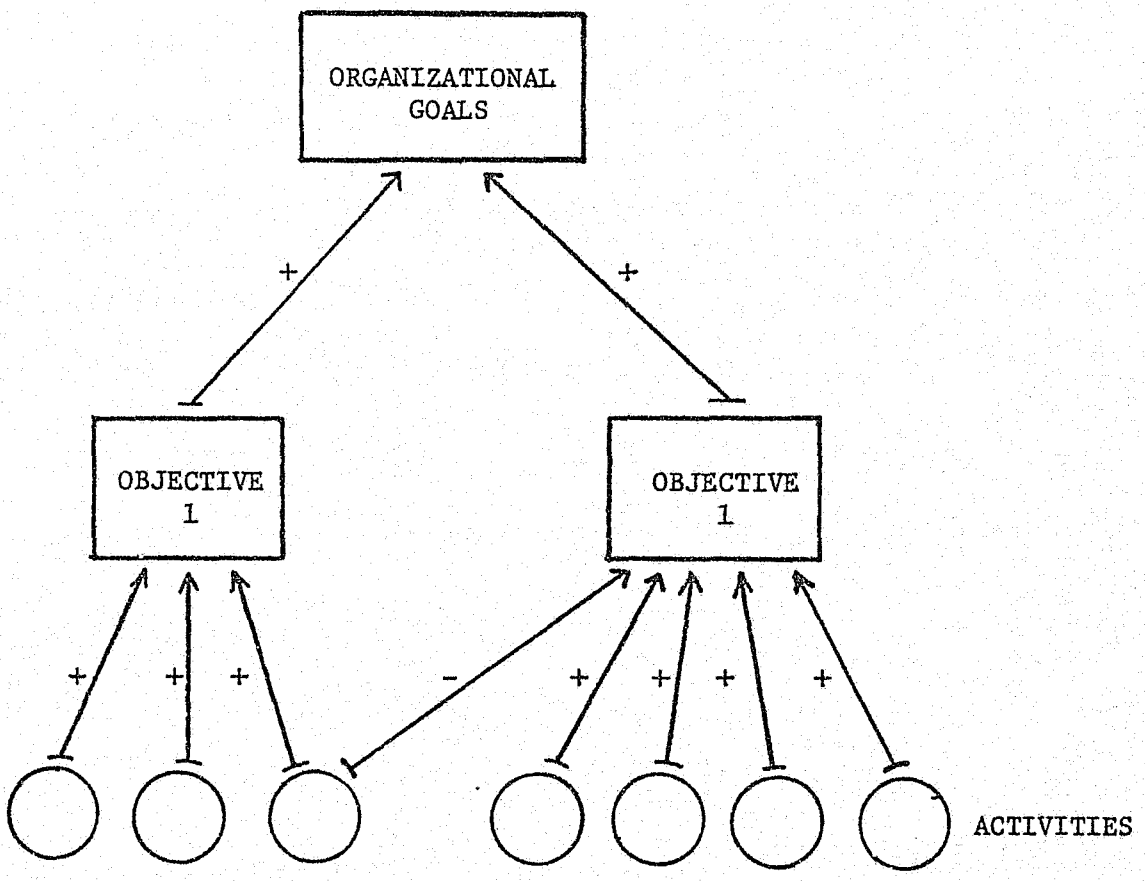


Figure 5
Activity-Objective Inconsistency
for Two Different Objectives

some outdoor crimes within the precinct were noted, the decrease was at least partially negated by an increase in the same type of crime in an adjoining precinct.

This serves to demonstrate the possibility of conflict between objectives and activities, in this case for two precincts. It also points out the need for applying a larger perspective to the measurement of effectiveness. In particular, it shows that for determining the effectiveness of any component organization, we must include some information about the relationship between the objectives of the organization and those of the CJS, and the contribution effectiveness of the organization makes toward the overall effectiveness of the CJS. If the objectives are not selected correctly, it might be possible to have a CJS with zero effectiveness composed of component organizations, each having 100% effectiveness but objectives or activities that conflict with those of other agencies.

Developing the Measurement Approach

The measurement approach is the group of measures, strategies, and the related measurement process chosen for a particular application, as shown in Figure 6. Certain practical requirements for information about the organization and its environment must be met for choice of an appropriate measurement approach. Those requirements dictate the process by which the measurement approach is developed. This simple means that, since different types of behavior call for different measures, and because the nature of the measurement strategy and process selected depend on when and how the organizational behavior to be evaluated can be observed, the elements of the measurement approach necessarily vary with the application. To design a test measurement approach, evaluators must make some

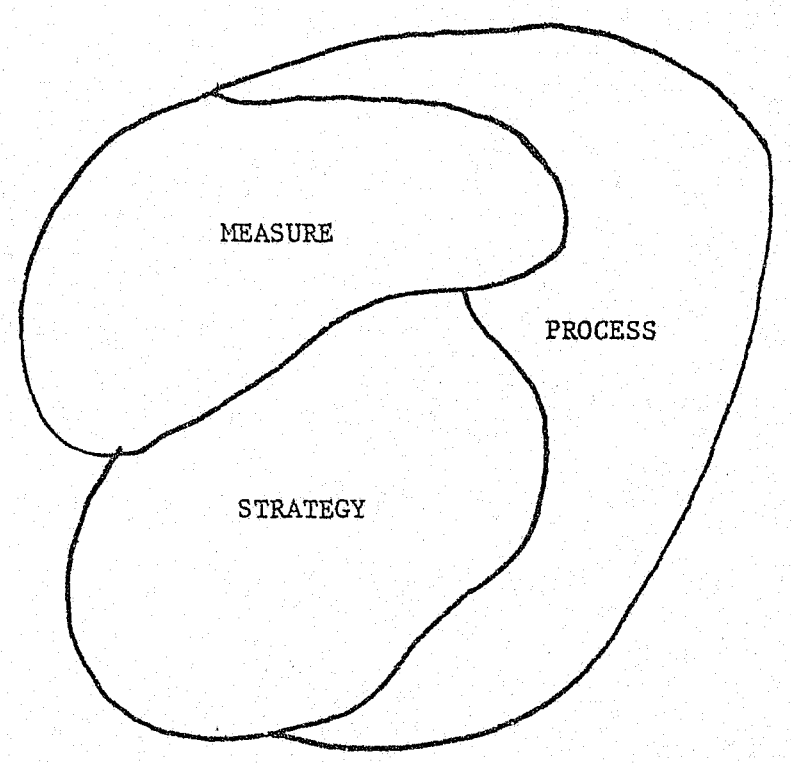


Figure 6

Elements of the

Measurement Approach

preliminary determinations, there must be a preliminary investigation of the organization and its environment. When the general nature of the organization (its behavior, and its environment) is known, evaluators can then proceed to select a measurement approach.

The overall development of the measurement approach, for which a general outline is given in Section 3, calls for the determination of the effectiveness measure, then the measurement strategy, and finally, the measurement process. In order to do this, there must be an orderly method for gathering information about the organization and its environment to assist in determining the appropriateness of various measurement approaches. One such method for analyzing a particular component organization is given as;

1) Identify and classify the goals, objectives, and activities of the organization.

2) Analyze and quantify the structure, flow of information, and flow of resources related to the activities of the organization and characterize the organization's relationship with and the nature of its environment.

Once these steps have been performed, possible measurement approaches are identified and then classified as to their suitable applications. Finally, the measurement approach is selected and applied. There can then be chosen some means for validation of both;

1) The applicability of the measurement approach, and

2) the accuracy of the results obtained.

Without validation, there can be no certainty placed in decisions made based on the measurement results. Validation, therefore is a crucial part of the feedback loop shown in the measurement process as illustrated in Figure 3.

Analysis of Purpose and Practice

A necessary step in the procedure for developing a measurement philosophy is the identification and classification of the goals, objectives, and activities (GOA) of the CJS. Because effectiveness is an objective-oriented measure, it is important that the objective be determined to give the result of any measurement meaning. Because the CJS is operating under a multiplicity of objectives, these must be checked against the goals for consistency. Since the activities are the observable behavior, these must be identified so that the evaluators know what is being evaluated. Classification is important because some measures of effectiveness and some measurement strategies are activity-specific. The classification is meant to apply primarily to activities, which are grouped together based on similarities in objectives and tasks related to the activity.

Analysis of Structure and Operations

The second step toward developing a measure of effectiveness and complementary measurement strategy, both within a measurement process, is the analysis of the organization and its environment with respect to structure, flow of information, and flow of resources. As the CJS is composed of component organizations and the relationships between them, the study most first examine the CJS from a systems perspective, and then proceed to consider the member agencies and their inter-relationships in fine detail. It is during the consideration of the component organizations that the environment of each organization is considered. It is during these analyses that criteria for selection of a measurement approach are implicitly stated.

Identification and Classification of Measurement Approaches

The third step toward developing a measurement philosophy lies in the

identification and classification of measures, strategies and policies that may be appropriate to the task of measuring effectiveness in the CJS. The philosophies are developed from a variety of sources. Classification is performed after the measures, strategies, and processes have been analyzed relative to their applicability to different types of activities. In other words, the measure, strategy, or process is considered in light of its characteristics and an assessment is made as to its suitability for various types of activities.

Given that a complete analysis of the structure and function of a component organization has been completed, and that its environment has been considered, there will be enough information about the activities of the organization to group them into sets of similar activities. This grouping is based on the objectives toward which each activity is directed and the types of tasks involved. Secondly, if the measures, strategies, and processes considered have been classified by the nature of applications for which they are suitable, then a selection of a measurement approach can be executed.

Validation of the results of the measurement can be done in only three ways. First, the results can be analyzed to see if they are intuitively reasonable. Second, the results can be compared to results obtained by different approaches applied to the same activities, or (third) by different approaches applied to similar activities in another organization. Further validation may consider the magnitude of the sampling and analysis costs incurred, and the relative level of effort required to perform the measurements, as per the strategy, and to operate the measurement process.

Section 5. IDENTIFICATION AND CLASSIFICATION OF CJS GOALS, OBJECTIVES, AND ACTIVITIES

A determination of the goals, objectives and activities (GOA) of the CJS must be made before an attempt is made to select a measurement approach. The goals must be analyzed in terms of their consistency, rationality, desirability, and equity. The objectives must be evaluated to determine whether or not they fairly serve the stated goals, that is, to determine whether or not the achievement of the objectives would result in realization of the target condition specified by the goals. Again, the objectives must also be checked for consistency. Finally, the activities engaged in by the individual law enforcement organizations in question need to be thoroughly identified and in some way classified. Only when these steps have been completed, will sufficient information be available to begin to identify potential measures of effectiveness.

This section serves to demonstrate how these steps can be executed. Several possible techniques for identifying and analyzing law enforcement activities are presented. These techniques are functional modeling approaches, as described in Section 4. They are representative of a large body of available methods of analysis which can be applied to both analyzing the activities of an organization and to developing an overall effectiveness function.

The System Perspective

Considering the CJS as a single system of components gives a broad perspective on the national law enforcement situation. The identification and classification of goals, objectives, and activities for the CJS depends on being able to describe the CJS. Once the system is described by

identifying its GOA, individual activities throughout the system can be classified by the nature of the activity and the objective toward which the activity is directed. This classification can be made fairly independently of the goal involved, as long as the objective is consistent with the goal.

The Criminal Justice System

The Criminal Justice System is a multi-component conceptual entity engaged in the production of law enforcement services. The system is described by its component organizations and the relationships between them. It may be possible to depict the CJS as the conceptual entity shown in Figure 7. In the illustration, the CJS is divided into three main levels, Federal, State, and Local, and five main technologies:

1) Police and Investigative
2) Prosecutorial
3) Judicial
4) Detention
5) Planning and Intelligence

There may be certain CJS activities that do not neatly fit into any single classification shown, but these are likely to be special programs involving a number of agencies in joint projects.

One important concept that this illustration does not show is the relationships between the component organizations. These relationships may be specified in terms of the flow of resources, flow of transactions such as crimes or criminals, the flow of information or intelligence, the flow of command, or in several other ways. These relationships do exist and do determine the structure and function of the CJS. These relationships not only determine how the output of each individual agency contributes to the output of the whole CJS, but to some extent determine how each individual agency produces its output.

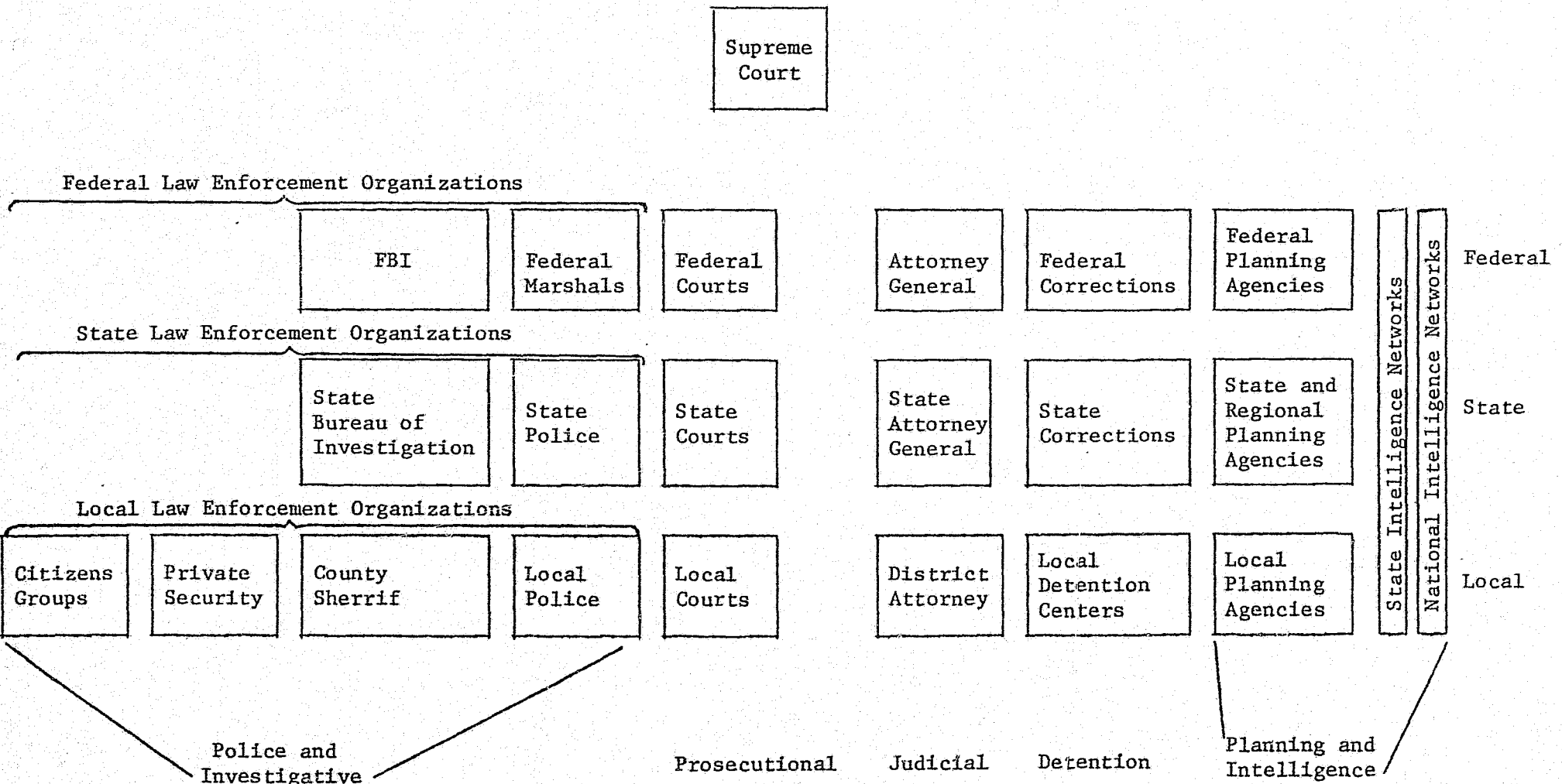


Figure 7
A Simplified Representation of the Criminal Justice System

There is no formal agency known as the CJS, but it is evident that the activities of the component organizations are the activities of the CJS. In fact, there is no formal set of CJS goals, objectives, and activities (GOA). However, it is clear that in attempting to define system-wide GOA, there is likely to be some relation to the GOA of the component organizations. The difficulty lies in the need for uniformity among the GOA's of the CJS. There is no guarantee that the GOA of component organizations are in harmony with one another. Although goals may be similar among the various member institutions, the very fact that each organization struggles to meet its own GOA rather than a single system-wide set of GOA ensures that there will be inconsistencies. This leads to the type of GOA inconsistencies described in Section 4. For example, a penal institution might determine its objective to be minimizing the number of inmates, on the theory that this best meets its chosen goal of rehabilitation. As an activity, it might choose to engage in early releases for certain convicted felons. If these released felons are committing a significant number of crimes in a local jurisdiction, then clearly the penal institution's GOA must conflict with the GOA of the local jurisdictions police agencies, a GOA which are likely to be deterrence-oriented.

In any event, there are two basic ideas to be considered when attempting to speak of the GOA of the CJS. The first notion is that the GOA of the CJS could be considered to be the GOA of the component organizations, conflicts, inconsistencies, and all. This is a "bottom-up" approach. The second notion, a "top-down" concept, is that because some optimal system-wide set of GOA could be constructed through diligent effort, it is possible to speak of the GOA of the CJS as if they were the optimal GOA.

It should be noted that, in the context of a systems behavior, both the composite GOA's of agencies is the CJS, and the "ideal" GOA have dynamic properties. The "ideal" GOA would likely be tailored for "optimal response" to changes perceived in the national environment. The "bottom-up" composite GOA change at the individual agency level in response to changes in environment perceived at that level.

CJS Goals

A determination of the goals of the CJS or a component organization within the CJS must be made prior to choosing an effectiveness measures. Determining the goals of a law enforcement agency within the CJS can be attempted by combining a formal statement of goals of the organization furnished by the organization with a compilation of opinion and evidence from researchers and from private citizens composing the constituency of the organization's jurisdiction. Once a list of goals has been prepared, the goals must be examined for consistency, rationality, desirability, and equity. All this is done to insure that any effectiveness measurements made are meaningful.

An effectiveness measurement cannot be undertaken, and results should not be presented without consideration of consistency of goals of the organization. Consider a hypothetical situation where it is desired to measure the effectiveness of a law enforcement agency, and that the agency has two conflicting goals. Furthermore, let us assume that the goals of the agency are fairly represented by two objectives, which of necessity are not consistent with each other. An objective-oriented performance measure is chosen for evaluating the effectiveness of the agency. The measure attaches equal importance to achieving each of the objectives. If the agency has entirely reached one of its objectives and

done nothing toward the other objective, the measure of effectiveness would indicate that the agency was only 50% effective. This result is entirely misleading, and therefore, meaningless. The cause of the misleading result was the inconsistency of objectives and goals.

CJS goals must also be rational and desirable before any measures of effectiveness will have meaning. Because law enforcement agencies are operated by ordinary human beings with a perspective on the role of law enforcement that is shaped primarily by their past experience in their occupation, it is possible that existing irrational or undesirable goals may never by questioned. Of course, an irrational goal is a goal that appears to have no reasonable foundation in terms of serving recognized human requirements.

The desirability of a goal should only be determined by the constituency of the law enforcement's jurisdiction, or indirectly by their elected representatives. As this is not always the case, the desirability of the goals relatively espoused by an organization is often in question. This is not to suggest that law enforcement organizations are purposely unresponsive to those they serve, but that often in the normal course of any organization's operations, it loses touch with its constituency (in terms of goals) by virtue of not having in operation workable procedures for continuous review of the role of the organization in the society it serves. When this occurs, the pressing day-to-day operational requirements of an organization can sometimes force it to make its own goals and objectives, which need not always be in the best interests of the society within the jurisdiction of the organization. In a similar way, the equity of the goals may become unbalanced in an unfavorable direction. This means that certain elements in a society receive significantly more benefits or better treatment from the law enforcement organi-

zation than do other elements. This may or may not be perceived by the society the organization serves, but can usually be identified by an impartial investigator willing to dig through budgetary statements, observe operations, and conduct interviews.

Both undesirable and inequitous goals can prevent a meaningful measurement of law enforcement effectiveness, because the very notion of computing effectiveness seems to imply that other aspects of performance are recognized. When seeking to evaluate performance in terms of effectiveness, we make the tacit assumption that the goals defined are in the best interest of the served society, since by evaluating performance we mean to consider the value the service provided by the organization. When the goals are considered undesirable or inequitous, then the activities of the organization are less likely to be perceived as services by the society. Hence, the justification for determining whether goals are desirable and equitous before seeking to analyze for effectiveness.

CJS Objectives

Like the goals of the CJS, the objectives of the CJS and its component organizations need to be identified and analyzed before choosing a measure of effectiveness. These objectives must be checked for consistency, rationality, desirability, and equity just as the goals were checked, and for the same reason. The major difference lies in that one additional check must be made. That is, it must be shown that the objectives of the organization are consistent with the goals of the organization.

CJS Activities

The activities of the CJS are those activities engaged in by the component organizations of the CJS. Generally, it is these activities for which a performance evaluation is desired. That is, some quantification of certain aspects of the observable behavior is computed and put forward as a performance measure. Certainly, some determination must be made about the rationality, desirability, and consistency prior to attempting to select measures of effectiveness. Both activity-activity conflict and activity-objective conflict must be acknowledged if they exist. This means that the objectives must already be identified, be they formal (stated) or induced. Conflicts must be detected. If an attempt is made to develop or apply a measure of effectiveness in a situation where there is conflict between goals, objectives, or activities, meaningless results will be produced, as was illustrated in Section 4.

The Component Organization Perspective

The need for identification and classification of law enforcement activities as a prerequisite to finding or applying a measure of effectiveness has been established. It has been shown that in order to develop measures of effectiveness suitable for use in the CJS, a good deal must be known about the component organization to assure that any measurements taken are meaningful. Among the characteristics of a component organization that should be determined are;

1) the Goals, Objectives, and Activities
2) the relationship to other CJS agencies
3) the technology and environment of the organization
4) the internal structure
5) the internal workflow
6) the internal flow of resources

Analysis of these characteristics yields the information necessary to classify the activities of the organization. What is suggested here is

that knowledge of the GOA alone is insufficient for selecting a measurement approach, because performance depends on both the organization and its environment.

Organizational Goals, Objectives, and Activities

For the CJS, the goals, objectives, and activities of each component organization must be identified and evaluated before an attempt is made to find and apply a measure of effectiveness. This identification and evaluation is required to assist in classifying the activities of the organization, and for ensuring that any measurements produced are meaningful. The classification of activities, to be discussed further, is necessary to selection of an appropriate measurement methodology. To determine the GOA for a component organization, formal policy statements from the organization can be combined with observations of the organizations activities and interview information from individuals working within or affected by the organization. Some judgment must be applied in differentiating between the primary organizational objectives achieved and the secondary effects of organizational activities. It is the existence of these secondary effects which account for a difference between stated objectives and induced objectives. In characterizing an organization, both sets of objectives need be identified.

The Relationship to Other CJS Agencies

There are several approaches to quantifying the relationships between different law enforcement organizations. It is doubtful, however, that the relationship between any two law enforcement agencies can ever be entirely known. For the purposes of finding and applying measures of effectiveness, and developing measurement approaches, we can narrow down the types of relationships to be examined to those that describe how one

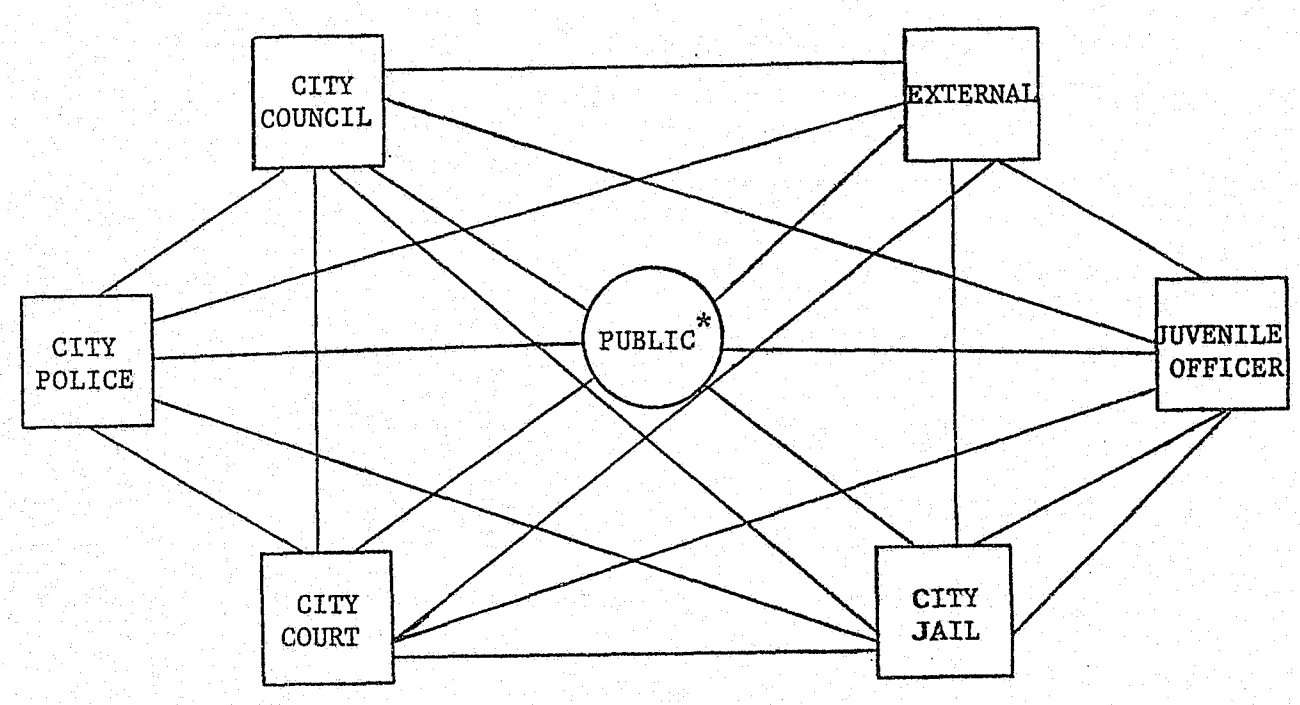
organization affects the effectiveness of another organization. These relationships obviously exist, and have been considered in Section 4. Such relationships between organizations have also been studied and modeled by such researchers as Avi-Itzhak [6].

When considering each of the component organizations of the CJS to be acting upon the effectiveness of the other member organizations, it can be seen that a complex set of relationships exist. As was suggested in Section 4, there are two basic strategies for finding those relationships. The first is the broad scheme known as multivariate analysis. Various correlative techniques are available, using such methods as regression analysis, for estimating the direction and magnitude of the effect of one agency upon another. However, as suggested in Section 4, the relationships established by the analysis do not necessarily have any cause-effect basis, causing the results to be somewhat suspect. Correlative models in the policy sciences also are often quite low in explantory power. In fact, two variables, such as the overall effectiveness of two law enforcement agencies, could be precisely determined by an independent variable, and still be poorly correlated. There are, however, certain applications of modern multivariate analysis in determining organizational interrelationships that prove useful. Multivariate analysis is particularly useful where no obvious or simple relationship exists between two organizations. In these situations, the multivariate analysis proves most useful for considering and attempting to untangle the direction of the relationships between a number of variables all at once.

A second basic strategy for determining the relationships between law enforcement organizations discussed in Section 4 is the strategy of

constructing functional models of structure of the organizations and the flow of resources, information, or command between them. This approach generally relies upon using flow elements that can be readily measured, such as crimes, criminals, or monetary resources. The main advantage to the analysis of the flow of measurable quantities between organizations is that the results are usually intuitively reasonable and are presented in an understandable model of a structure with flows. No unusual correlations are obtained, and the complete model obtained is simple to manipulate for purposes of policy experimentation. The greatest disadvantage is that some important relationships which are not intuitively obvious may not appear in the completed model. As an example of an application of functional modeling to the problem of describing relationships between CJS organizations, consider an hypothetical City Criminal Justice System as shown in Figure 8. The flow of money for the system is described by the formal budgetary statement shown in Table 1. The input-output representation in Table 2 shows not only the flow of money downward from the sources, but also lateral transfers of manpower resources measured in dollars through use of salary rates. For example, it is seen in Table 2 that the Juvenile Officer worked part time in ordinary City Police work, and spent another substantial portion of time either testifying in or assisting the bailiff at the City Court. City Police are shown to have utilized a portion of their time either in testimony, waiting to testify, or otherwise assisting the bailiff.

Figure 9 graphically depicts the flow of resources listed in Table 2. This figure is particularly revealing, showing the resource flow relationships between the component organizations. The $105,000 shown entering from external sources is meant to show grants made by State and



PRIVATE CITIZENS
SOCIAL GROUPS
BUSINESS
INDUSTRY

} PUBLIC*

NON-RESIDENTS
EXTERNAL LAW ENFORCEMENT
STATE GOVERNMENT
COUNTY GOVERNMENT
FEDERAL GOVERNMENT
EXTERNAL BUSINESS & INDUSTRY

} EXTERNAL*

Figure 8

City Criminal Justice System

Income for Law Enforcement:

Federal Grant		$105,000.
City Tax Revenue	$455,000.	
Less: non law enforcement expenditures	$310.000.	
Net Law Enforcement Tax Income		$145,000.
Law Enforcement Income		$250,000.

Law Enforcement Expenditures:

City Council		$ 10,000.
City Police		$ 70,000.
City Court		
Magistrate	$35,000.	
Expenses	$45,000.	
Total Court Expenses		$ 80,000.
City Jail		$ 65,000.
Juvenile Officer		$ 25,000.
Law Enforcement Expenditures		$250,000.

Table 1

City Law Enforcement

Income and Expenditures

Sources \ Recipients	External	Public	City Council	City Police	City Court	City Jail	Juvenile Officer
External	–	105,000	–	–	–	–	–
Public	40,000	–	–	–	–	–	–
City Council	–	–	–	70,000	80,000	65,000	25,000
City Police	–	–	–	–	3,000	–	–
City Court	–	–	–	–	–	–	–
City Jail	–	–	–	–	–	–	–
Juvenile Officer	–	–	–	5,000	1,000	–	–

Table 2

Input - Output Representation

of Resource Flow in the

City Criminal Justice System

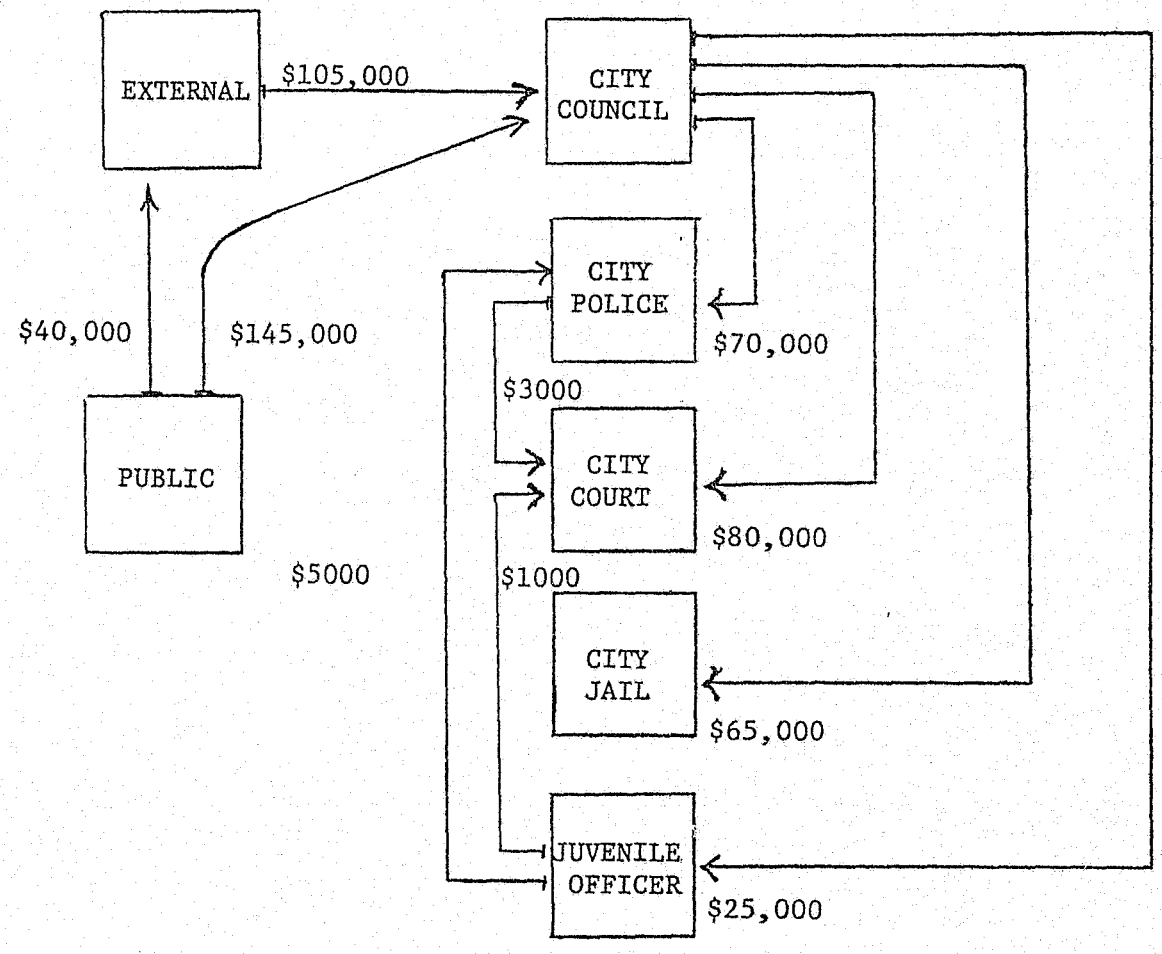


Figure 9

The Flow of Resources

in the

City Criminal Justice System

Federal agencies. Figure 10 goes one step further by showing how the resource flow relationships between the component organizations can be modeled using a signal flow graph technique developed by Mason [82] and described by Kuo [80]. The signal flow graph model represents the relationship between the component organizations as "gains". The direct flow of resources from one agency to another is equal to the sum of the resources entering the block representing the first agency multiplied by the gain shown on the directed line segment connecting the first agency's block to the second agency's block. The signal flow graph technique for relating the component organizations in terms of resource flow can be very helpful for planning resource requirements, since varying amounts of capital can be applied at the node marked "s". The major limitation is that the resource flow relationships between any two component organizations must be approximately linear in some decision interval being examined.

As a simplistic example of using the signal flow graph model with another type of flow between component organizations, consider the flow of cash through the hypothetical City Criminal Justice System. Here the assumption is made that one case represents one criminal. It is assumed that the only available information consists of

1) records showing that the destination of each criminal or suspect leaving one of the component organizations in the City CJS, and

2) an estimate of recidivision for those prisoners remanded to detention centers outside the jurisdiction, where the repeat offenses occured within city limits.

This is rather simple information to compile, particularly if all arrest reports are maintained in one place rather than at separate precincts. From these records, simple branching probabilities are computed, as shown

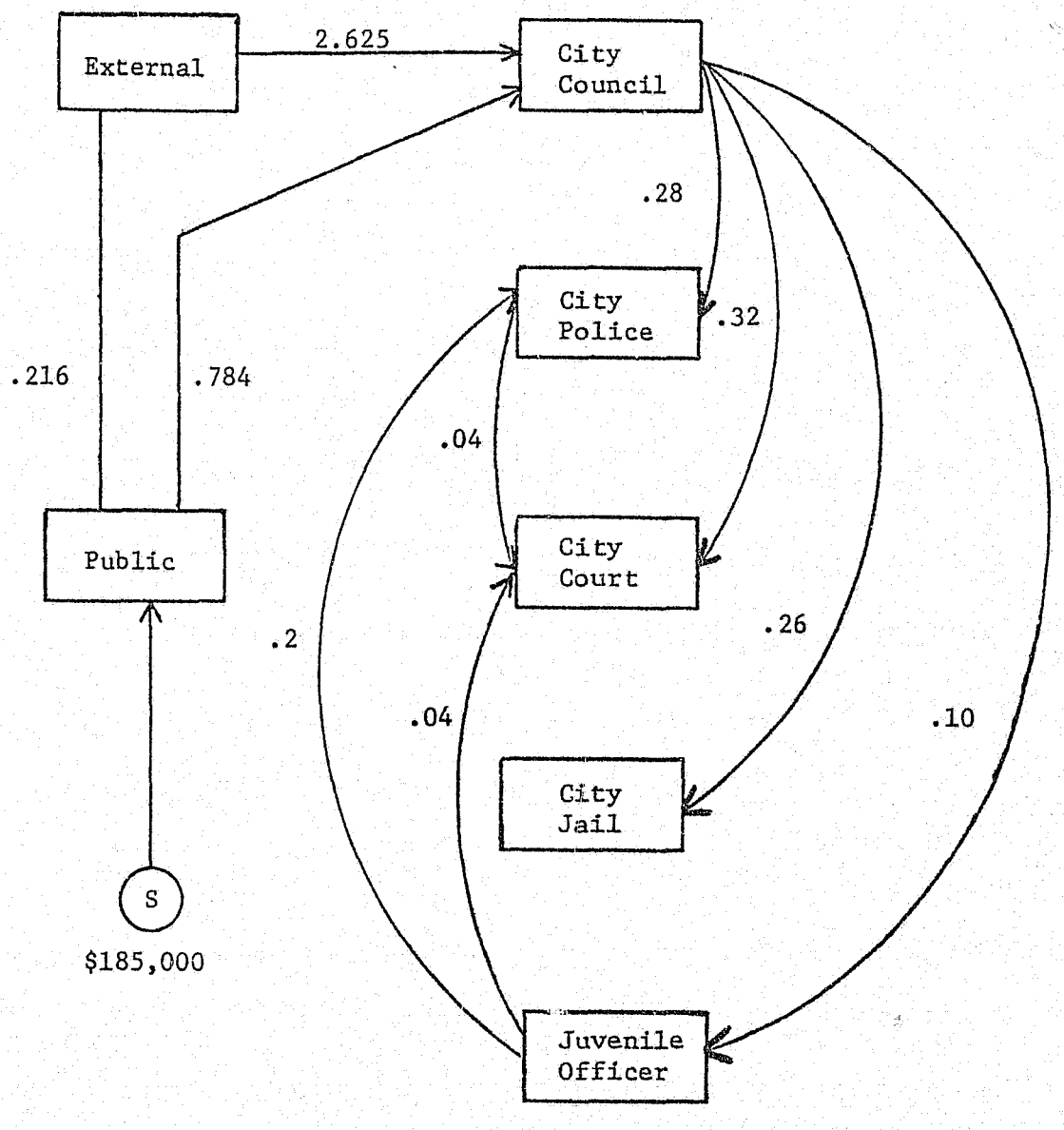


Figure 10
A Signal Flow Graph Model
of Resource Flow in the
City Criminal Justice System

in Figure 11. For example, it can be seen that the probability of another crime occurring given that a "case" has been transferred from the City Jail to an external detention centers or other entities external to the community is .80. In the model, police are seen as involved in 80% of the cases. Figure 12 displays a signal flow graph representation of the system shown in Figure 10, with an additional feature. The additional feature is the input of an arbitrarily chosen number of criminals, 500, who are assumed to commit one or more crimes against the public during a typical study period. This supply of crimes (cases) is entered into the City Criminal Justice System through the node S. It is desired to display the relationships between the component organizations by determining the impact the commission of the crimes makes on each organization. More precisely, 500 crimes are initially entered into the system, and then the number of cases handled at each component organization is computed. The computations for the signal flow graph model are listed and explained in Appendix 1. The calculations take advantage of Mason's Rule, a well-known technique for network evaluation. These computations yield the information shown in Table 3. The Gain Vector shows the "gain" between the block representing the "Public" and other blocks representing the other component organizations of the City Criminal Justice System. The "gain" between the "Public" block and any other block "B" the ratio of the output (in cases handled) at block B to the input of crimes at the "Public" block. The Volume Vector shows a particular realization of system activity where the 500 crimes have been "entered" into the system. The Volume Vector has been obtained by multiplying the Gain Vector by 500. The results shown in the Volume Vector indicate City Police will eventually

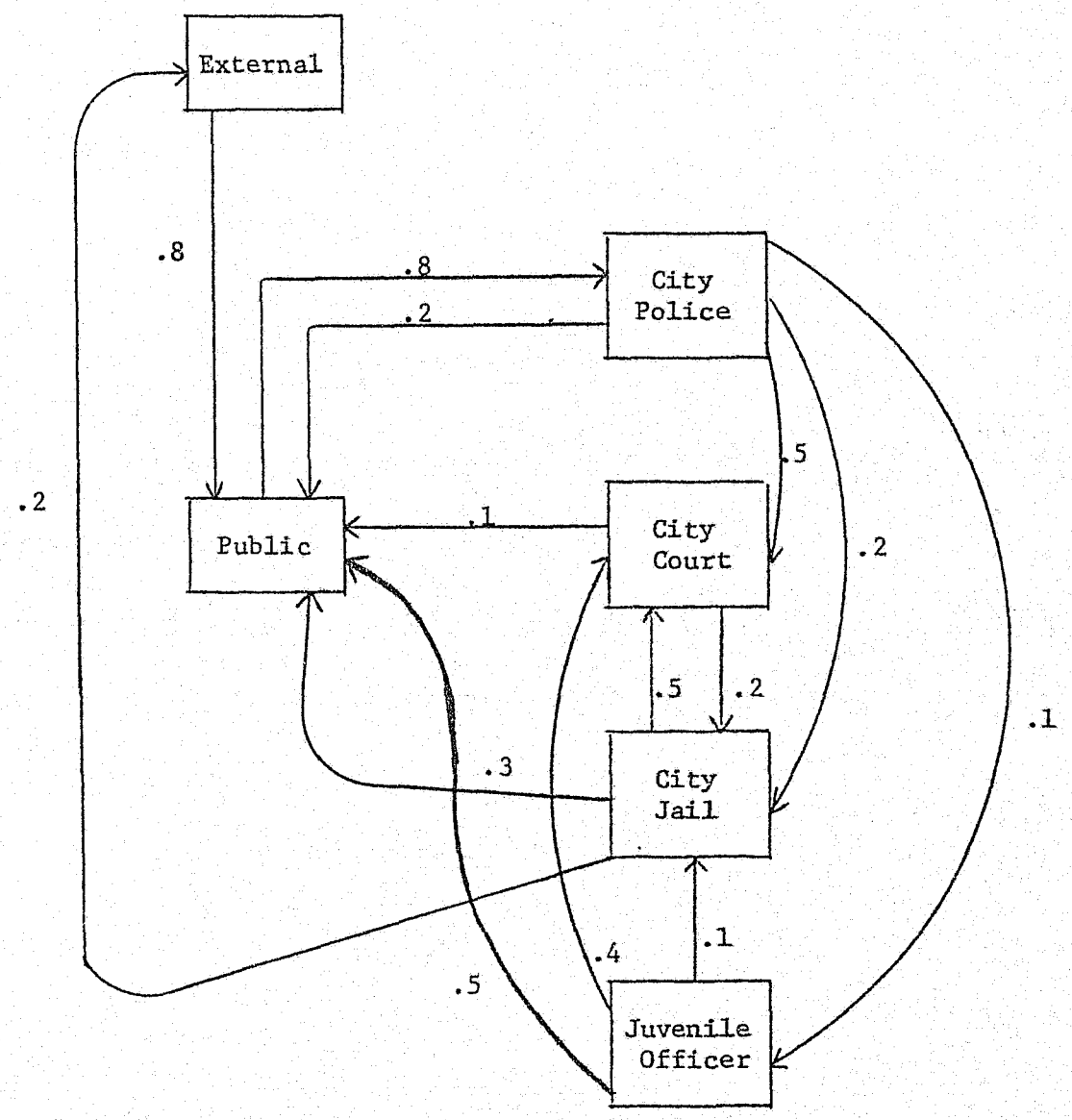


Figure 11
Probabilistic Flow of Criminals
in the City Criminal Justice System

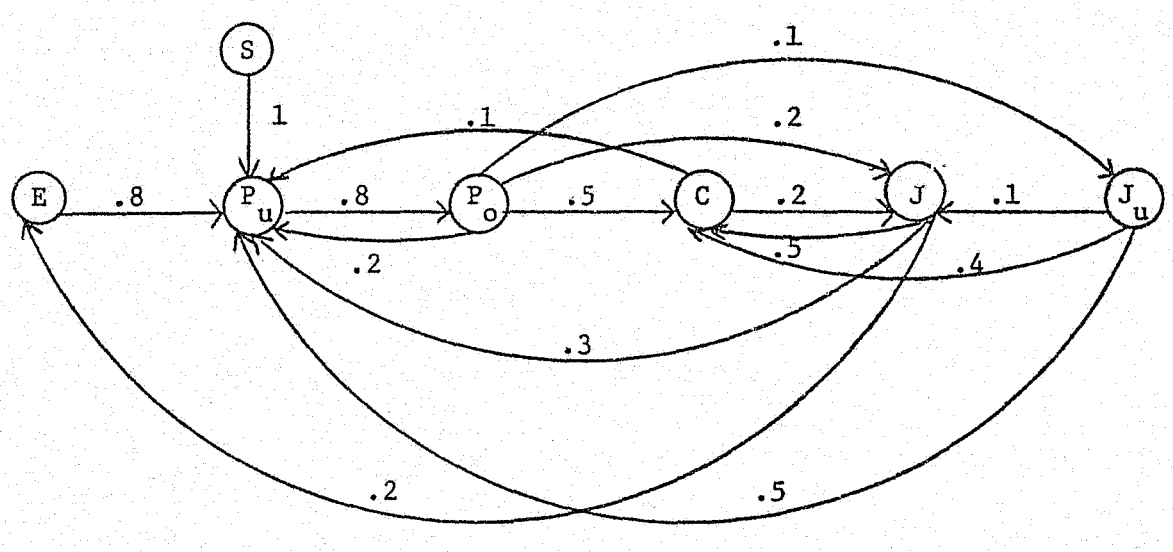


Figure 12
Signal Flow Graph Model of the Flow
of Criminals in the City Criminal Justice System

GAIN VECTOR

	External	Public	City Council	City Police	City Court	City Jail	Juvenile Officer
Public	.0778	1	0	1.1302	.6844	.3893	.1130

VOLUME VECTOR

External	Public	City Council	City Police	City Court	City Jail	Juvenile Officer
39	500	0	565	342	195	57

Table 3

Gain Vector and Volume Vector
for the Flow of Criminals in
the City Criminal Justice System

have to handle 565 cases. This shows that due to the feedback process known as recidivism, the number of total cases eventually handled by an agency may be much larger than the given initial case load, even though no new input of crimes is introduced. This has some important planning implications for a multi-component law enforcement system.

In summary, several methods of evaluating the relationships between component organizations are available. Two such approaches described in Section 4 have been demonstrated here. The two demonstrated approaches can be used not only to demonstrate relationships between organizations, but for estimating the effects of added crimes on the individual component organization within the CJS. An important consideration, which has been mentioned, is that the relationships between the organizations must be known in quantitative terms in order to assess both the effectiveness of the component and the effectiveness of the system as a whole. In addition, there must be available some means for measurement or reasonably accurate estimation of the quantitative relationship. The fact that practical measurement schemes measure observable events implies that the quantitative relationships must be expressed in terms of tangible, observable elements, or that the quantitative relationships be expressed in terms of intangible elements closely associated to tangible elements.

Technology and Environment

According to Weissenberg [77], the technology of an organization is a broad element that determines to some extent the procedures and policies of the organization. The technology of the organization must be considered during any measurements or evaluations, particularly in that the technology and environment are the limiting factors on the absolute value of the output of the organization. When evaluating an organization, the appropriate-

ness of the technology to the achieving the goals of the organization should be considered, particularly if there is difficulty in achieving these goals.

In a serious investigation, the technology of the organization is best determined by an on-site assessment. This would involve an on-site inspection of equipment and facilities, and a review of policies and procedures to determine their technological content. Within an organization, different technologies are applied for different functions or activities. The technologies involved in a law enforcement organization can be easily determined by examining the technologies evident in each of the organization's functions. Consider a very simple example. Figure 13 contains a simplified write-up of a police dispatching job, showing how the technologies involved in a job or function can be identified merely by considering the tasks involved. The write-up of the dispatcher's tasks contains enough information about the techniques involved to allow immediate identification of the technologies utilized. Indeed, these technologies and the job tasks suggest how the dispatching operations might be evaluated in terms of effectiveness. In an evaluation, the evaluators would benefit from examining job write-ups, so that this would be a logical practical step in the measurement process.

The environment of a law enforcement organization consists of those elements external to the organization. The immediate tangible environment consists of other agencies in the CJS, the criminals the agency deals with, and the individuals and organizations affected by the existence of the organization. In that the immediate environment partially or entirely determines the goals, and together with the technology the objectives, and because the notion of effectiveness is based on goal achievement, it is imperative that in the evaluation of a law enforcement agency or program

JOB TITLE; DISPATCHER

JOB DESCRIPTION: OPERATE RADIO DISPATCHING STATION FOR EMERGENCY AND NON-EMERGENCY RESPONSE. RELAY INFORMATION AND INSTRUCTIONS BETWEEN PATROLMEN AND BASE.

JOB TASKS: ANSWER PHONE. OPERATE RADIO. TAKE VERBAL MESSAGES OVER PHONE. ASSIGN PRIORITIES TO REQUESTS FOR ASSISTANCE. DISPATCH PATROL CARS.
RECEIVE REPORTS AND RELAY INSTRUCTIONS OVER RADIO.
RECORD MESSAGES FROM PATROLMEN.

TECHNOLOGIES UTILIZED: SPOKEN AND WRITTEN ENGLISH.
RADIO COMMUNICATIONS.
TELEPHONE COMMUNICATIONS.
INFORMATION HANDLING.
RECORD-KEEPING.

Figure 13

An Example Job Write-Up

a full characterization of the agencies environment be made. This characterization should constitute a "situation paper" about the environment. The document should contain complete information about the structure and function of other agencies in the CJS, and their operational and budgetary relationships to the law enforcement organization being studied. In addition, a complete demographic, historical, social, and economic picture of the immediate jurisdiction should be included. If the law enforcement organization were at the local level, pecularities of the local terrain and data on local business, industry and social organizations should also be included. Care should also be taken to show how and where the organization interfaces with its parent government, both operationally and financially. This serves to identify service consumers.

Internal Structure

The internal structure of an organization is perhaps one of the simplest features of the organization to be determined. Formal organizational hierarchies usually identify significant features of the organization, but do not necessarily give a complete picture of the work flow and the flow of resources within the organization. However, in practice, a formal organizational chart showing all operational entities within an organization and identifying the chain of command is sufficient for describing internal structure.

As an example, consider how a committee seeking a means for evaluating local safety services in a hypothetical city might portray the internal structure of the city police department. An example of this portrayal is given in Figure 14. On a larger perspective, if one considered the CJS as the organization to be studied, showing internal structure would be accomplished by combining the formal hierarchies of each of the component

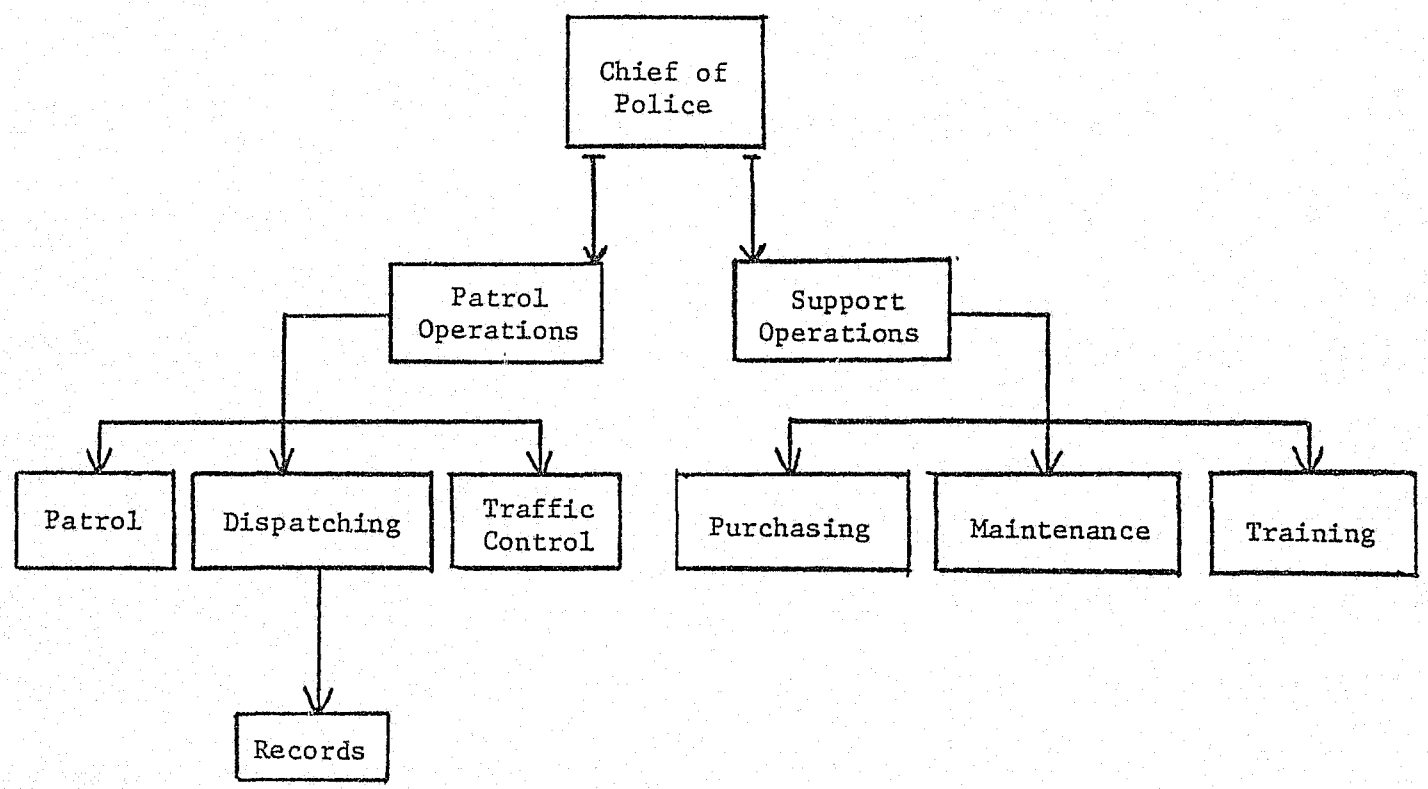


Figure 14
Internal Structure of a
City Police Department

organizations into a single hierarchy. The jurisdictional relationships between the component organizations would serve to tie each component's organizational hierarchy into the combined hierarchy.

Internal Work Flow

Internal work flow in a law enforcement organization is the flow of information and individuals to be processed through the organization. Internal work flow is best identified by a rigorous flow-charting of formal and informal procedures used for processing of paperwork and formal policies on the conduct of operations. The flow chart series to reduce a large asynchronous operation to a series of simpler sequential operations. The simpler sequence of operations more readily yields to analysis.

When diagramming the work flow internal to a law enforcement organization, care should be taken to make note of informal procedures and deviations from normal operations. Any flow-chart should display these procedures as it is requisite to show actual operation of the organization rather than only the formally prescribed procedures. Effectiveness is likely to depend on actualities rather than a formal policy statement. Verification of the flow-chart prepared is best accomplished by following the path of individual cases or jobs through the organization. This identifies the paperwork produced, points out any informal procedures in use, and also shows the nature and quantity of resources expended on a case or job.

Internal Resource Flow

The internal flow of resources is the movement and expenditure of capital, man hours, equipment hours, and other resources in the operations of the organization. This flow is closely identified with internal work flow to the extent that a major portion of expended resources involves man and equipment hours expended on individual cases or jobs. The portion of resources expended not directly connected to volume-oriented measures of work flow can be considered as overhead or burden. In seeking to determine the nature of the internal resource flow in an organization and to quantify the

flow, two methods are frequently applied. Each method serves as a check on the other, and both usually give different yet similar results. Both methods require knowledge of the internal structure of the organization, and rely upon the important conceptual principle that expended resources are related to work performed.

The first applicable method, a "bottom-up" approach, involves tracing individual transactions through the organization and recording the estimated resources expended at each step. Several such transactions are traced through the organization, and a generalized indicator of types and quantities of resources expended per transaction or case is produced. The various resources are converted into a dollar figure and summed, providing an estimate of the "variable" cost of operating the law enforcement agency or, in a larger system, groups of agencies. The "fixed" cost of operations is best developed from budgetary statements. As a simple example of the "bottom-up" approach, consider a hypothetical State Bureau of Investigation whose internal structure is shown in Figure 15. During a 3 month study period, investigators follow the handling of 10 requests for assistance from state and local law enforcement agencies. For each request for assistance, the investigators recorded, as best they could discern, the resources expended during each operation. Manpower costs were derived from the individual employee's salary and benefits. Automobile expenses were based on mileage and fuel costs. Crime lab expenses included not only manpower expenses but the approximate costs for chemical reagents and disposable equipment expended. The results of the sampling are summarized as shown in Table 4. The flow of resources corresponding to the 10 sampled assistance requests is represented in Figure 16. It can be seen that an average of about $500 was expended for each of the 10 cases. This figure is a useful operational indicator of the

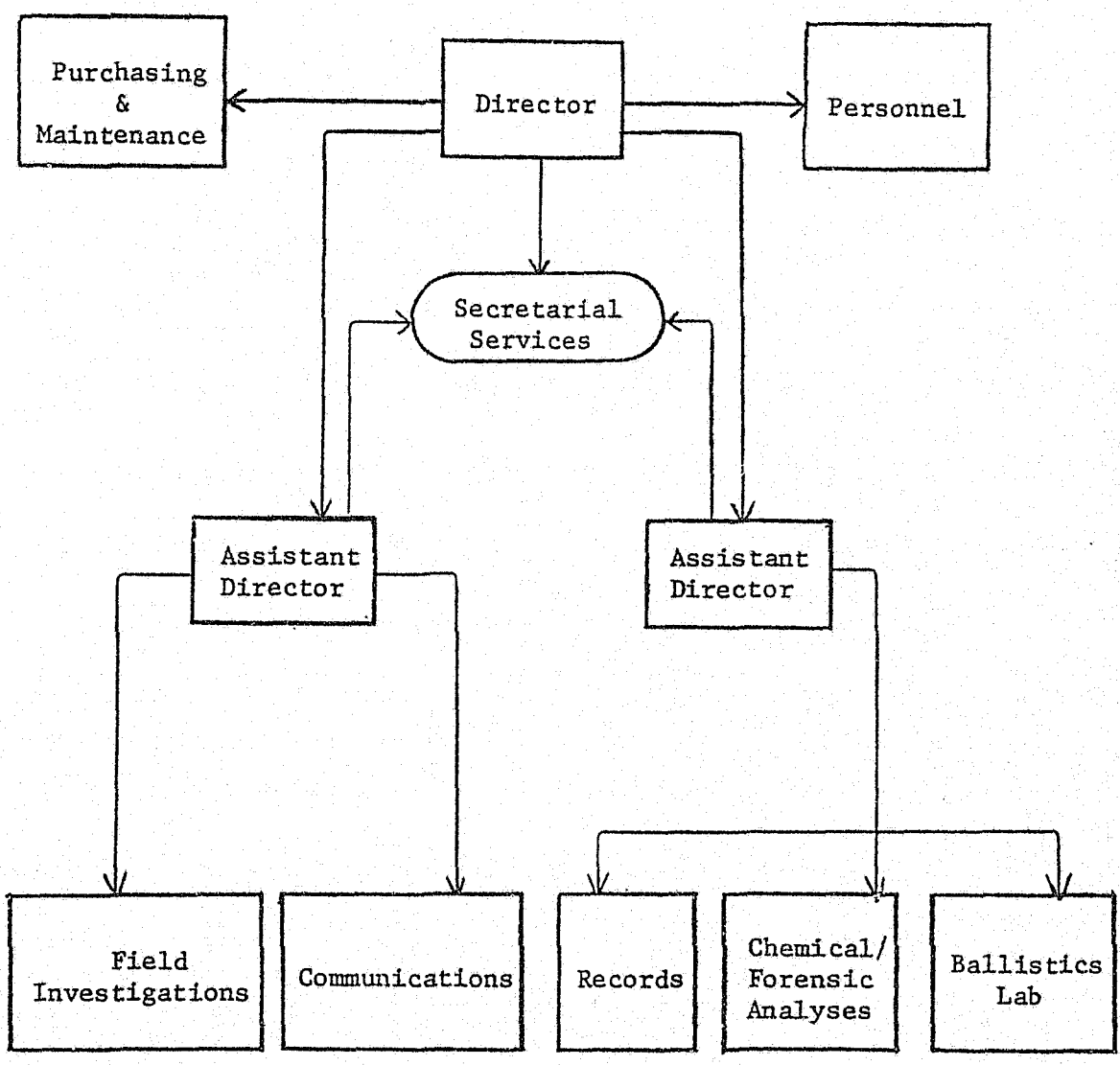


Figure 15
Internal Structure
of the
State Bureau of Investigation

	Quantity	Cost
Investigator Hours	400	$3100.
Investigator Supplies	$ 140	140
Auto Mileage	1000	100
Fuel Expanded	70	39
Messages Relayed	72	7
File Clerk Hours	6	36
Filing Supplies	$ 8	8
Laboratory Hours	40	320
Laboratory Supplies	$ 75	75
Secretarial Hours	85	595
Secretarial Supplies	$ 27	27
Supervisory Hours	65	550
		$4970

Table 4

Resources Expended on 10 Cases
by the State Bureau of Investigation

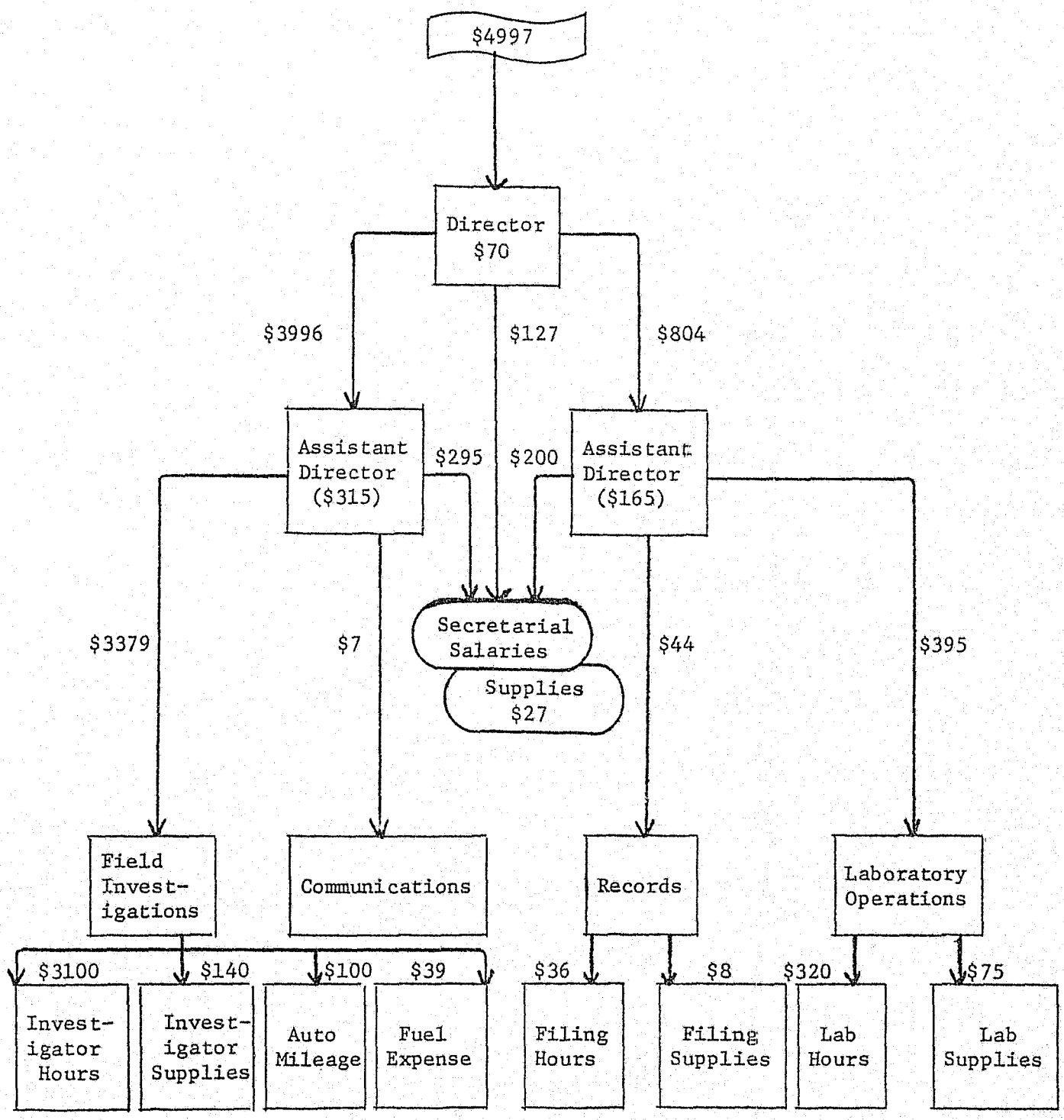


Figure 16
Internal Resource Flow for 10 Cases Investigated
by the State Bureau of Investigation

flow of resources. This sort of indicator would be best applied in short range planning for changes in the number of requests for assistance. The real shortcoming of the indicator is that the indicator does not include overhead items, such as the cost of hiring and firing, and the larger capital expenditures for equipment and facilities.

The second means for analyzing internal resource flow in an organization is a "top-down" approach involving examining the official budget of an organization, including the distribution of funds within the various elements of the organization, and equating those funds to some level-of-effort indicator, such as the number of cases or transactions. The approach is almost entirely a quantification of resource flow because only general information about the route of the resource flow is developed. It is also possible with this "top down" approach to remove budgetary overhead from the dollar costs before computing a "dollars per transaction" figure. Removing the fixed operating expenses from the total dollar costs leaves the volume dependent costs. Limits must be set on the inferences to be drawn from this sort of analysis, however, because, to some extent, all costs are volume dependent. The notion of "fixed costs" arises because many of the expenditures are insensitive to modest changes in the level of effort. For example, a correctional facility constructed to house 2500 inmates will incur essentially the same heating and lighting costs whether it has 2300 prisoners or 2400 prisoners in residence. The variation in number of inmates would most likely be felt in the commissary budget as a variable cost.

As a simple example of the "top-down" approach to describing resources flow, consider once again the hypothetical State Bureau of Investigation. From an accountant's report the data in Table 5 is prepared. The data shows total expenses incurred by the Bureau during a recent month. During this

EXPENSES	$28,041.
Purchasing & Maintenance	
Clerical Salaries	800.
Supervisory Salaries	700.
Equipment and Supplies	70.
Personnel	
Clerical Salaries	400.
Supervisory Salaries	800.
Equipment and Supplies	22.
Directors - Asst. Directors	
Secretarial Wages	1,500.
Supervisory Salaries	3,300.
Equipment and Supplies	100.
Departmental	
Investigators Salaries	8,000.
Invesigators Supplies	980.
Depreciation Expenses: Autos	860.
Fuel Expense: Autos	380.
Communication Salaries	1,000.
Communication Equipment Repairs	23.
Records Room Salaries	800.
Records Room Supplies	17.
Lab Technician Salaries	4,200.
Laboratory Supplies	415.
Secretarial Salaries	3,000.
Secretarial Supplies	89.
General	
Building Depreciation Expense	300.
Gas, Electric, and Water	285.

Table 5

Monthly Expenses for

the State Bureau of Investigation

time a daily average of 47 investigations were underway. This shows a total monthly expenditure of about $597 per investigation. About $525 per investigation in Departmental and Director's expenses resulted. Considering that an investigation usually requires 3 to 5 weeks for completion, it can be seen that the $525 estimate of operational expenses obtained by the "top-down" approach is within 5% of the $500 estimate obtained by the "bottom-up" approach.

Both approaches to analyzing internal resource flow mentioned have been demonstrated by rather simple examples, but the methods themselves can be applied to problems of virtually unlimited scale, once the two concepts are understood. Most every real system can be divided into lesser components and analyzed for internal resource flow. For example, were we to consider the Criminal Justice System as a group of component organizations as was shown in Figure 7, we could describe and quantify the internal resource flow of the system by

1. identifying the resource flow internal to each component organization, and
2. determining the resource flow relationships between each of the component organizations.

These are two separate questions, each of which can be answered through the philosophies of analysis described above.

Classification

In order to collect the activities of law enforcement organizations into groupings of activities of a similar nature, it is necessary to classify the activities of each organization by some universal set of criteria. The criteria are universal to all agencies within the CJS because the criteria are developed to serve the potential measures and measurement strategies, rather than the various activities of the component organizations. This arises because measures and strategies are activity-specific, that is, being more valid for measurement of certain types of activities than other activities.

Characterizing an Activity

There are several ways of analyzing and classifying an activity of an organization. One simple model of an activity is depicted in Figure 17. In the figure, an activity is shown to be the process that relates the operator and the operand. Here, both the operator and the operand may be people, machines, or information. The operand is that input to the process that the operator processes to produce an output. Both the inputs and outputs from the activity may consist of resources and information. Using this simple model of activity, it is seen that an activity can be characterized by three general types of information: the identity of the operator, the nature of the process, and the identity of the operand. For purposes of classification of law enforcement activities, the activities of an individual organizations can be typified and grouped using the three classes of information described.

The identity of the operator refers to a specification of the element of the component organization that is responding to certain stimuli to perform the operation. The operator can be an individual or a machine.

In larger models, such as in the instance where it is desired to evaluate an entire department at once, it may be possible to look upon the department as the operator. The nature of the process is the most important information about the activity because most effectiveness measures are objective-oriented. The specification of the nature of the process consists not only of describing the objectives served by the activity, both stated and induced, but by describing the means by which resources are expended in the process, and by describing the flow of information related to the process. In addition, the technology of the process must be defined. The technology is both the description of the logical procedures forming a process and an inherent limiting factor on how well the process works. The identity of the operand is specified in terms of both the input and the output of the activity. This includes stating the types of resources and information for which the flow is described in specifying the process.

The individuals participating in the activity as part of the operand can conceptually be lumped in with the resources, particularly with regard to the concept of flow. A question that must be considered in identifying the operand is the question of determining how extensively the identity operand is to be discovered. In other words, the process may operate directly on a particular set of persons, resources, and information, and indirectly on many other such sets. At a minimum, the first order elements of the operand, those elements directly affected by the process, must be known. It can probably be shown that those elements indirectly affected need to be identified in the operand in order to fairly assess the effectiveness of the activity. This relates to the concept of the induced objective in the sense that in both these situations, there may be a significant difference between formality and actuality. In fact, an observer attempting

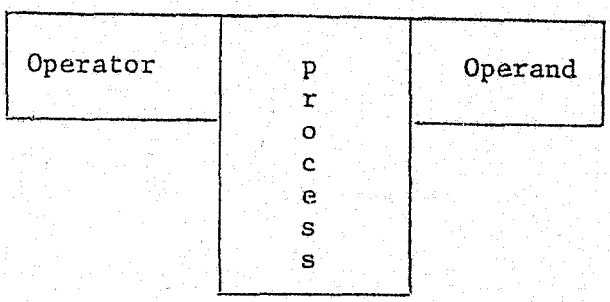


Figure 17

Simple Model of an Activity

to develop a set of induced objectives for an organization by analysis of its activities would be more interested in the indirectly affected elements in the operand than the first order elements, because there is intuitively more opportunity for variance from the desired results of stated objectives nested in the behavior of the indirectly affected elements of the operand.

Classifying Activities of the CJS

Once the individual activities of each component organization within the CJS have been identified and typlified as to the nature of the activity, a cross-organizational classification of activities can be performed to group all similar CJS activities. Figure 18 depicts the classification of CJS activities by their characteristics apparent in the Characteristics x Activities Matrix form is used, one indicating that a particular characteristic describes the activity, and a zero indicating that it does not. Along the activity scale, the activities may be part of local, regional, state, or national CJS component organizations. By arranging all the CJS activities along a linear scale and by classifying them by the characteristics of the activity alone, it is possible to arrange the activities into groupings of similar activities. It can be seen in Figure 18 that groupings can be made for activities 1 and 3, activities 2 and 4, and activities 5, 6, and 7. For example, activities 1 and 3 might represent patrol activity by local police and by local sherrif, respectively. A second matrix that will be utilized during the selection process in the matrix showing the relationships between the component organizations and the various activities. The matrix shows the component organization along one dimension, and the activity along the other; the entries within the matrix are again one or zero, depending on whether or not the particular agency engages in the activity. A simple example is shown in Figure 19.

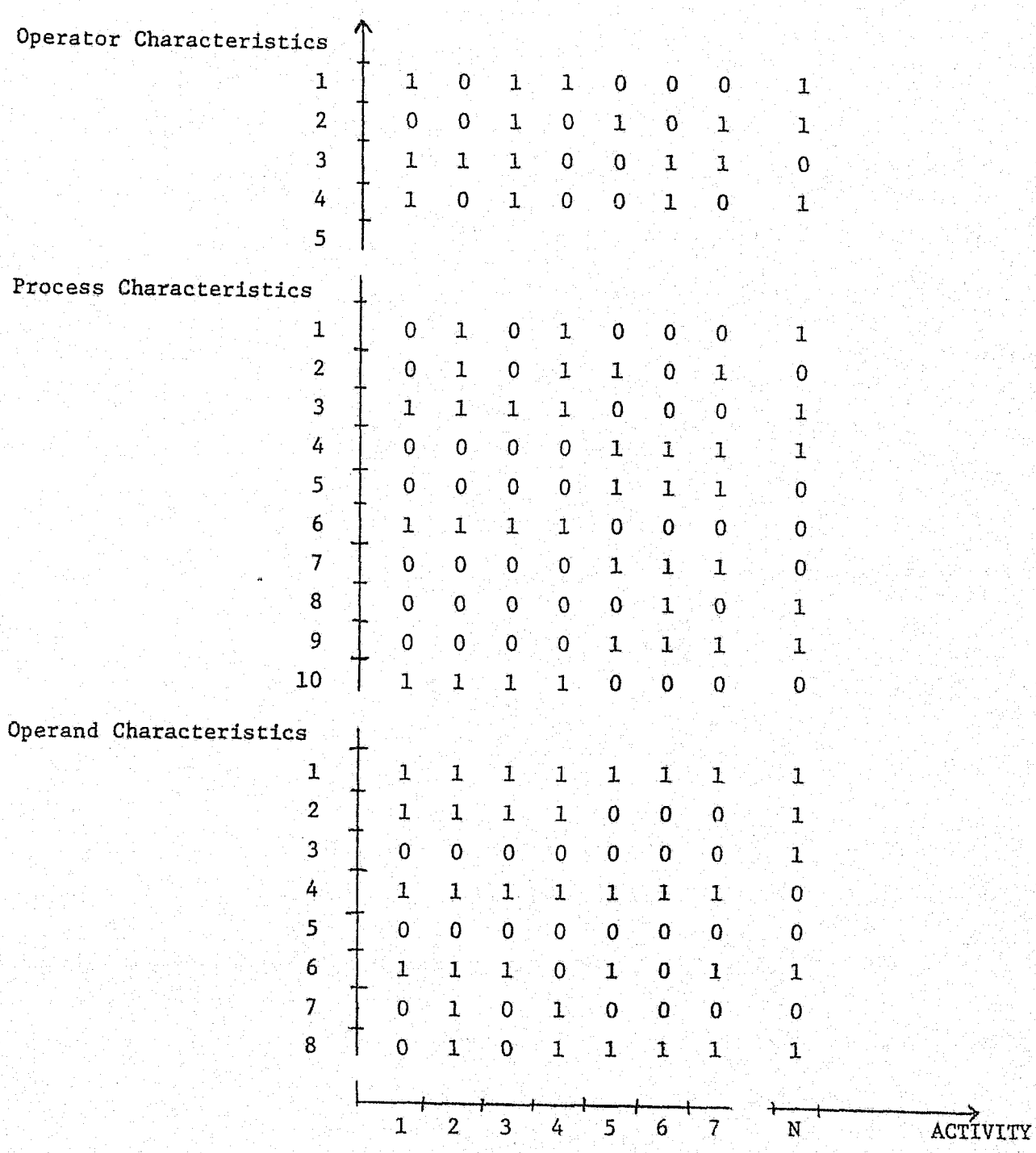


Figure 18

Classification of Activities

by Their

Characteristics:

The Characteristics X Activities Matrix

	PATROL	DISPATCHING	SURVEILLANCE	COUNTY AMBULANCE
SHERRIF	1	1	0	1
CITY POLICE	1	1	1	0

Figure 19
A Simple Example
of the
Activity X Organization Matrix

In developing the information about the characteristics of the activities, special care must be taken to include information about the nature of the interface of the activity with the public. To an extent, determination of effectiveness depends on the factors of choice, divisibility, and exclusion as explained by Ostrom [35], and Freeman [56]. This basically is the problem of trying to fairly estimate who benefits from public services, when in fact all do. Because it is hard to exclude an individual citizen from benefits of law enforcement services, and because it is difficult to determine exactly what amount of resources are expended for each individual in a society, special care must be taken when proceeding to choose a resource-oriented measure of effectiveness. Therefore, the evaluator must include within the operator, process, and operand characteristics some information as to the benefits flowing to the public as a result of the activity. Here, of course, specification of the type of benefits, rather than quantity or quality, is required.

Section 6. IDENTIFICATION AND CLASSIFICATION OF MEASURES, STRATEGIES, AND PROCESSING

There are a wide variety of approaches for effectiveness measurement currently available, some of which have been described in this paper. In addition, countless methodologies have not yet been devised but await further development effort. Once a set of possible measurement approaches for a particular application has been extracted or developed from available sources of measurement techniques, the approaches can be classified by the nature of the applications to which they best apply. This then enables evaluators to select a measurement approach suitable to their application.

Identification

The identification of a measurement approach consists of a specification as to the nature of a particular measure of performance, or more specifically effectiveness, a complementary strategy for measurement, and a measurement process to embody them. This identification process follows the general outline of approach selection given in Section 3.

Sources of Measurement Approaches

There is a multitude of existing approaches for measurement of organizational effectiveness. However, in developing new measurement approaches, the labor lies in three areas; survey of existing procedures, observations in the field, and ideation for creation of new methods. Figure 20 depicts some of the possible sources for new measures of effectiveness, measurement strategies, and measurement processes. The first source of information about possible measures of effectiveness is the relatively large body of literature dealing with individual and organizational performance measurement. This literature deals with both applications and theory. The applications

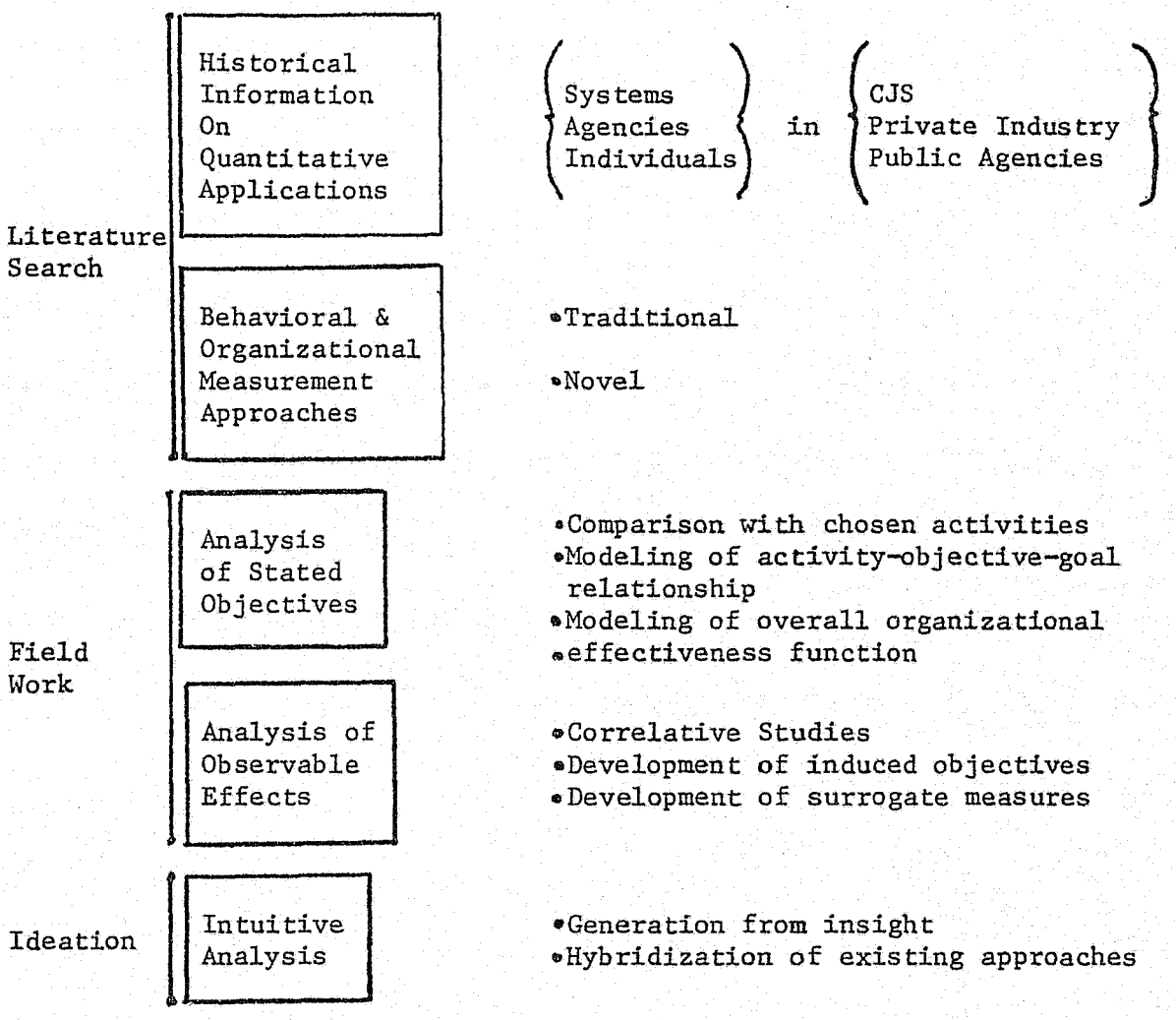


Figure 20
Sources of Measurement Approaches

cited are in private industry, public agencies, and the CJS. The second source of information is also a body of literature, but it is a rather distinct type of reporting about the theory of human and organizational behavior. It is known that behavioral scientists have developed some rather unusual methodologies for evaluating an organization. There is great possibility that some of these techniques, such as measurement of the resolution of conflict within an organization, can become the inexpensive yet accurate surrogate measures. In addition, the literature, by suggesting a means for analyzing human and organizational behavior, provides clues as to what features in the organization contribute to or detract from organizational effectiveness. A typical study by Child [28] involves identifying several of these features.

There are two distinct types of analysis that can be performed in the field that will yield the bases for measurement approaches. The distinction between the two rests in the definition of induced objectives, given in Section 2. The analysis of the stated objectives of the organization and the activities chosen to reach those objectives will yield information about the relationships between the objectives and the activities, and hopefully insight into possible measurement methods. The analysis of observable effects looks not at the objectives and activities, but attempts to identify observable effects of the activities. This includes secondary effects as well as the principle effects. From these observations it may be possible to develop a set of induced objectives, which could be compared to stated objectives as a means of measurement in itself. Within the detailed analysis of observable effects also lies the possibility of use of multivariate correlative studies for the eevelopment of reliable surrogate measures. In any case, field

observations serve to inform the evaluator as to nature of the types of behavior to be evaluated. This runs parallel to the idea of making a preliminary investigation into behavior to be evaluated, as presented in Section 5. However, in this case, the objective is to generate a variety of applicable measurements rather than to select a small group of measurement approaches.

The final source of measurement approaches is the thought process of the researcher. Totally novel concepts in measurement can result from insight into the workings of the CJS afforded by a survey of literature and observations in the field. There are a number of highly-organized methodologies utilized by the research and development community for stimulating the flow of inventive thought; one of these methods may prove useful in generating new measurement approaches. It is logical to believe that in generating new measurement approaches, or in seeking to identify existing approaches, the flow of effort should be from literature to the field, and finally to the thought process. The experience provided by the field work and the information developed from a thorough survey of literature is a sound basis for new ideas. The inclusion of both literature and data from the field in the search program ensures a good coverage of possible measurement approaches.

Classification

Once the possible measurement approaches have been enumerated, they require analysis as to the way in which they work and as to the applications to which they are best applied. Obviously, the whole point is to develop something that can serve as a valid indicator of a law enforcement organization's effectiveness. Therefore, applications within the CJS need be considered as the target applications.

Types of Approaches

Before attempting to select a measure of effectiveness and the strategy for applying it, there should be some examination of the general relationships between the different types of measurement approaches developed from the various sources discussed. It is likely that diagramming these relationships as is done in Figure 21 would be helpful. The figure shows primarily that measurement approaches, as defined in Section 4, can be categorized by the area of science which embodies the principles that are used in the approaches.

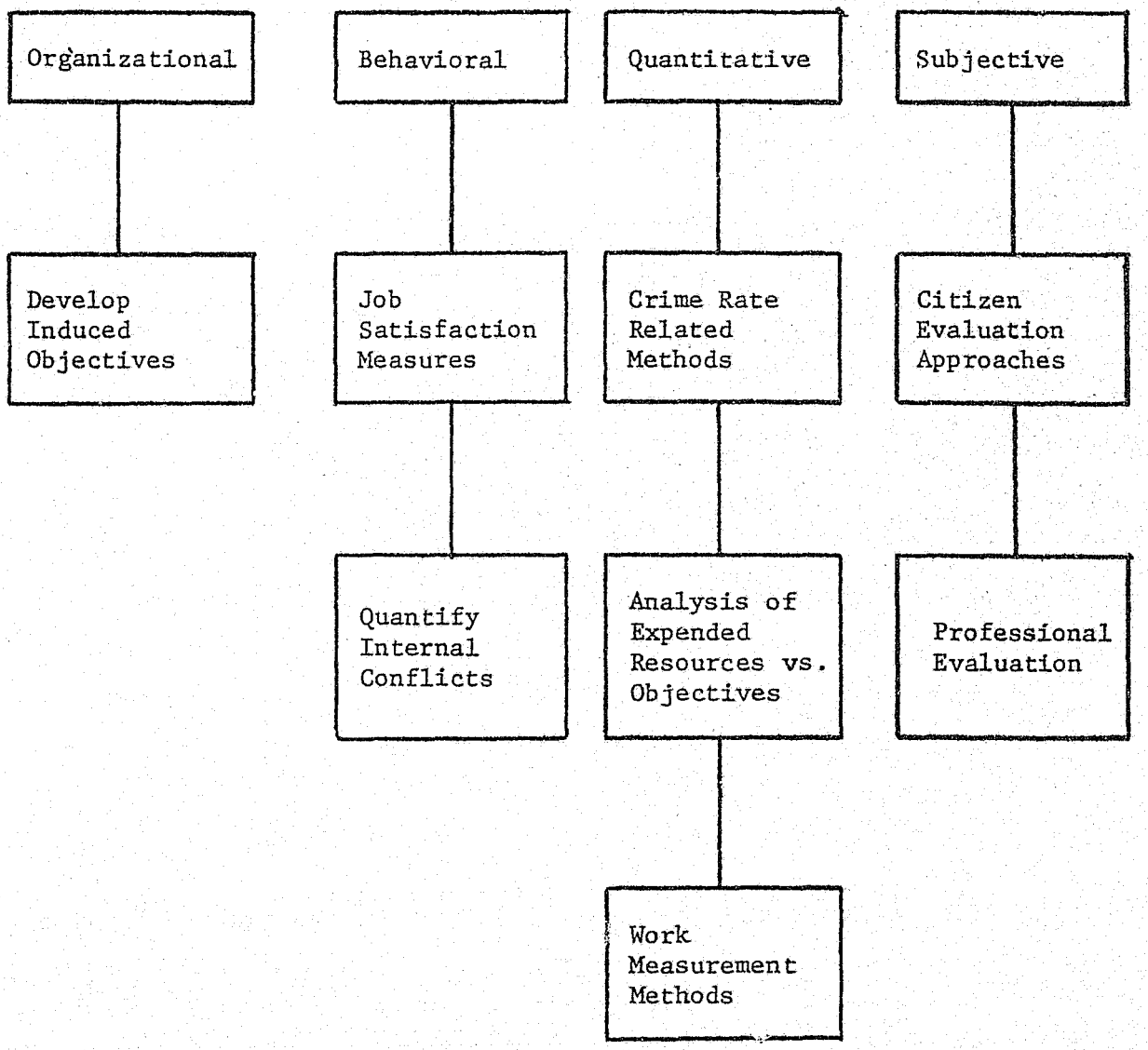


Figure 21
Describing Relationships Between
Types of Measurement Approaches

Specifying the Approach

A good deal of information is needed about an approach (measure, strategy, and process) to adequately fit the approach to an appropriate application, because not all approaches give meaningful results for every application. The information needed is primarily about;

1) the data required (its source, cost, handling),
2) the effort required for operation,
3) the nature of the output of the approach, and
4) the strengths and limitations of the technique.

The first three items of information are inherently stated in the definition of the measure and the strategy involved in the process. The fourth item is to be determined by the evaluators in light of the first three items, and their experience in using the approach. This points out the need for validation of the approach, as discussed in Section 7.

When the four data items are known for each of several approaches, they can be related to their respective approaches through a simple matrix representation as is shown in Figure 22. The final column in the Approach Matrix is produced from the previous columns, as it is an expression of fourth data item as mentioned above, in terms of best applications. The most useful applications are specified in terms of the characteristics of the activities that can be examined with the approach. These are the same types of characteristics of an activity as were used in Section 5 to describe the various CJS activities. It can be seen that, as defined previously, each measurement approach consists of a measure, a strategy, and a measurement process. This specification of the measurement approach is analogous to the specification of the characteristics of an activity, as was described in Section 5. In both cases, it is demonstrated that a simple matrix

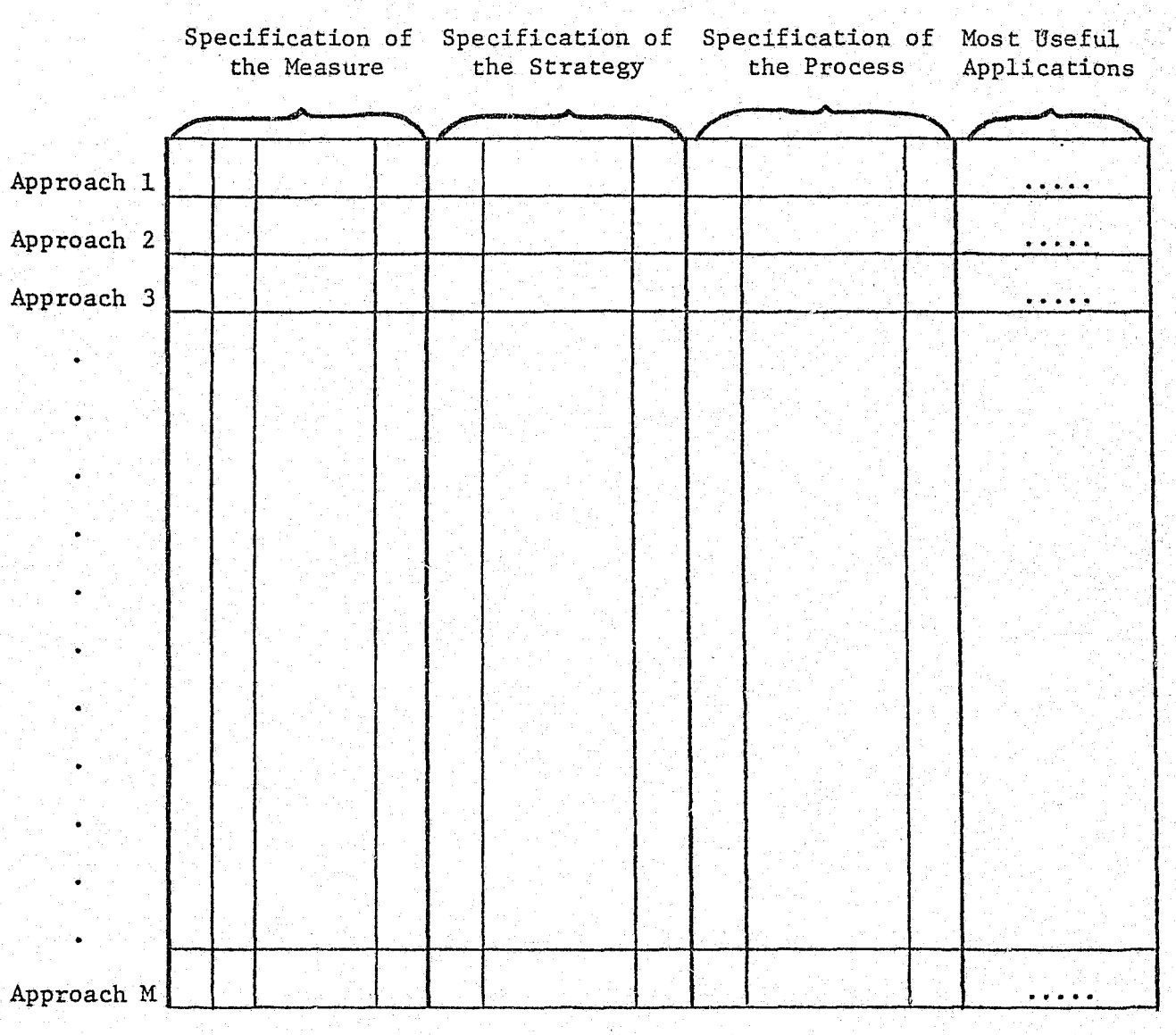


Figure 22
Specification and Classification
of Measurement Approaches:
the Approach Matrix

representation can be utilized for handling the multiplicity of specifications regarding each activity or measurement approach.

Determining Best Applications

The specification of the "most useful applications" (MUA) column in the Approach Matrix is equivalent to the determination of the relationship between a particular set of characteristics of an activity, and the "best" measure, strategy, and process for evaluating the behavior represented by that activity. The exact process for determining which measure, strategy, or process is most appropriate has not yet been developed. Section 3 demonstrated the distinctions between various types of measures, and Section 4 proposed a general method for overall approach development for a given application. Several different types of information have been identified as being necessary to relate a particular problem to the approaches which might apply.

However, what is not presented here is the decision process for making the exact determination of a best measure, strategy, and process. Only the elements necessary for or involved in the decision have been identified. This points out the need for a decision process to allow for construction of the MUA matrix. This need for a decision process is further discussed in Section 8.

Section 7. SELECTION, APPLICATION, AND VALIDATION OF APPROACH

Once the measurement approaches are known and the activities in question specified, the measurement approach can be selected. Following the selection of the measurement approach, a period of testing and validation must occur in order for the measurement approach and the results it produces to be verified.

Selection and Application

The selection of measurement approach is expedited through use of several simple matrices containing information about the characteristics of the activities best evaluated by each measurement approach or the first step in choosing the measurement approach, once the groundwork has been completed, is the construction of a "most useful applications" matrix. This is done by splitting off the most useful applications column from the measurement approaches characteristics matrix shown in Figure 22, and transposing it. This yields a Characteristics x Approach matrix. This is the MUA matrix which contains a column of characteristics for each approach. When this has been done, the characteristics of activities contained in the MUA' matrix are compared to these characteristics used to describe each of the activities of the organization(s), as discussed in Section 5. These characteristics are contained in the Characteristics x Activities matrix shown in Figure 18. Each column of the MUA' matrix is sequentially overlaid with each column of the Characteristics x Activities matrix where there appears to be a match between the two sets of characteristics, a "1" is inserted into the appropriate location in an Activity x Approach matrix. Zeros are inserted elsewhere. An example of the resulting Activity x Approach matrix is given in Figure 23. this resulting matrix contains the information needed to show which measurement approaches apply to which activities.

Time Study	Citizen Survey	Behavioral Method	
1	1	0	Patrol
1	0	0	Dispatching
0	0	1	Surveillance
1	1	0	County Ambulance

Figure 23
A Simple Example of the
Activity x Approach Matrix

To complete the selection process by showing which approaches apply to which organization, the Activity x Approach matrix is given in Figure 23. This resulting matrix contains the information needed to show which measurement approaches apply to which activities.

To complete the selection process by showing which approaches apply to which organization, the Activity x Approach matrix is premultiplied by an Organization x Activity matrix, such as the one shown in Figure 19. The resultant Organization x Approach matrix shows how many applications have been discovered for each measurement approach within each organization. In the example depicted in Figure 24, it can be seen that there are 2 potential citizen survey applications for the sheriff, and a single application of the behavioral approach exists within the city police activities.

Time Study	Citizen Survey	Behavioral Method	
1	1	0	Patrol
1	0	0	Dispatching
0	0	1	Surveillance
1	1	0	County Ambulance

Figure 23
A Simple Example of the
Activity x Approach Matrix

To complete the selection process by showing which approaches apply to which organization, the Activity x Approach matrix is given in Figure 23. This resulting matrix contains the information needed to show which measurement approaches apply to which activities.

To complete the selection process by showing which approaches apply to which organization, the Activity x Approach matrix is premultiplied by an Organization x Activity matrix, such as the one shown in Figure 19. The resultant Organization x Approach matrix shows how many applications have been discovered for each measurement approach within each organization. In the example depicted in Figure 24, it can be seen that there are 2 potential citizen survey applications for the sheriff, and a single application of the behavioral approach exists within the city police activities.

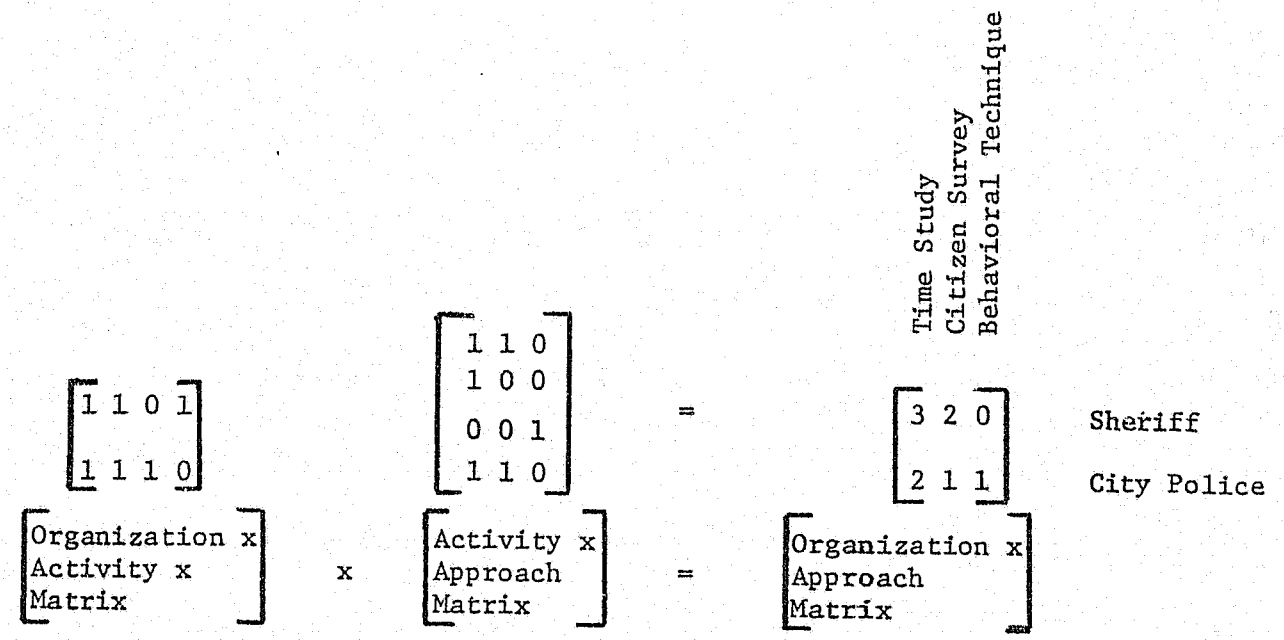


Figure 24
Measurement Approaches
Selected for Each Organization

Improving the Selection Process

It can be seen that the process of overlaying each MUA's column with each column from the Characteristics x Activities Matrix and checking for correspondence is a rather tedious process. However, with the availability of digital computers, there is no excuse not to leave this to matrix-handling algorithms. This would enable evaluators to make much precise definition of the characteristics of an activity. That is, a larger number of characteristics could be utilized without sacrifice of ease in computation.

Additionally, there is also the possibility of allowing selection of measurement approaches not perfectly suited for an application to be selected. This would be required when no "best" measurement approach is available for a particular set of characteristics describing an activity. With a little sophistication, the available measurement approaches could be evaluated, using the previously described matrix techniques, and approach with the best "fit" selected. One way this might be done is by first inserting some sort of weighting factors into both the MUA' columns and the columns of the Characteristics x Activities Matrix. For the MUA' columns, each weighting factor replacing a zero or a one would represent some measure of the magnitude of the applicability of the approach to activities described by the corresponding characteristic. For the columns of the Characteristics x Activities Matrix, the weighting factor inserted for each characteristic would represent the degree to which that characteristic described the activity. Once weighting factors are inserted, each MUA' column can be multiplied by the transpose of each column from the Characteristics x Activities Matrix, yielding a new sort of Activity x Approach Matrix which, unlike Figure 24, consists of non-zero scalar values representing the degree

CONTINUED

2 OF 5

applicability of each measurement approach to each activity. For each activity, the measurement approach with the highest degree of applicability would then be selected.

Application

The method of application of a specific measurement approach is entirely dictated by the elements of the approach. The measure of effectiveness chosen determines what quantity is to be measured and what terms the results will be given in. The measurement strategy specifically identifies the source and type of data, how it is to be gathered, and when it is to be sampled. The measurement process dictates the procedure by which the evaluators apply the strategy and interact with the results. The measurement process must necessarily be capable of allowing for the alteration of the measurement strategy, such as might be necessary due to changes taking place in the data source during the course of the evaluation.

There is reason to believe that a measurement process can be designed to be self-improving. This refers to both increasing the efficiency of the overall measurement process, and to improving the quality and significance of results obtained. Surely selection of the measurement process design is a critical factor in determining the success of the evaluative effort. If there are general principles that can be utilized to design self-improving processes that tend to gravitate toward optimal operation, then researchers ought to develop the principles for application in CJS performance measurement.

Validation of Approach

Once a particular measurement approach has been proposed as a good way to assess the effectiveness of a certain organization's activities, there need be some validation to ensure that;

1) the approach yields accurate and reliable results, and

2) the results are being properly interpreted.

Accuracy and Reliability

The real validation of a measurement approach lies in the accuracy and reliability of the results its application produces. There are at least three fundamental ways to assess the accuracy of results produced by an approach. A different measurement approach is utilized in each. To test for accuracy, the results obtained by the first approach on the original problem can be compared to;

1) the results obtained by a different approach applied to the same activities or

2) the results obtained by a different approach applied to very similar activities.

A third possibility is indirectly applying the measurement approach in question to similar activities of another organization that has been previously evaluated using a different approach. Although the choice of a validation approach lies with the evaluators, the same principles that determine a best measurement approach determine a best validation technique. The only distinction is that researchers would prefer to apply a dissimilar validation approach if possible.

Reliability is merely the consistency of the results obtained. A measurement approach that produced significantly large variations in its assessment of an organization's effectiveness over a few relatively short time intervals could be identified as unreliable. In general, the fundamental nature of many of the approaches available leads to consistency of results. Accuracy is more likely to be a problem than reliability, as it is an absolute problem, not a relative problem. That is, it is possible to apply a measurement approach to a particular organization and produce

results over a period of time that are consistent relative to each other. There is no guarantee, however, that the measurement approach is tainted by a systematic bias, which destroys accuracy without affecting reliability.

Proper Interpretation of Results

A measurement approach is of no value if in the measurement process faulty conclusions are drawn as to the significance of the results. This will especially be a problem when the evaluators are not agreed as to the accuracy of the data, such as when victimization survey data is used in place of official crime statistics. A second problem of requirement is that the results must be available in an understandable form. Obviously, human elements within the evaluated organization will hold the results suspect unless it is clear just how the measurement approach works. Therein lies the need for avoidance of difficult and cumbersome procedures in the measurement process itself, and the need for results specified in workable terms. This sort of avoidance is best had through preparation of flexible "packaged" evaluation programs for internal uses, and through good communications when evaluators are external to the organization.

A major requirement of packaged programs is that they fit the organization. The problem for the researcher designing such a packaged evaluation program is to introduce flexibility while eliminating ambiguity. Managers do not respect packaged efforts that are vague and do not seem to apply to their organization. There is no doubt that managerial cooperation is essential to any internal evaluative effort.

Similarly managerial cooperation is of great value to external evaluation groups. This cooperation is of great assistance during the preliminary investigations, and is useful for identifying potential sources of data concerning the organization's activities. Finally, the organization's

efforts to make improvements on the basis of evaluations are likely to depend on the degree of credibility of the results among managers. Giving them a meaningful part in the evaluation promotes this sense of credibility.

Section 8. RESEARCH NEEDS IN EFFECTIVENESS MEASUREMENT FOR THE CJS

Thus far, there has been established a general methodology for selecting and applying measures of effectiveness in the CJS. Throughout the paper there have been mentioned a number of significant concepts which deserve further study. The fruition of such work will be a gain in the insight into effectiveness measurement and understanding of the component interrelationships in the CJS so that many new improvements to the operation of the CJS can be made.

In this section, several current works on performance measurement for social services are discussed. Each of these works deals with a particular subset of the overall conceptual package presented thus far. Following a brief analysis of these papers, some of the research needs identified by this paper are described.

Current Measurement Efforts

There is a great deal of current literature available on performance measurements for the social services. A representative sample of this literature is discussed here to identify some of the inherent weaknesses of the approaches described and to demonstrate how the lack of a coherent conceptual basis for measurement has reduced the significance and value of these efforts. In general, current literature is flawed by the lack of the proper preparatory work in understanding the meaning of measurements. Many efforts totally ignore the difference between measures of effectiveness, measurement strategies, and measurement processes. In point of fact, most reported applied research is entirely concentrated on selecting performance measures, and of these papers, there seems only to be the desire to distinguish between quantity and quality of services provided.

In general, few papers develop measurement schemes that give any effectiveness ratings to the organization on the basis of its making a positive contribution to the overall system in which it operates. Finally, several of the measures of effectiveness offered contain terms representing behavior not in the control of the measured organization. This fault has been discussed previously in Section 4.

Holzer [8] presents a management-oriented productivity measurement paper for application at the urban police force level. Although the paper has a number of interesting points, such as a recognition of the need for a conceptual framework for measurement, there are a number of serious problems with the paper. The first and most basic is the use of productivity as a performance measure for police services. The very use of the word "productivity" implies that there is a significant positive relationship between police efforts and the production of services to the community. As has already been demonstrated in Section 1, there is a great deal of evidence to the contrary.

Within this paper the word "nondeterrence" appears in relation to effectiveness measurement. The implications of the use of such a word are astounding, yet go unnoticed. In short, there are two implications involved. The first and foremost is that police control victimization levels, and that greater police effectiveness reduces these levels. As mentioned in Section 1, this is not generally the case. Secondly, until such policy experiment approaches as those discussed by Jones [17] and Ross [18], (approaches for detecting and analyzing the effects of a public policy change over time), are fully developed and tested in CJS applications, there can be no certain analytical basis for attempting to attribute the non-occurence of certain events to changes in police behavior. Certainly there are many cases where police behavior changes such as the addition of new

programs cause decreases in reported crime. The problem lies in determining the degree to which the changed behavior actually changed reported crime. In other words, if there is a variation in reported crime, finding the part of the variance due to a change in police behavior is the problem.

Another basic problem with the content of the paper is the near complete exclusion from consideration of the measurement strategy and the measurement process. This leaves the potential user of some of the suggested measures of productivity without any suggestions as to the means for collecting and analyzing the data, and using the results. Perhaps this is the real bridge between theory and application that must be crossed to reach conceptually sound applications yielding meaningful and useful results. There must be considerable development of measurement strategies and processes before the application of any performance measure can yield consistent, meaningful results at reasonable cost.

A final criticism of Holzer's paper is that there is no attempt to consider the police force as a component organization in a larger CJS. Instead, the paper is oriented toward maximizing productivity at the component level. As has been previously suggested in Section 4, optimizing performance at the component level does not always lead to optimal system performance. The overall optimization of CJS performance requires that a set of system goals and objectives be developed, and that optimization of component organization behavior be performed in relation to the system objectives. This will be discussed further in this section.

Police Productivity: Hirsch and Riccio

Hirsch and Riccio [16] discuss some of the aspects of productivity measurement for the police patrol. The goals, objectives, and activities of the police patrol are identified, and several popular performance

measures are proposed as means for tracing poor productivity to its source. Throughout the paper, several ratios are offered as indicators of productivity, quality of arrests, efficiency, and effectiveness. In addition, the need for applying multiple measures is recognized as a means for obtaining a more detailed picture of organizational performance.

Among the shortcomings of this effort is the primary assumption that productivity is a good measure of police performance. As was previously discussed, productivity is quite limited in its meaning for applications in the CJS. In addition, it should be pointed out that productivity measurements rapidly develop into unpopular quotas, a problem recognized by Hirsch and Riccio [16].

Another major difficulty apparent in the paper is that no groundwork is laid for selection of performance measures. Instead, the various types of measures are recognized and examples of each are suggested. Little is done to demonstrate a process for selection of these measures, and no consideration is given to measurement strategies or the design of a general measurement process. The measures suggested are not related to the major methods of applying them.

Finally, many of the measures proposed are contaminated by factors outside the control of police. This does not refer to the question of police control over crime. Rather, several of the measures given depend on the behavior of other CJS components. For example, the percentage of arrests that result in convictions is proposed as a measure of the quality of arrests. This percentage has a great deal to do with the judicial process and the effectiveness of the prosecutor in pursuing his caseload. Again, the main cause for this deficiency is that no logical decision process was developed for selecting performance measures, and that there

was no consideration given to providing complementary measurement strategies and corresponding measurement processes.

Output Measurement: Fisk and Winnie

Fisk and Winnie [37] discuss the current status of output measurement in the United States, particularly at the local government level. The emphasis in their paper is on the growing use of quantity and quality output measures by local government. The foremost feature of the paper that is missing from other works is the recognition of the need to develop both the measurement and the measurement strategy. There is mention of the need to select a measurement strategy that is affordable and new types of inexpensive measurement strategies are discussed. In their discussion of output measurement, an attempt is made to classify output by quantity and quality. Quality is defined as both impact and what Hatry [79] calls effectiveness. Several illustrative examples of quality and quantity measures are given.

Although the difference between the measure and the measurement strategy is mentioned, there is no discussion of the measurement process. Instead, there is a discussion of a limited feature of the measurement process, that is, the ultimate use to which information from measurements is applied. There is no specific given as to how a local government might design a measurement process for its measurement requirements. Finally, there is no particular demonstration that quantity and quality measures of the type shown are the best measures available, or that they had been selected by a decision process that generates good performance measures. Specifying quality as effectiveness is not wholly unreasonable, but equating effectiveness to impact is another question. In fact, effectiveness is impact only when the selected objectives refer to behavioral changes to be exhibited by the environment

in response to the activities of the organization. In general, there is a distinction between objective-oriented measures and impact-oriented measures, as described in Section 3.

PPBS: Mushkin and Cotton

The idea of characterizing the performance of public agencies through use of volume and quality indicators is at the very heart of the PPBS System as described by Mushkin and Cotton [75]. Briefly, Mushkin and Cotton simply list a number of what they consider to be volume and quality indicators suitable for analyzing and evaluating public expenditures. The measurement concepts are similar to but more simplistic than those given by Fisk and Winnie [37]. Because of this, each and every criticism cited for the previous paper applies here, particularly the fault of not presenting the measurement strategy and measurement process concepts. The fact that only the measures themselves are given causes misapplication of these measures, which in turn denigrates the entire PPBS concept. Misuse of measurements due to the lack of development of measurement strategies and processes that are practical and understandable has led to the view of PPBS procedures as obstacles by such authors as Conrad [53]. A basic problem with the PPBS approach is that it relies upon output and quality indicators as performance measures. Certainly these quantities can be utilized in certain circumstances, but total reliance upon these measures reduces the degree of accuracy with which organizational performance can be determined.

The general methodology of the PPBS approach seems to lie in the maximization of the volume of output per dollar expended while maintaining the quality of output within certain limits. The deficiency with this approach is the reduction of "quality" from an objective to a constraint. This reduction is closely associated with a desire to reduce the value of

all services flowing from an agency to equivalent dollar value, either by presuming that increasing dollar expenditure is synonymous with improved or increased output or by attempting to affix a price tax to the execution of each type of activity of the organization. Where the evaluators cannot express benefits in monetary terms, there seems to be the tendency to apply a cost/benefit ratio, which of course leaves no room for "quality" as an objective, and tends to reduce multi-dimensional measures of performance to a single ratio. This results in unparalleled ambiguity and lack of real specificity of meaning. Hatry [79] and others have identified this weakness in measurement practices. It is quite clear that PPBS is inappropriate for CJS applications due to its weaknesses described here.

Output and Efficiency: Ostrom

Ostrom [14] develops definitions of output and efficiency, and demonstrates potential output and efficiency measures and their application. There seems to be a desire to rely upon output and efficiency as performance measures simply because these measures "have most frequently been utilized in describing private sector relationships," Immediately it can be seen that the tacit assumption has been made that output and efficiency are the finest performance measures available, and that there has been no recognition given to developing a decision process for selecting performance measures. It is suggested within Ostrom's paper that police do not have complete control over crime, and that crime actually is determined by a complex interaction between the populace and social and private institutions.

Within Ostrom's paper, an attempt is made to demonstrate and classify some of the different types of police activity by identifying the consumption process and production process associated with the benefits produced by the activities. It is suggested that the case of measuring output and efficiency

depends on which of four types of police activity specified is being measured, and the importance of using both producer and consumer data in these evaluations is mentioned. Some special "shadow prices" are suggested as measures of the private costs of crime used for estimating the value of services provided by police. Finally, there is some suggestion made as to how to apply the proposed measures. It is suggested that similar systems studies are easier to operate than other measurement schemes, and that ordinal rankings, although not as precise as exact measures, do provide a sufficient basis for drawing inferences regarding the performance of an evaluated agency. The need for identifying and classifying police activities mentioned by Ostrom has been explained in Section 4 of this paper as a logical step in the decision process for selecting a measurement approach. Further discussion of the step is given in Section 5. The classification of activities has been shown to be a prerequisite for the selection of the performance measure, and subsequently the measurement strategy and measurement approach. This is a considerably larger endeavor than is discussed in Ostrom's paper, particularly in that Ostrom only distinguishes four basic types of police behavior.

The similar systems studeis suggested by Ostrom's paper are but one of the many possible approaches to determine the values of performance measures for a given application. There is merit in the argument that this type of study can be of value when the numerical values of the performance measures chosen are difficult to find. The problem is, once again, that the lack of a well defined decision process for selecting performance measures, measurement strategies, and measurement processes has led to the rather arbitrary selection of a performance measure and measurement study.

Social Service Measurement: Mantel, et.al.

One of the most recently reported social service measurement efforts concerned an evaluation of a group of 16 related community service agencies in Cleveland, Ohio. The agencies offer a variety of services from medical treatment to vocational counseling. The objectives of the evaluation were to develop data on the services and to produce evaluative models capable of rendering a standardized assessment of the services. There was a literature search conducted, the conclusion being parallel to what has been reported in this paper, that reported efforts have been limited in scope and rather superficial in their treatment of the problems associated with performance measurement for public services.

An effort was made by the researchers to identify the goals of the community service agencies, but no distinction was made between goals and objectives, particularly because of differences in agency environments. This lack of distinction between goals and objectives is primarily a matter of definition. The researchers provide two system goals which, on close examination, can be seen to be objectives, as described in this paper.

The researchers made an attempt to classify each and every service (activity) of the agencies by preconceived definitions, but failed. What did result was the development of a set of general categories into which all the services could be located. From this, the notion of a service "package" was developed, showing the agency, the consumer, and the service. Each agency was then to be rated on how well it delivered these packages, with results to be weighed heavily on the more "important" packages.

An elaborate scheme was developed to use the Delphi Method to identify the most important services provided by the agencies studied, and then to assign two sets of weighting factors for use in a utility function that

relates the satisfaction of a multielement criteria score to a utility score. The Delphi was then invoked again to assign weighting factors to convert the computed utility factors into a quality index. In concluding the paper, the authors stated that even though an intricate model of performance had been developed, it did not in fact guarantee that each agency would be fairly rated by each of the selected criteria.

Overall, the paper contains several of the important elements of the process for selecting a measurement approach, as discussed in Section 4. There are however significant omissions. For example, although objectives and activities were identified, there was no mention of any check to determine whether or not these objectives and activities were mutually consistent and desirable, or, in the terminology of Poland [32], there was no check for "appropriateness" of objectives. There is little mention of any decision process behind the selection of the six "criteria" used in the measurement process, nor is it brought to the attention of the reader that these six "criteria," such as efficiency and accessibility, are usually considered as performance measures in themselves. Therefore, there is no measure selection decision process as defined in the paper.

A major weakness also is apparent in the construction of a single quality function for use among all the agencies. The same six weighting factors were used, regardless of the agency, to transform the six utility functions (corresponding to the fullfilment of the six criteria) to a quality index. This represents the implicit assumption that it is as important for a recreational program at a gymnasium to be effective as it is for nursing and health care services to be effective. The authors, as previously mentioned, acknowledge this weakness but do not offer any remedy.

Although the researchers in the reported investigation report trying to integrate their investigation report trying to integrate their measurements to existing record-keeping systems, and although there is mention of the need for keeping measurement costs within reach of the agencies, there is no real effort at designing or even presenting a measurement strategy, not to mention development of a measurement process. There is not even mention of any decision process for selecting measurement strategies and processes.

In brief, there seems to be an implicit recognition on the part of the researchers that there is the need to identify the goals, objectives, and activities of an organization, and that multiple complementary measures are better than solitary measures. There is also the tacit recognition of the measurement process, although it is not mentioned or developed. The basic problem, as with most papers on performance measurement for public services, is that the conceptual groundwork for measurement lies ignored and undeveloped due to the headlong rush to identify and find new ways of calculating measures.

Areas for New Research

A good deal of existing research into performance measurements for CJS applications has been shown to be fragmented and superficial. In general, past efforts do not form a coherent body of work because new applications were made as each new measure or measurement approach became available, or popular, as the case may be. Little effort has been expended toward developing a truly universal foundation for the measurement of organizational behavior. As these preliminary attempts to apply quantitative and qualitative measures developed, however, legislative and public interest in performance has markedly increased, giving rise to the need for new and better measurement approaches.

Perhaps now it is time to reject many of the superficialities of the first wave of analysis attempts, and make a concerted effort to get at the really basic concepts behind measurement. There is real need to begin with the fundamentals, formulate new basic concepts, develop logical approaches for selection of measures, strategies, and processes, and to apply and test the new developments. Throughout these new development efforts, there must be ongoing concern for maintaining a uniformity and completeness of thought. There must be particular attention paid to using the broadest of perspectives in developing this new basis for measurement; the basis must be general enough so as to apply to all CJS applications, yet complete, consistent, and free of unfounded assumptions.

Besides development of the general rationale for measurement, research efforts must follow through from the development of the conceptual hardware for measurement to the more specific developments and investigations required at the CJS and component organization level. As a great deal of performance measurement hinges on organizational and CJS objectives, there is also a requirement for specific studies in this area, almost to serve as a bridge between the general concepts of measurement to be developed and the more specific aspects of CJS and component organization performance measurement. The research efforts to be described in the remainder of this section represent not only gaps in current knowledge, but the development of the basis for new and more powerful measurement approaches. These efforts are suggested as means of eliminating the current confusion about the meaning of measurements, and to bring practical measurement methodologies into greater numbers of more significant applications.

The format for the remainder of this section is directed by the numbering scheme developed for identifying the logical hierarchy of different new research programs depicted in Figures 24, 25, 26, and 27. There are

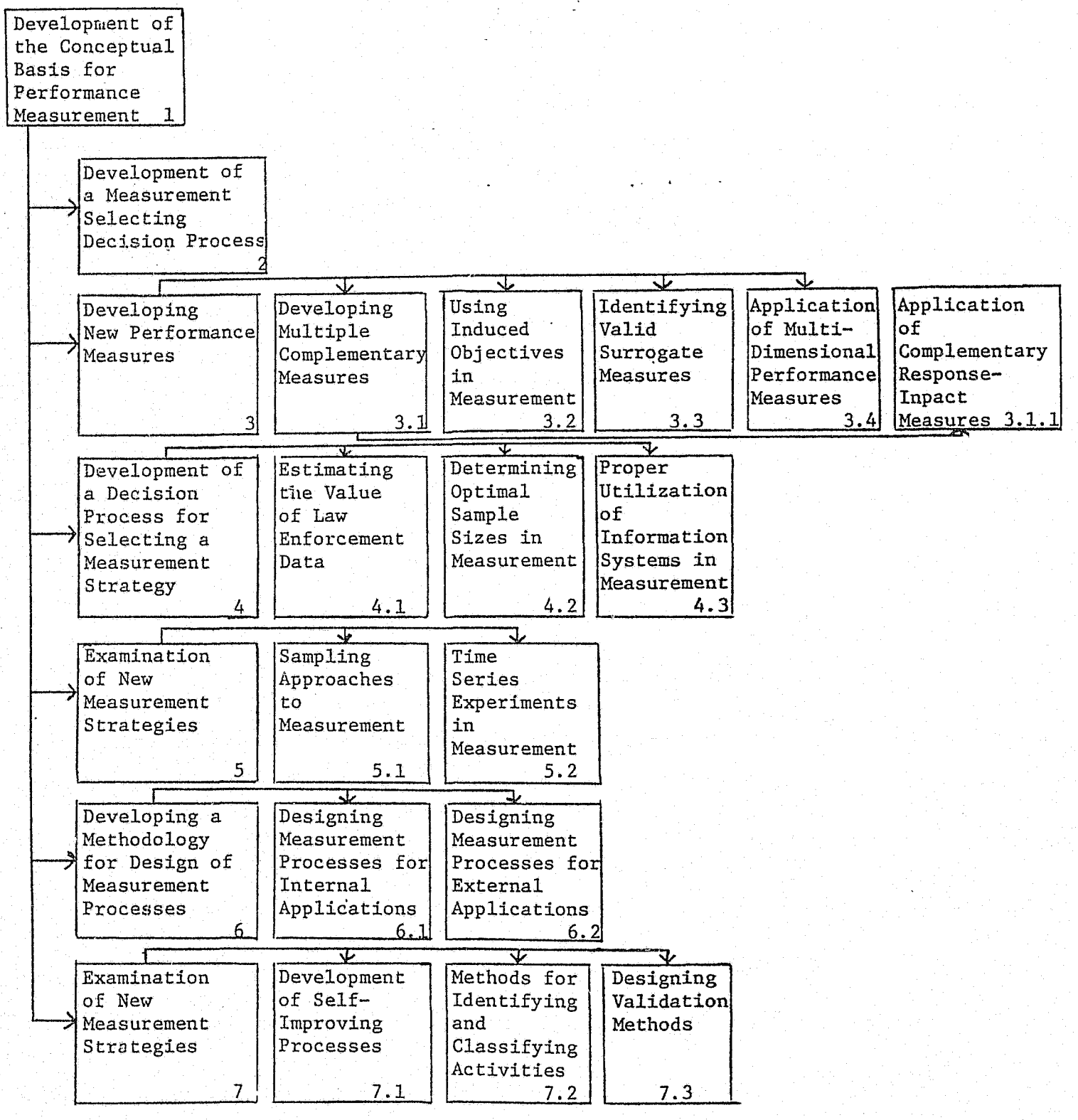


Figure 24
Development of the Conceptual Basis
for Performance Measurement

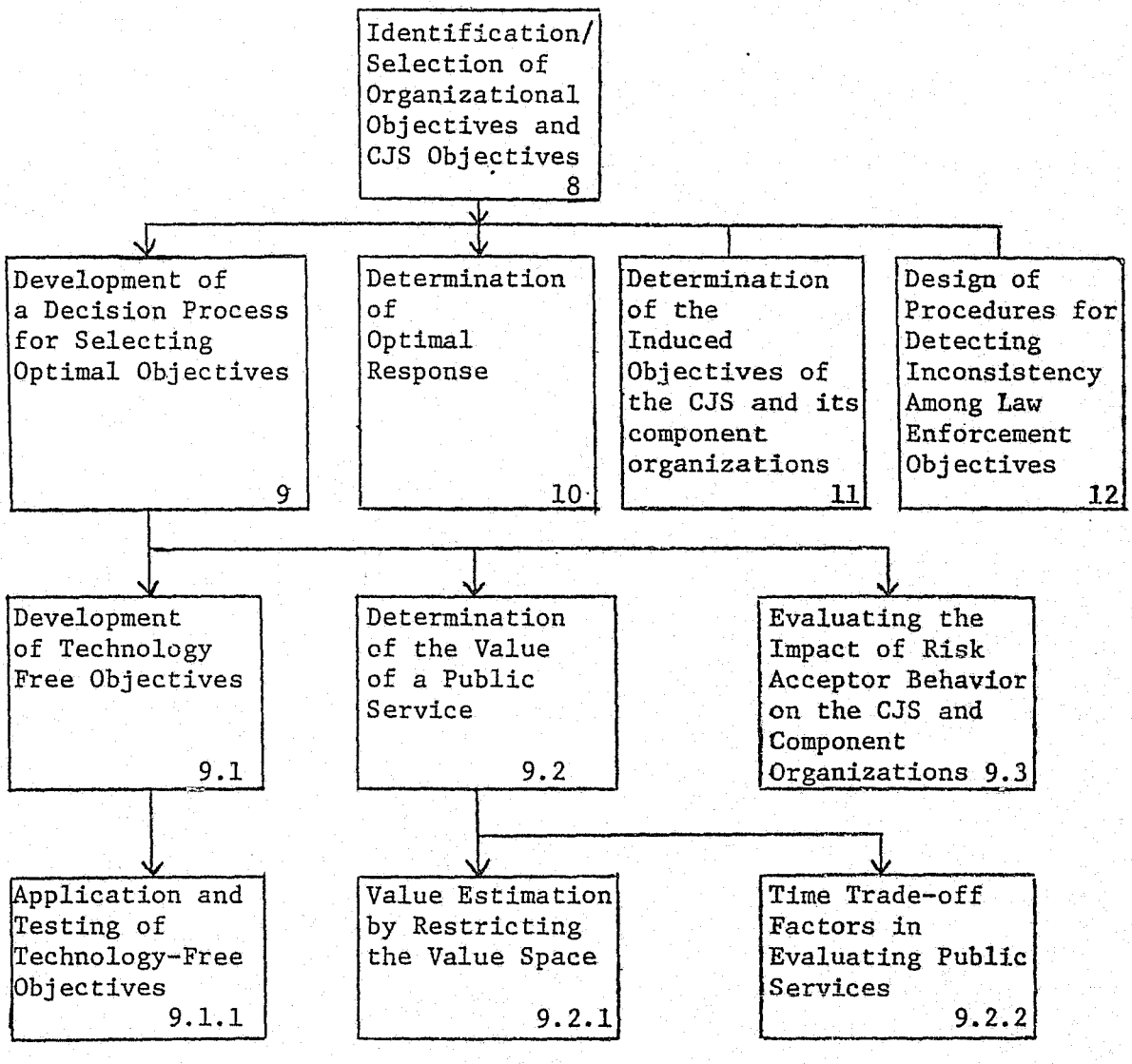


Figure 25

Investigating the Identification and

Selection of CJS and Component Organization Objectives

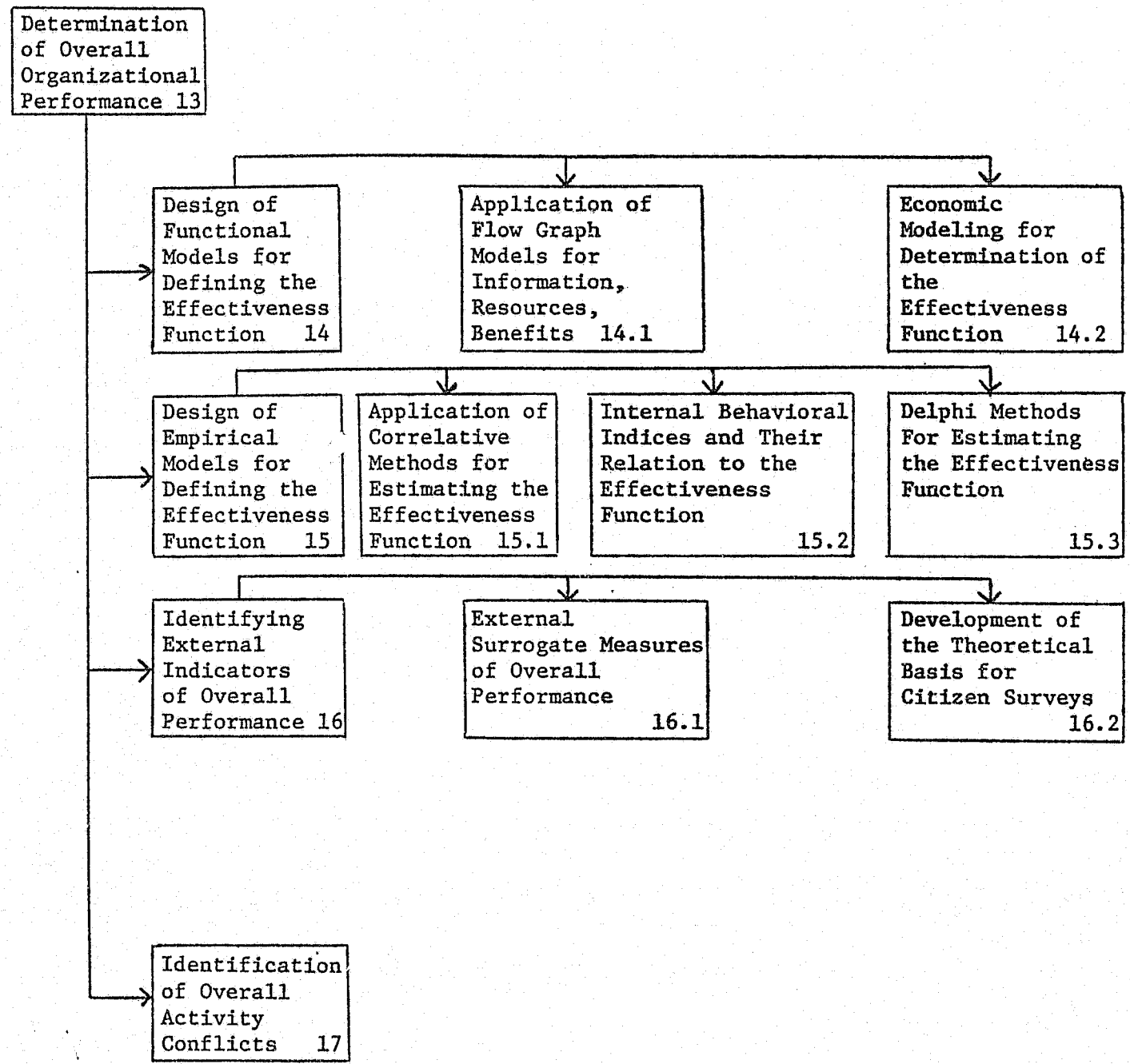


Figure 26

Determination of

Overall Organizational Performance

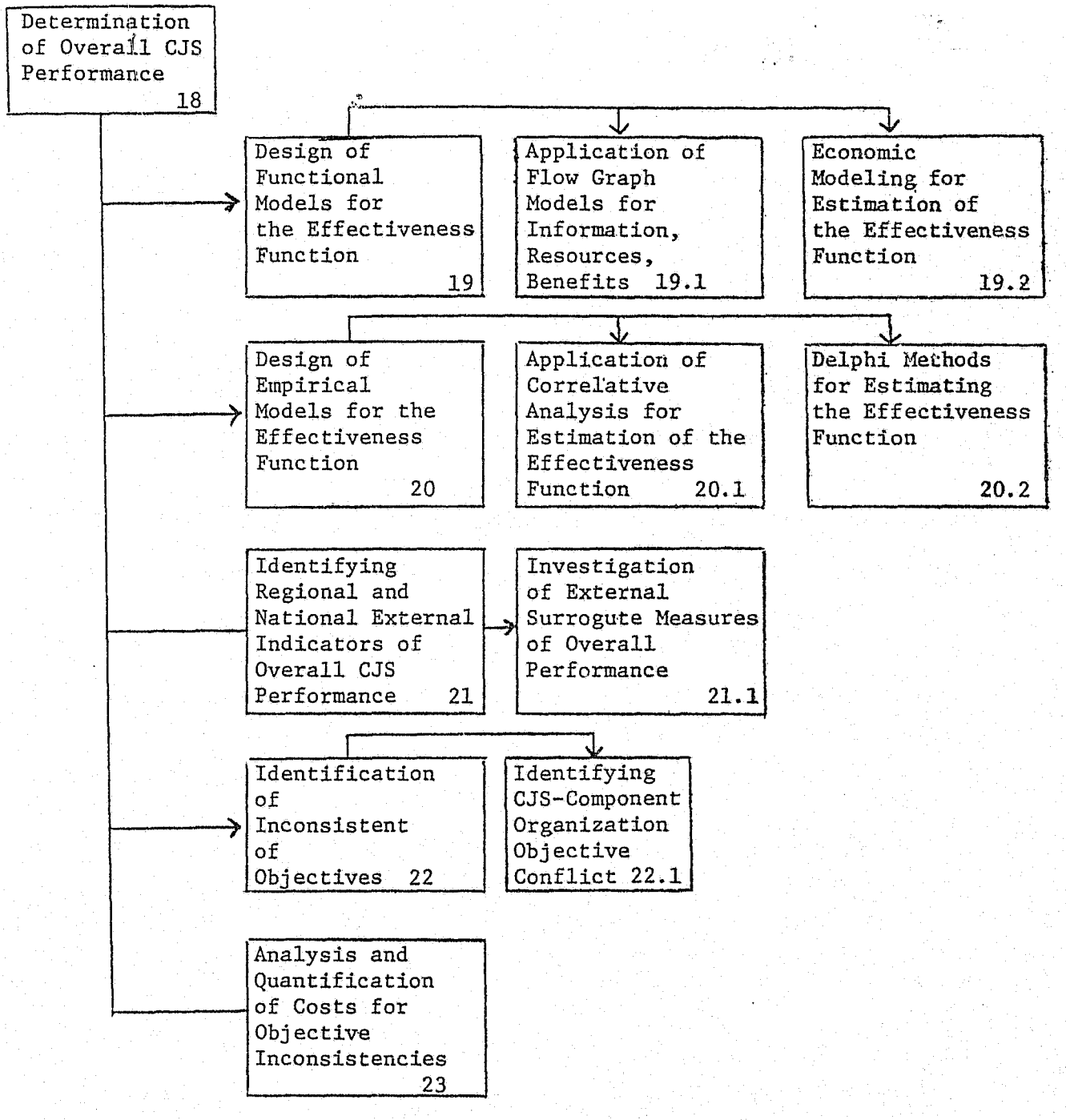


Figure 27

Determination of

Overall CJS Performance

21 major research programs designated along with 36 supportive programs. It is believed that these programs specify and define the scope of new research requirements for performance measurement in the CJS. These 56 programs are divided into four major categories, one category per each of the following figures. These major developmental areas are;

1) Development of the conceptual basis for performance measurement.

2) Identification and selection of Organizational Objectives and CJS Objectives.

3) Determination of overall organizational performance.

4) Determination of overall CJS performance.

It is believed that the most pressing requirements for new understanding of evaluative methodologies lie within these four broad categories, and that these following programs of investigation are the most likely means of of producing better and more powerful measurement approaches, to yield new insights into the operation of the CJS.

Development of the Conceptual Basis for Performance Measurement

As has been previously discussed in this section, the real shortcoming of current evaluative efforts is the lack of a coherent set of measurement principles. The suggested research and development programs identified in the subsequent pages have been designated as means for developing those measurement principles. The programs as shown in Figure 24 are based on various developments presented in this paper.

Program 1. Development of the Conceptual Basis for Performance Measurement

There is no coherent rationale for performance measurement available to those who would seek to evaluate the behavior of the CJS and its component organizations. Presently, there are many diffuse and disparate research efforts which have been able to expose a few of the relevant performance measures, such as those discussed in Section 4, new ways to gather data, such as the victimization surveys discussed in Section 1, and different ways to process the data, such as the efforts of Mantel et. al. [78]. Without doubt, it is time to execute a truly comprehensive study of performance, including both

1) the theoretical basis for measurement, and
2) the practical application of measurement theory.

The study would need to define the relevant variables involved in performance and study the behavior of these variables with relation to the taking of measurements. The real nature of a performance measure would be characterized in behavioral terms. Subsequently, the interactions involved

in the measurement process, such as those between evaluators and the evaluated agency, and a general basis for design of the measurement process, as described in Section 3, would be developed. Finally, a general framework could be structured from the behavioral characteristics of the variables involved in the measurement process. The general framework would encompass the entire concept of measurement: the measure, the strategy, and the process.

Several needed principles will be developed, such as the concept of of measurability. In other words, there will be a definite effort to determine what can be measured, and to what degree of accuracy. Only a few efforts have been made into this area by authors such as Ostrom [35] as discussed in Section 4. Other principles, concerning measures, the measurement strategy, and the measurement process will be set forth for consumption by law enforcement evaluators who need guidelines in their efforts to design measurement systems. This uniform set of principles would be the coherent basis for measurement, expressed in terms that lend themselves to application within the CJS.

Program 2. Development of a Measure-Selecting Decision Process

As described in Section 3, there are an amazing number of possible measures that may be applicable to a given type of behavior. Each measure is best supplied with a different type of data, which means some measures are less expensive to apply than others. As suggested in Section 3, the very first step in designing the measurement approach should be the selection of a measure. As has been identified in the references cited in this section the measures selected by current CJS researchers are not chosen by any particular decision process that assures meaningful measurements at low cost. There is no well developed decision process for choosing measures that are compatible with the measurement strategy and process.

An investigation should be undertaken to develop a decision process for selecting performance measures for law enforcement applications. The process designed should be based on the principles of measurement as determined in Program 1, and should be able to consistently select least cost performance measures that are strategy and process compatible. Such a decision process would necessarily contain some of the features of the prototype model for selection of a measurement approach proposed in Section 4. The decision process defined by the research efforts within the scope of Program 2 would strive toward developing a general structure of the measure selecting decision process, and show how specific features of the process are adapted to different applications.

Program 3. Developing New Performance Measures

Reliance on existing measures of performance has nearly destroyed the value of performance measurement to those within measured agencies, and to those managers responsible for application of the measures. One cause of this is that currently available measures of performance are being applied in situations and to behavior they were not designed to measure. In Section 4, there were several examples given as to how currently available measures are being misused, mostly through the measure used being grossly inappropriate for the behavior measured.

As a necessary step in advancing the value of performance measurement in the CJS, there should be undertaken a program of research to produce new measures that are not available at present. Such a study would incorporate new developments in the policy sciences, behavioral sciences, economics, systems engineering, and other relevant fields. The new measures need to be specified with reference to the behavior they best evaluate, so that they interact effectively with the selection process described in Section 3 and exemplified in Section 7.

Program 3.1 Development of Multiple Complementary Measures

Section 3 described the value of applying multiple measures of performance to a particular observed activity. The basic idea is that no single performance measure completely describes an activity, but that a pair or group of measures that complement each other give a very clear picture of performance. There is currently very little reported work on finding those measures of performance that best complement each other, nor are there many such pairs or groups of measures identified as complementary measures in current use within the CJS.

Given an adequate effort including behavioral modeling, studies into organizational dynamics, and field work, and other areas given in Figure 16, a multiplicity at new measures could be derived. The major requirements for such a study would be that it be comprehensive in scope, in order that none of the new developments in organizational evaluation be overlooked, and that the developed measures be tested in the field for validity.

Program 3.1.1 Application of Complementary Response-Impact Measures

As the description of the parent program, Program 3.1, points out, complementary performance measures can give extended significance to results obtained from performance evaluations of a law enforcement agency. Within this context, it is noted that there is a great deal of difference between evaluating an agency's performance in terms of the response of the agency to various stimuli, and in evaluating that performance in terms of the environment's behavior resulting from that response. The difference can be clearly seen by consulting the behavioral diagrams given in Section 2.

Because these measures refer to the same behavior as the part of the agency, and because they are so radically different, there is a significant synergistic effect involved when the two measures are used simultaneously.

In order to make the most of measurements for CJS agency performance, a program of identification, development, and testing should be undertaken to assess the validity of complementary response-impact measures. The research should also be directed toward developing a method for selecting the proper complementary response-impact measures for a given CJS application, and toward discovering the behavioral principles involved in determining whether or not a particular application will yield complementary response-impact measures.

Program 3.2 Using Induced Objectives in Measurement

There are likely to be many evaluative applications in the CJS for the induced objective concept as described in Section 2. Indeed, the very definition of the induced objective relates it to behavior that can be observed, and therefore measured. With regard to effectiveness measurement, which is one of the main topics of this paper, identifying induced objective based upon measurement will be of particular value.

There are two justifications behind the identification of induced objectives and the use of these objectives in effectiveness measurement. First, the mere identification of induced objectives provides telling evidence about the agency's performance and its enviornment. Evaluators can compare induced objectives to formal objectives to point out any variance between desired results and actual input. Secondly, evaluators may be able to base effectiveness measurement on induced objectives where there is reason to believe the stated objectives infeasible. A program of investigation to develop induced objective based measures of performance should be begun to find both the principles behind the measures and to identify applications in the CJS for which these measures are best suited.

Program 3.3 Identifying Valid Surrogate Measures

It is known that many aspects of an organization's performance are quite difficult or costly to measure directly, and that, in these cases, surrogate or "proxy" measures, as described in Section 3, may yield significant and representative results at reasonable cost and effort. Several authors such as Fisk and Winnie [37], Hirsch and Riccio [16], and Krug [4] mentioned the use of these measures, yet there is little available in the way of a consistent unified theory behind their use, particularly with regard to the principles of measurement in the CJS as described in relation to Program 1.

From correlative studies and other analysis of investigations involving field observations a general method for selecting surrogate measures could be developed. In addition, a multiplicity of such measures useful in CJS applications would be developed, applied, and evaluated to ensure that the general method for selecting surrogate measures is valid. One of the most obvious areas of application of any measures developed would be in testing effectiveness for some of the various law enforcement activities which, when successfully undertaken, are thought to result in no criminal behavior on the part of the environment. That is the developed surrogate measures might first be tested in detecting those environmental changes that occur in response to deterrence programs.

Program 3.4 Application of Multidimensional Performance Measures

Arguments in Section 4 suggest that there are many situations where performance is not adequately represented by a single measure. In almost all cases, it is highly questionable to compress different measures into a single common denominator such as dollar value. For this reason, it is often quite useful to employ a multidimensional measure of effectiveness to

properly represent the observed behavior. Unfortunately, most existing research and most applications seek only to measure output, efficiency, and some measure of quality, which is directly converted to dollars on a utility value. This points out the need for the development of multidimensional measures of performance for CJS applications. The study indicated would necessarily attempt to develop both the procedure for constructing the measure and the methodology of using the measure for evaluative purposes.

Such a multidimensional measure of performance might incorporate several groups of complementary measures, as owuld be developed in Program 3.1. The measures would no doubt be varied in orientation, but all chosen to describe the same behavior. It is clear that such measures can be chosen, but another prupose of this program of research is to show how multidimensional measures can be applied in evaluating CJS performance. There are many promising possibilities here, including the performance profile approach discussed in Seciton 4. What remains is to establish, through investigation, however, is how these measures are to be applied to evaluation programs in the CJS.

Program 4. Development of a Decision Process for Selecting a Measurement Strategy

Heretofore, much effort has been expended debating which source of data best suits CJS measurement needs. Little attention has been given to the development of an overall measurement strategy, such as is defined in Section 3. There is a pressing need for a research effort directed toward charting the relationships between behavior and the measurement strategies best suited for performance measurement of that behavior. In addition such a research effort should strive to build a decision process for selecting measurement strategies. Availability of such a process would make it

possible for managers and planners to avoid selecting deficient or expensive strategies when better strategies are available.

Program 4.1 Estimating the Value of Law Enforcement Data

There is a limit to the value of data. This fact seems to have been overlooked in the planning of new data centers for criminal information. As discussed in Section 3, the value of compiling additional data depends on the marginal advance toward achieving organizational objectives that would be made possible by the additional information.

There is justification for commencing a study into determination of the value of crime-related data. The study should encompass both determination of how an organization in the CJS uses information and what that information costs. Given this knowledge, some quantification of the marginal advantage gained by collection of additional crime-related data by a CJS organization or group of organizations can be developed.

Program 4.2 Determination of Sample Sizes in Performance Measurement for CJS Applications

The art of sampling is well developed and there is a sorrel statistical foundation for selecting sample sizes that minimize cost and maintain or maximize accuracy. Unfortunately, there seems to be a marked disregard of this well developed science, which has been utilized successfully by private industry for a number of years, by those involved in performance measurement in the CJS. As was pointed out in Section 3, measurement can be equated to sampling, so that the determination of measurement effort involves setting sample sizes. This program, a twin to Program 4.1, is proposed as a means of modifying existing statistical procedures and developing new procedures for selecting the optimal quantity of data in a CJS evaluation.

Program 4.3 Proper Utilization of Information Systems in Measurement

When any new technology meets its first application in an operational environment, there is a preliminary wave of proposals for utilizing the technology in novel ways. The schemes envisioned are usually either too superficial and/or too simple to be of real value in the operational environment, or too massive and costly that, although begun in good faith, the attempted application results in a frivolous waste of resources. Now that computer technology has begun to meet its first successful applications, it is time to pull together all that is known about the interface between a data handler and data user and produce a coherent set of principles to both,

1) govern the current use of computers in the CJS, and
2) direct new attempted applications.

Program 5. Examination of New Measurement Strategies

Without a doubt there are an enormous number of promising new measurement strategies that, although available in theory, have not been applied to performance measurement within the CJS. This is because either there is no historical basis for their use in CJS applications or because, as explained in Section 3, researchers have been so awed by the problem of selecting a source of data that they have overlooked the other essential components of a measurement strategy such as the time span the data is to cover, the time span of the collection effort, and the quantity of data to be gathered. Indeed, the fact that data costs money is ignored in a large proportion of existing evaluative efforts.

It is proposed that a search for new measurement strategies be initiated, for purposes of,

1) identifying a general classification scheme for all measurement strategies,

2) surveying measurement strategies employed in non-CJS applications, and

3) applying the principles of measurement developed in Program 1 and the information gained in 2) above to design a multiplicity of new and useful measurement strategies for each of the categories defined by the classification scheme developed in 1) above.

Once these strategies are developed and tested, their characteristics, such as cost to operate, will be known. Then the strategies will be available for selection by the decision process of Program 4 for an individual application in the CJS.

Program 5.1 Sampling Approaches to Measurement

Section 3 equates many of the basics of measurement to sampling. This suggests that many of the principles of sampling correlate with or should be incorporated with the principles of measurement discussed in relation to Program 1. Program 5.1 would seek to determine the involvement of sampling principles in measurement, and then proceed to develop a number of sampling-oriented measurement strategies. These strategies would be designed such that the basic concept of sampling, selecting a representative subset of all available data, directs their execution.

Program 5.2 Time Series Experiments in Measurement

The role of time series experiments in CJS performance measurements has not been established. What has been shown is that time series methods stand to make a significant contribution to the performance measurement abilities of those who would seek to evaluate CJS organizations. In Section 3, one method reported by Ross [18] proved to be a highly successful means for assessing the impact of a policy change. This points out the need to develop new time series measurement strategies which would provide a new approach to performance measurement, particularly in conjunction with the complementary response-impact measures to be developed in Program 3.1.1.

Program 6. Developoing a Methodology for Design of Measurement Process

Section 3 offered a prototypical model to the measurement process, and therein the basic components of the process were identified. Furthermore, it has been shown that evaluators, in their haste to find performance measues and select data sources, have virtually ignored the larger consideration of designing the measurement process to guide the application of the performance measure through the measurement strategy. There has been little effort made to apply modern systems thinking to the performance measurement process in the CJS.

The purposes of Program 6 are to explore the factors that determine what composes a best measurement process, and to develop a decision process for CJS evaluators to apply in selecting such a measurement process. The program would necessarily consider the different process requirements of the CJS activities where measurement is applied. Additionally, the decision process developed would be structured to allow the flexibility necessary to accommodate the great variety of CJS applications.

Program 6.1 Designing Measurement Processes for Internal Applications

There is a significant difference between the requirements between measurements carried out for internal control and those performance measurements executed for the benefit of elements external to the organization. This difference is reflected in the six-step prototypical measurement process given in Section 3. Program 6.1 is proposed as a vehicle for recognizing and analyzing the important characteristics of an internal evaluation, and developing from the analysis a comprehensive enumeration of the effects of the existence of these distinguishing characteristics on the structure and operation of a well-designed measurement process. This enumeration would be of significant value in adapting the general measurement process to a particular internal application.

Program 6.2 Designing Measurement Processes for External Applications

Program 6.2 is a logical twin to Program 6.1. The purpose of this twin is to develop a complete enumeration and evaluation of the characteristic make up of the well-designed measurement process for external applications. This means that the program is intended as a means of discovering the particular features that must be present in a well-designed measurement process intended for use by evaluators external to the CJS component organization. The program would produce results identical in format but distinct in content from Program 6.1. The results would prove extremely useful to those external to organization who would be seeking to apply the procedure for selecting a measurement process designed in Program 6.

Program 7 Examination of New Measurement Processes

Section 3 makes reference to only one report about measurement processes for the CJS. The report by Caputo [21] shows a procedure similar to the six-step prototype given in Section 3 as a model for a measurement process. There are, however, new evaluative processes being developed for other applications. It is certain that, given a reasonable survey and analysis of these other new evaluative processes, a number of new measurement processes suitable for CJS applications could be synthesized, particularly for application at the organizational level. This then, is the objective of Program 7.

Program 7.1 Developing Self-Improving Processes

The objective of this program is to develop an understanding of the elements of a measurement process that influence its operation in such a way that the process is self-improving. This amounts to identifying those elements that, when included in the measurement process (as illustrated

in Section 3), contribute to a gradual refinement of the structure and operation of the process os that more meaningful measurement results are obtained at lower cost and effort. Such elements may be as simple as a provision for continual reassessment of data analysis methods, or as complex as a scheme for rotating the membership of the group of evaluators. The only requirement made on an element is that it need contribute to the proper operation of the measurement process over time as does feedback in a control system.

Program 7.2 Methods for Identifying and Classifying CJS Activities

One of the basic premises of performance measurement is that the behavior in question can be identified. This is the first step in the model measurement process given in Section 3. The basic informational requirements for classifying CJS behavior are discussed at length in Section 5, and Section 7 demonstrates how this classification is used for selecting an appropriate measurement approach.

It is the task of Program 7.2 to develop a number of methodologies for the identification and classification of law enforcement activities. Such methodologies are to produce classifications that allow for and ensure the selection of the measurement approach most appropriate to the behavior being evaluated. In addition, the methodologies need be universal in scope so that all CJS activities can be classified by functional, rather than jurisdictional, characteristics, because, as been previously demonstrated, many measurement approaches are activity-specific. This means simply that different types of behavior call for different types of measurement approaches.

Program 7.3 Designing Validation Techniques

The objective of Program 7.3 is to determine the nature of a best validation technique for a given measurement approach. As discussed in

Section 7, validating the results obtained through one measurement approach are usually most thoroughly validated through application of some other measurement approach. This is not always the case, however. The successful completion of Program 7.3 should yield a number of different validation techniques. Each technique would be identified by the nature of the measurement situation to which it applied. These techniques could then be used to verify results obtained through application of a measurement approach to a CJS application.

Identifying and Selecting Organizational Objectives and CJS Objectives

A second family of related research programs is discussed here. The real justification for this group of programs lies in the confusion of infeasible and conflicting objectives pursued by CJS component agencies, and in the lack of a uniform and coherent set of system-wide objectives visible to the public. The notion of the induced objectives, as discussed in Section 2, is a central element in this family of programs. This family of programs is given before the remaining two groups of research programs because they depend on some identification of objectives being made. The necessity of identifying or selecting objectives prior to measurement is central to the measurement process as discussed in Section 3 and then examplified in Section 5.

Program 8 Identificantion and Selection of Organizational Objectives and CJS Objectives

The primary objective of this program is the development of a classification scheme for describing both organizational objectives and CJS objectives. The notation to be developed would be capable of demonstrating the type of objective, and its relationship to organizational goals, other organizational objectives, and system-wide objectives. The application of the classification would lie in its use in assisting organizations with

law enforcement goals to select appropriate objectives, particularly through selection procedures to be developed in subsequent programs.

Program 9. Development of a Decision Process for Selecting Optimal Objectives

Section 4 details a few of the ways that law enforcement organizations can fail in their selection of objectives. That this occurs is not surprising, particularly in light of the limitations current technology places on the organization's perception of its objectives, as described in Section 2. Program 9 is proposed as a means for developing a logical decision process for use by CJS component organizations in selecting objectives that, in light of the capabilities of the organization, the nature of its environment and the system-wide objectives, will most nearly bring about the set of target conditions described in the goals of the organization and of the CJS. Furthermore, a process for selecting system-wide objectives for the CJS would be developed, as a complement to the organizational objective-selection process.

Program 9.1. Development of Technology-Free Objectives

Program 9.1 is an enumeration and analysis of specific adaptations required in the decision process for selecting optimal objectives developed in Program 9 so as to force the decision process to produce objectives that are essentially "technology-free." That is, Program 9.1 will explore methods for removing the perception-limiting effects of "current technology" from the objective selection process, and to discover how to further promote the selection of objectives whose achievement is relatively independent of the technology of the organization.

Program 9.1.1 Application and Testing of Technology-Free Objectives

Program 9.1.1 involves supplying technology free objectives developed by the decision process adapted in Program 9.1 to a test group of CJS component organizations. The utility of the objectives, which would be

apparent after a test period, would then reflect on the adequacy of the adapted decision process.

Program 9.2 Determination of the Value of a Public Service

Section 2 points out the desirability of having some estimate of the value of a public service, particularly if one is designing a decision process for selecting objectives which may affect the degree or extent to which each of a number of services is provided. Program 9.2 is designed to develop a general framework for attaching some measure of value or utility to the services provided by CJS component organizations.

Program 9.2.1 Value Estimation by Restricting the Value Space

One promising technique for estimating the value of a law enforcement service is the method by which successively tighter upper and lower bounds are placed on the scalor region on which value is defined. Once a sufficient number of iterations have been made, or a sufficient number of bounds been applied, the result is an interval estimate of the value of the service in question. Program 9.2.1. is directed toward developing this philosophy and preparing a useful methodology, based on the concept, for applications within the CJS.

Program 9.2.2 Time Trade-Off Factors in Evaluation of Public Services

The consideration of the value of a public service in the setting of objectives calls for the recognition of time dynamicity. The discounting of economic benefits is an accepted practice, yet little work has been done to establish factors to express the increased utility of receiving more immediate benefits. Program 2.2.2 is proposed as a means for laying the groundwork for use of time trade-off factors in evaluation of public services, and providing necessary information for Program 9.

Program 9.3 Evaluating the Impact of Risk Acceptor Behavior on the CJS and Its Component Organizations

Section 1 described a class of criminals who either block out thoughts of risks associated with crime or choose to accept these risks, regardless of their magnitude. Program 9.3 proposes to investigate the particular effect the existence of this particular type of criminal has on the determination of optimal law enforcement objectives.

Program 10 Determination of Optimal Response

Program 10 is proposed as an exploration of the concept of optimal response as it applies to organizational behavior, particularly with regard to the setting of new objectives and designation of activities to support those objectives as a response to changes in the environment. The objective of the project would be to develop a set of principles or relations that would serve a CJS organization in determining how to adapt to changing conditions that may have destroyed the optimality of its original objectives.

Program 11 Determination of the Induced Objectives of the CJS and Its Component Organizations

The concept of induced objectives was described in Section 2, and subsequently it was suggested that determining the induced objectives of an organization would yield a myriad of beneficial analytical information. Program 11 is directed at developing from available data a set of induced objectives for the CJS and for a selected group of its member organizations. A general technique for identifying the induced objectives of a law enforcement organization would also be developed as a logical prerequisite for execution of the other parts of the program.

Program 12 Design of Procedures for Detecting Inconsistency Among Law Enforcement Objectives

The nature of conflicts among law enforcement objectives has been discussed in Section 4, and it has been established that individual CJS agencies select objectives that best suit their own needs, regardless of system needs. Program 12 would be directed toward producing a set of procedures for making a simple determination as to the consistency of a set of organizational objectives. An effort would be made to package the procedures in such a manner as to be immediately available to CJS managers and evaluators.

Determination of Overall Organizational Performance

As discussed in Section 4, the overall performance of an organization depends on how well it performs each of its activities. Programs in this family attempt to develop the basis for an overall measure of performance.

Program 13 Determination of Overall Organizational Performance

Program 13 will seek to integrate the concepts of Section 4 and to develop an overall set of principles for relating performance levels for the individual activities of an organization to an overall measure of performance.

Program 14 Design of Functional Models for Defining the Effectiveness Function

Under this program, functional models relating flow of resources, information, and casework through a CJS agency and its environment to overall organizational performance would be developed. As described in Section 4, achievement of individual objectives would be expressed in terms of flow so that their contribution to overall performance could be quantified.

Program 14.1 Application of Flow Graph Models for Information, Resources, Benefits

Flow graphs were shown to be particularly useful in determining the relationships between variables in complex feedback models such as for those shown in Section 5. Program 14.1 is directed at developing such models for use in determining the effectiveness function. Such developed models might then serve Program 14.

Program 14.2 Economic Modelling for the Determination of the Effectiveness Function

Program 14.2 represents a second avenue of analysis for estimating the effectiveness function. There are a wide variety of economic interpretations that can be made for the attainment of various organizational objectives, and it is the objective of this program to incorporate these economic effects into an overall model of effectiveness.

Program 15 Design of Empirical Models for Defining the Effectiveness Function

As described in Section 4, it is possible to develop empirical models for relating the effectiveness of an agency in its individual activities to overall effectiveness. This program would develop the principles behind this empirical modeling and test validity of the empirical approach.

Program 15.1 Application of Correlative Methods for Estimating the Effectiveness Function

Program 15.1 would seek to apply several of the currently available correlative methods to the building of an empirical model for estimating the effectiveness function. The perceptions of evaluators and other indicators of overall effectiveness would be correlated with the effectiveness of an agency in achieving each of its objectives.

Program 15.2 Internal Behavioral Indices and Their Relation to the Effectiveness Function

This program would seek to identify some of the observable behavior internal to the organization that can be correlated with its overall effectiveness. Among these are such indicators as "resolution of conflict" and "distribution of authority." These indicators would be fitted into an overall effectiveness function of the type described in Section 4.

Program 15.3 Delphi Methods for Estimating the Effectiveness Function

The empirical modeling approach described in Section 4 as a means for determining the effectiveness function requires several subjective evaluations of overall performance. This program is proposed as a means of testing the applicability of Delphi approaches to the problem of consolidating the subjective performance evaluations.

Program 16. Identifying External Indicators of Overall Performance

Section 3 points out the existence of measures of performance based on the effects on the environment of the organization's operation. This program would seek to develop the notion of measuring overall performance in terms of these environmental effects.

Program 16.1 External Surrogate Measures of Overall Performance

This study would seek to identify external surrogate measures, that is, an attempt would be made to identify and measure the indirect effects of an agency's operations as a means of estimating overall performance. Surrogate measures, as described in Section 3, would be used for the instances where the impact of certain aspects of organizational behavior cannot be directly measured.

Program 16.2 Development of the Theoretical Basis for Citizen Surveys

Section 1 established that there is currently great interest shown in citizen surveys as data collection means. This program is designed to identify the basic concepts behind this data gathering technique, and to build a behavioral model to enable evaluators to judge the accuracy of their results and the limitations on the meaning of data from such surveys.

Program 17 Identification of Objective-Activity Conflicts

This program is similar to Program 12, except that the emphasis is on the activity-objective interaction. There would be a general method for detecting such conflicts developed and some theoretical adaptations offered for inclusion in the models of overall performance. Several field investigations would be relied upon for data relating conflict to overall performance.

Determination of Overall CJS Performance

This final family of development efforts is undertaken for purposes of describing the relationship between the achievement of organizational objectives and the achievement of CJS objectives as discussed in Section 4. This description could then serve for purposes of allocating resources among the various CJS component organizations and for identifying conflicting sets of objectives. In addition, information about the contribution that individual organizations make to overall performance will assist in evaluating proposed policy changes.

Program 18 Determination of Overall CJS Performance

This program is directed at developing the general principles underlying system-wide performance measurement, and to analyze several models

that relate component organization performance to overall performance. The objective is to select a model that most closely corresponds with measurement principles and yields the most accurate results in its application.

Program 19 Design of Functional Models for the Effectiveness Function

Section 4 points out the concept of using functional models as an approach to estimating an effectiveness function which relates the achievement of individual objectives to overall performance. This study program is intended to develop models similar to those discussed in Program 14, only adapted to consider the CJS as the organization, with its objectives being those of its component organizations.

Program 19.1 Application of Flow Graph Models for Information, Resources, and Benefits

This program serves to extend the results of Program 14.1 to cover the entire CJS. Flow graph models would be devised to account for information, resource, and benefit flows both within component organizations and between them. The objective is to develop an adequate definition of the effectiveness function for the entire CJS.

Program 19.2 Economic Modeling for Estimation of the Effectiveness Function

This program is proposed as a means for extending the models obtained in Program 14.2 so that the entire CJS falls within the scope of an economic model. The objective is to develop an effectiveness function from the economic considerations involved in the operation of the CJS.

Program 20 Design of Empirical Models for the Effectiveness Function

This program adapts the models developed in Program 15 so that the overall performance of the CJS can be considered. Again, the objective is to develop a relationship between performance at the component level and overall CJS performance.

Program 20.1 Application of Correlative Analysis for Estimation of the Effectiveness Function

As an extension of the results of Program 15.1, this program would seek to develop the empirical relationship between component organization performance and overall CJS performance. The program would rely upon correlative analysis as a means of describing the relationship. This description would then serve as an empirically-derived effectiveness function.

Program 20.2 Delphi Methods for Estimating the Effectiveness Function

This program, which is related to Program 15.3, is proposed as means for exploring the possible application of the Delphi Method and similar methods in making an empirical description of the effectiveness function for the CJS. Several such techniques would be tested for use in making subjective evaluations of overall performance.

Program 21. Identifying Regional and National External Indicators of CJS Performance

This program is directed at identifying the direct impacts of CJS activities and deriving measures of overall performance based on these impacts. This is a logical extension of Program 16 from the organizational level to the regional and national levels.

Program 21.1 Investigation of External Surrogate Measures of Overall Performance

This program is directed at developing measures of overall CJS performance, such measures being based on the more indirect impacts of CJS activities. This program is a logical extension of Program 16.1 and the concept of surrogate measures presented in Section 3.

Program 22 Identification of Inconsistent Objectives

This program extends the results of Program 12 to handle system-wide objectives. This program is a necessary step in the development at an overall effectiveness function for the CJS. The execution of the program is required to lend measuring to any overall performance measurements made through application of methods developed in preceding programs. The importance of this program is clearly established by the arguments regarding inconsistent objectives and performance measurement given in Section 4. The completion of this program would yield a number of principles regarding identifying inconsistency among CJS objectives, and a procedure for applying these principles in the detection of inconsistencies.

Program 22.1 Identifying CJS-Component Organization Objective Conflicts

This program seeks to make application of the procedures for detection of inconsistency of objectives developed in Program 22 and information about CJS objectives developed in Program 8. An attempt would be made to identify a number of existing conflicts between system-wide objectives and component organization objectives as a method for testing the procedures developed in Program 22.

Program 23 Analysis and Quantification of Costs for Objective Inconsistencies

The purpose of this program is to attach a cost to the relative lack of coordination among CJS component organizations. This lack of a chain of command, as described in Section 5, ensures that conflicting objectives are designated. The quantification of the costs of the conflicts, as developed by this program, would certainly prove useful in both the determination of overall CJS performance, Program 18, and in the selection of CJS objectives, Program 8. The program would be necessarily broad enough

in scope to consider both CJS-component organization objective conflicts and objective conflicts between component organizations. The costs developed may be expressed in economic or non-economic terms, so long as the overall magnitude of the problem is truly reflected. The basis for these quantifications will be the functional interrelationships of the component organizations, particularly those considered in Program 19.

Summary of Research Needs

In short, there is a profound need for the development of the concepts of measurement strategies and measurement processes for CJS applications. Additionally, there must be some effort made to develop models relating effectiveness achieved in relation to individual objectives to overall organizational effectiveness. Finally, the process by which CJS objectives are chosen must be studied to improve the overall performance of the CJS and to eliminate many of the existing conflicts.

It is believed that the programs of research and development which have been generally outlined in this section represent the best means for meeting measurement needs in the CJS. The programs may prove to be the most effective way of preparing to bring performance measurement in the CJS up to date by introducing some of the newer technologies available for performance measurement and by giving the entire measurement concept a thorough analysis. It is expected that the research can lead to new and more useful measurement approaches that depart from the rather superficial treatment previously afforded CJS applications.

LIST OF REFERENCES

Text Ref. No.		
(6), (9)	1.	Avi - Itzhak, Benjamin, and Shinnar, Revel. "Quantitative Models in Crime Control." Journal of Criminal Justice, 1 (Fall, 1973), 185-217.
(59)	2.	Beattie, R. H. "Criminal Statistics in the United States." Journal of Criminal Law, Criminology, and Police Science 51 (May, 1960), 49-51.
(26)	3.	Becker, Gary S. "Crime and Punishment: An Economic Approach." Journal of Political Economy, 76 (March, 1968), 169-217.
(5)	4.	Belkin, Jacob, Blumstein, Alfred, and Glass, William. "Recidivism as a Feedback Process: An Analytical Model and Empirical Validation." Journal of Criminal Justice, 1 (March, 1973), 7-26.
(50)	5.	Berk, Richard A. "Performance Measures: Half Full or Half Empty?" Social Science Quarterly, 54 (March, 1974), 762-764.
(12)	6.	Blumstein, Alfred, and Larson, Richard. "Models of a Total Criminal Justice System." Operations Research 17 (March, 1969), 199-232.
(21)	7.	Caputo, David C. "Evaluating Urban Public Police: A Developmental Model and Some Reservations." Public Administration Review, 33 (March, 1973), 113-19.
(13)	8.	Cho, Yong Hyo. "A Multiple Regression Model for the Measurement of the Public Policy Impact on Big City Crime." Policy Sciences, 3 (1972) 435-55.
(53)	9.	Conrad, John. "Law, Order, and Corrections." Public Administration Review, 31 (November, 1971), 596-602.
(25)	10.	Ehrlich, Issac. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." Journal of Political Economy, 81 (May, 1973), 521-65.
(37)	11.	Fisk, Donald M., and Winnie, Richard E. "Output Measurement in Urban Government: Current Status and Likely Projects." Social Science Quarterly, 54 (March, 1974), 725-40.
(74)	12.	Grant, Eugene L., and Ireson, W. Grant. Principles of Engineering Economy. 5th ed. New York: Ronald Press, 1970.
(23)	13.	Harris, John R. "On the Economics of Law and Order." Journal of Political Economy, 78 (May, 1970), p. 165.

(79) 14. Hatry, Harry P. "Measuring the Effectiveness of Nondefense Public Programs." Operations Research, 18 (September, 1970), 772-84.

(16) 15. Hirsch, Gary B., and Riccio, Lucius J. "Measuring and Improving the Productivity of Police Patrol." Journal of Police Science and Administration, 2 (1974), 169-84.

(8) 16. Holzer, Marc. "Police Productivity: A Conceptual Framework for Measurement and Improvement." Journal of Police Science and Administration, 1 (December, 1973), 459-67.

(24) 17. Ireland, T. R. "Optimal Enforcement of Laws." Journal of Political Economy, 80 (March, 1972), p. 421.

(17) 18. Jones, E. Terrence. "Evaluating Everyday Policies - Police Activities and Crime Incidence." Urban Affairs Quarterly, 8 (March, 1973), 267-79.

(81) 19. Keller, Fred S. Learning: Reinforcement Theory. New York: Random House, 1969.

(80) 20. Kuo, Benjamin C. Linear Networks and Systems. New York: McGraw-Hill, 1969.

(2) 21. Larson, Richard C. "On the Modeling of Police Patrol Operations." IEEE Transactions on Systems Science and Cybernetics, 6 (October, 1970), 276-81.

(19) 22. Logan, Charles H. "General Deterrent Effects of Imprisonment." Social Forces, 51 (September, 1972), 64-73.

(78) 23. Mantel, Samuel J. Jr.; Service, Allan L.; Reisman, Arnold; Koleski, Raymond A.; Blum, Arthur; Dean, Burton V.; Reich, Robert; Jaffee, Miriam; Rieger, Howard; Ronis, Richard; and Rubinstein, Judah. "A Social Service Measurement Model." Operations Research, 23 (March, 1975), 218-39.

(82) 24. Mason, S. J. "Feedback Theory: Some Properties of Signal Flow Graphs." Proc. IRE, 41 (September, 1953), 1144-56.

(39) 25. Nanus, Burt, and Perry, Luther. "A Planning Oriented Measure of Crime and Delinguency." Journal of Criminal Justice, 1 (Fall, 1973), 259-63.

(52), (51) 26. New York City - Rand Institute. "Research in 1970-1971." Operations Research, 20 (1972), 474-515.

(35) 27. Ostrom, Elinor. "Exclusion, Choice and Divisibility: Factors Affecting the Measurement of Urban Agency Output and Impact." Social Science: Quarterly, 54 (March, 1974), 691-99.

(14) 28. Ostrom, Elinor. "On the Meaning and Measurement of Output and Efficiency in the Provision of Urban Police Services." Journal of Criminal Justice, 1 (Summer, 1973), 93-112.

(27) 29. Ostrom, Elinor. "Metropolitan Reform: Propositions Derived From Two Traditions." Social Sciences Quarterly, 53 (December, 1973), 474-493.

(54) 30. Ostrom, Elinor, and Whitaker, Gordon. "Does Local Community Control of Police Make a Difference? Some Preliminary Findings." American Journal of Political Science, 17 (1973), 48-76.

(7) 31. Rardin, Ronald L., and Gray, Paul. "Analysis of Crime Control Strategies." Journal of Criminal Justice, 1 (Winter, 1973) 339-46.

(56) 32. Reppetto, Thomas A. "Crime Control Management and the Police." Sloan Management Review, 14 (Winter, 1972-72), 45-54.

(18) 33. Ross, H. Laurence; Cambell, Donald T.; and Glass, Gene V. "Determining the Social Effects of a Legal Reform." American Behavioral Scientist, 13 (March, 1970), 493-509.

(60) 34. Sellin, T., and Wolfgang, M. E. Delinquency in a Birth Cohort. Chicago: University of Chicago Press, 1972.

(40) 35. Skogan, Wesley, G. "The Validity of Official Crime Statistics: An Empirical Investigation." Social Science Quarterly, 55 (June, 1974), 25-38.

(61) 36. Stigler, George. "Optimal Enforcement of Laws." Journal of Political Economy, 78 (May, 1970), 526-36.

(10) 37. Swimmer, Gene. "The Relationship of Police and Crime - Some Methodologies and Empirical Results." Criminology, 12 (November, 1974), 293-314.

(20) 38. Tittle, Charles R., and Rowe, Alan R. "Certainty of Arrest and Crime Rates: A Further Test of the Deterrence Hypothesis." Social Forces, 52 (June, 1974) 455-62.

(76) 39. U. S. Congress. Joint Economic Committee. Project Design and Evaluation with Multiple Objectives, by A. Myriek Freeman III, Joint Committee Print. The Analysis of Public Expenditures, Vol. 1, (Washington, D. C.: Government Printing Office, 1969), pp. 565-91.

(55) 40. Votey, Harold L. Jr., and Phillips, Llad. "Social Goals and Appropriate Police for Corrections: An Economic Appraisal." Journal of Criminal Justice, 1 (Fall, 1973), 219-40.

(77) 41. Weissenberg, Peter. Introduction to Organizational Behavior. Scranton, Pa.: Intext Educational Publishers, 1971.

(9) 42. Wellford, Charles R. "Crime and Police - A Multivariate Analysis." Criminology, 12 (August, 1974), 195-211.

(49) 43. Whitaker, Gordon P. "Who Puts the Value in Evaluation?" Social Science Quarterly, 54 (March, 1974), 759-61.

SUPPLEMENTAL REFERENCES

1. Blumstein, Alfred. "Management Science to Aid the Manager: An Example from the Criminal Justice System." Sloan Management Review, 15 (Fall, 1973), 35-48.

2. Blumstein, Alfred, and Cassidy, R. Gordon. "Benefit - Cost Analysis of Family Planning." Socio - Economic Planning Science, 7 (1973), 151-60.

3. Chackerian, Richard. "Police Professionalism and Citizen Evaluation: A Preliminary Look." Public Administration Review, 34 (March, 1974), 141-8.

4. Child, J. "What Determines Organizational Performance - Universals Vs. It All Depends." Organizational Dynamics, 3 (1974), p. 2.

5. Davidson, Frederick. "Dimensions of Utility in a Regional Planning Context." Decision Sciences, 5 (1974), 91-101.

6. Drake, A. "Plainfield Crime Prevention Unit." Journal of Criminal Justice, 1 (1973), p. 74.

7. Eisenstein, James, and Jacob, Herbert. "Measuring Performance and Outputs of Urban Criminal Courts." Social Science Quarterly, 54 (March, 1974), 713-23.

8. Greenwood, Peter W. "Evaluating the Outcome of Organizational Change." Workshop on Decision Analysis, IEEE-ORSA Joint National Conference on Major Systems. Anaheim, Calif., 1971. IEEE Transactions on Systems, Man, and Cyternetics, 2 (April, 1972), p. 275.

9. Hartenstein, Annette A. "A Cooperative Approach to Police Management Training." Journal of Police Science and Administration, 1 (December, 1973), 433-9.

10. Heller, Nelson B., and McEwen, J. Thomas. "Applications of Crime Seriousness Information in Police Departments." Journal of Criminal Justice, 1 (Fall, 1973), 241-53.

11. Hill, Morris. "A Goals - Achievement Matrix for Evaluating Alternative Plans." American Institute of Planners Journal, 34 (January, 1968), 19-29.

12. Hill, Morris. "A Method for the Evaluation of Transportation Plans." Highway Research Record, (No. 180, 1967), 21-34.

13. Helmes, J. C. "An Ordinal Method of Evaluation." Urban Studies, 9 (1972), 179-92.

14. Jones, P. M. S. "Determining Priorities and Investment Levels in Scientific Research and Development," Policy Sciences, 1 (1970), 299-309.

15. Kimberly, John R. "Environmental Constraints and Organizational Structure: A Comparitive Analysis of Rehabilitation Organizations." Administrative Science Quarterly, 20 March, 1975), 1-9.

16. Knowles, Patrick, and Peterson, Rolf. "Measurement of Flexibility in State Police Officers." Journal of Police Science and Administration, 1 (June, 1973), 219-23.

17. Krug, Donald; Andima, Haron; Bergner, Lawrence; Patrick, Sherman; and Whitman, Steve. "A Prevalence Estimation Model of Narcotics Addiction in New York City." American Journal of Epidemiology, 98 (1973), 56-62.

18. Larson, Richard C. "Decision - Aiding Tools in Urban Public Safety Systems." Sloan Management Review, 14 (Winter, 1972-73), 55-73.

19. Larson, Richard C., and Chaiken, Jan M. "Methods for Allocating Urban Emergency Units: A Survey." Management Science, 19 (December, 1972), 110-30.

20. Larson, Richard C., and Stevenson, Keith A. "On Insensitivities in Urban Redistricting and Facility Location." Operations Research, 20 (1972), 595-612.

21. Lichfield, Nathaniel. "Evaluation of Urban and Regional Plans: A Review." Regional Studies, 4, (1970), 151-65.

22. Lyden, Fremont James. "Using Parson's Functional Analysis in the Study of Public Organizations." Administrative Science Quarterly, 20 (March, 1975), 59-70.

23. Meyer, John C. "Police Attitudes and Performance Appraisal: The Forest and Some Trees." Journal of Police Science and Administration, 1 (June, 1973), 201-8.

24. Molnar, Daniel, and Kammerud, Marshall. "Developing Priorities for Improving the Urban Social Environment," Socio-Economic Planning Science, 9 (1975), 25-9.

25. Munro, Jim L. "Towards a Theory of Criminal Justice Administration: A General Systems Perspective." Public Administration Review, 31 (November, 1971), 621-31.

26. Nash, Christopher; Pearce, David; and Stanley, John. "Criteria for Evaluating Project Evaluation Techniques," American Institute of Planners Journal, 41 (March, 1975), 83-9.

27. O'Leary, Vincent, and Duffee, David. "Managerial Behavior and Correctional Policy." Public Administration Review, 31 (November, 1971), 603-16.

28. Ostrom, Elinor; Parks, Roger B.; and Whitaker, Gordon P. "Do We Really Want to Consolidate Urban Police Forces? A Reappraisal of Some Old Assertions." Public Administration Review, 33 (September, 1973), 423-32.

29. Parker, L. Craig, and Roth, Marvin C. "The Relationship Between Self-Disclosure, Personality, and a Measure of Job Performance of Policeman." Journal of Police Science and Administration, 1 (September, 1973), 282-6.

30. Poland, Orville F. "Program Evaluation and Administrative Theory," Public Administrative Review, 34 (July, 1974), 222-8.

31. Rutherford, G. Scott; Schofer, Joseph L.; Skutsch, Margaret; and Wach, Martin. "Goal Formulation for Socio-Technical Systems." Journal of the Urban Planning and Development Division, Proceedings of the American Society of Civil Engineers, 99 (September, 1973), 157-69.

32. Said, Kamal E. "A Policy-Selection/Goal Formulation Model for Public Systems." Policy Sciences, 5 (1974), 89-100.

33. Skutsch, Margaret, and Schofer, J. L. "Goals-Delphis for Urban Planning: Concepts in Their Design." Socio-Economic Planning Science, 7 (1973), 305-13.

34. Stanley, J. K. "A Cardinal Utility Approach for Project Evaluation." Socio-Economic Planning Science, 8 (1974), 329-38.

35. U. S. Congress. Joint Economic Committee. Systematic Analysis and Grants-in-Aid in a Federal System, by John Cotton and Selma Mushkin, Joint Economic Print, The Analysis of Public Expenditures, Vol. 1, (Washington, D. C.: Government Printing Office, 1969), pp. 332-54.

36. Weddle, Robert S. "City of Tyler (Texas) Crime Prevention Project." Journal of Criminal Justice, 1 (Summer, 1972), 173-76.

37. Zacker, Joseph, and Bard, Morton. "Effects of Conflict Management Training on Police Performance." Journal of Applied Psychology, 58 (1973), 202-8.

APPENDIX 1

MASON'S RULE CALCULATIONS FOR THE SECOND EXAMPLE APPLICATION IN SECTION 3

The calculations shown here are for the second example application of signal flow graphs to describing interrelationships between component organizations as described in Chapter 3. The mathematics used here are applicable to any network in which the direct relationships (gains) between any two components in the network is approximately linear over some decision interval. The signal flow graph does allow for indirect relationships to be evaluated, such as the situation where a first variable affects a second, and the second affects a third. The approach also allows for feedback, transfer of the output of an agency back to its input. This feedback is not dynamic, however. This means there is no delay between the time the output is produced and the time it is received at the input. Also, the feedback coefficients that describe the magnitude and direction of the feedback are constants.

The gain between any two component organizations is found by calculating the gain between the representative blocks or nodes on the signal flow graph. Generally, gain is defined as the output of the second node divided by the input at the first. Because of feedback and indirect relationships, a special formula called Mason's Rule is required to find the gain between two organizations. The formula was developed by Mason [82], and the form shown here is from Kuo [80].

$$G = \frac{1}{\Delta} \sum_{N} G_N \Delta_N$$

where

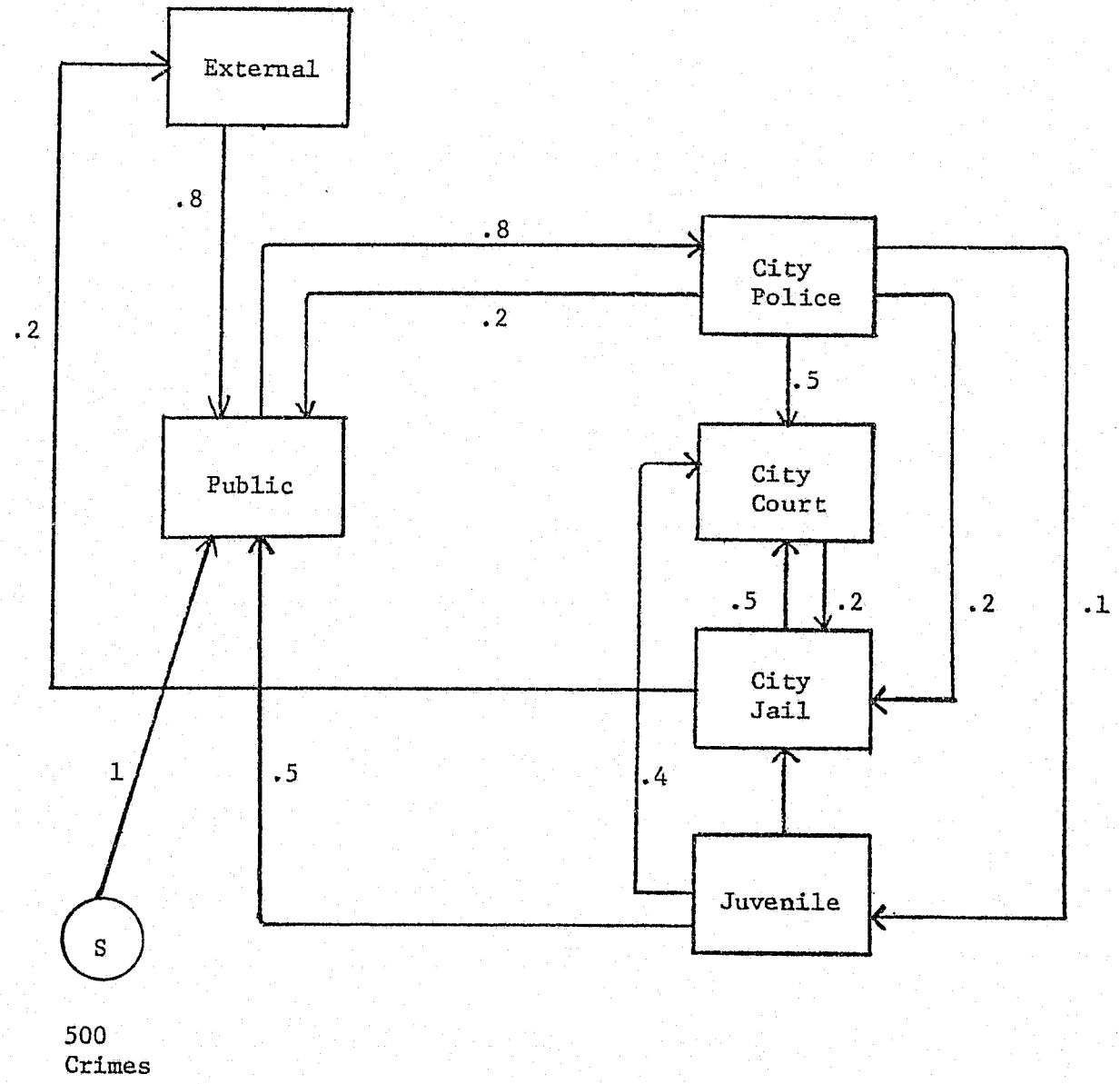
G_N = the gain of the nth forward path

and $\Delta = 1 - \sum_m P_{m1} + \sum_m P_{m2} - \sum_m P_{m3} + \ldots$

where P_{mj} = product of the gain the mth possible combination of j nontouching loops (loops which do not share any common nodes)

and Δ_N = the value of Δ for that part of the signal flow graph not touching the nth forward path.

G is the gain between the two notes (representing agencies). A combined probabilistic flow and signal flow diagram is shown in the following figure. An input of 500 crimes is given. Note that the input is assumed to be converted to arrested criminals with the rather high probability of 0.80. The rest of the signal flow graph deals in terms of criminals. In the calculations, the relationship between the input of crimes against the public and volume of criminals handled by every other component agency is determined.



Combined Probabilistic Flow
and Signal Flow Diagram

$\Delta = .63704$

$G_1 = (.8)(.5)(.2)(.2) = .016$

$G_2 = (.8)(.2)(.2) = .032$

$G_3 = (.8)(.1)(.1)(.2) = .0016$

$\Delta_1 = 1-0 = 1$

$\Delta_2 = 1-0 = 1$

$\Delta_3 = 1-0 = 1$

$$G = \frac{1}{\Delta}\Sigma_N G_N \Delta_N = \frac{1}{.63704} \quad (.016 + .032 + .0016)$$

$G = .07786$

CALCULATING GAIN
OF NETWORK BETWEEN
"PUBLIC" AND "EXTERNAL"

$$\Delta = 1 - [(.8 \cdot .2) + (.2 \cdot .5) + (.8 \cdot .5 \cdot .1) + (.8 \cdot .5 \cdot .2 \cdot .3) + (.8 \cdot .1 \cdot .5) + (.8 \cdot .1 \cdot .4 \cdot .1) + (.8 \cdot .1 \cdot .1 \cdot .5 \cdot .1) + (.8 \cdot .1 \cdot .1 \cdot .3) + (.8 \cdot .1 \cdot .1 \cdot .2 \cdot .1) + (.8 \cdot .5 \cdot .2 \cdot .2 \cdot .8)] + [(.8)(.2)(.2)(.5) + (.8)(.1)(.5)(.2)(.5)] = .63704$$

$G_1 = (.8)(.5) = .4$

$G_2 = (.8)(.1)(.1)(.5) = .004$

$G_3 = (.8)(.1)(.4) = .032$

$\Delta_1 = 1-0 = 1$

$\Delta_2 = 1-0 = 1$

$\Delta_3 = 1-0 = 1$

$$G = \frac{1}{\Delta} \sum_N G_N \Delta_N = \frac{1}{.63704} [.4 + .004 + .032] = .6844154213$$

CALCULATING GAIN
OF NETWORK BETWEEN
"PUBLIC" AND "CITY COURT"

$\Delta = .63704$

$G_1 = .8$

$\Delta_1 = 1-(.2)(.5) = .9$

$G = \frac{1}{\Delta} \Sigma\ G_1\Delta_1 = (.63704)^{-1}\ (.8)(.9) = 1.130227301$

CALCULATING GAIN
OF NETWORK BETWEEN
"PUBLIC" AND "CITY POLICE"

$\Delta = .63704$

$G_1 = (.8)(.5)(.2) = .08$

$G_2 = (.8)(.2) = .16$

$G_3 = (.8)(.1)(.1) = .008$

$\Delta_1 = 1-0 = 1$

$\Delta_2 = 1-0 = 1$

$\Delta_3 = 1-0 = 1$

$G = (.63704)^{-1}(.08 + .16 + .008) = .3893005149$

CALCULATING GAIN
OF NETWORK BETWEEN
"PUBLIC" AND "CITY JAIL"

$\Delta = .63704$

$G_1 = (.8)(.1) = .08$

$\Delta_1 = 1 - (.5)(.2) = .9$

$G = (.63704)^{-1} (.08)(.9) = .1130227301$

CALCULATING GAIN
OF NETWORK BETWEEN
"PUBLIC" AND "JUVENILE OFFICER"

PERFORMANCE MEASURES FOR EVALUATION OF LEAA AND CJS PROGRAMS

by

Richard C. Larson

Arnold Barnett

Amedeo Odoni

October, 1975

TABLE OF CONTENTS

I. INTRODUCTION

Each year the Law Enforcement Assistance Administration (LEAA) spends hundreds of millions of dollars on programs aimed at improving the components of the criminal justice system throughout the United States. Since LEAA's inception (1968), there has been widespread concern that the monies being spent should be for innovative new programs whose underlying hypotheses could be tested in order to advance our knowledge of CJS operation and the entire social environment impinging on the problem of crime. It is well known that much of the money has not been spent in this manner. While indeed some funds had to be allocated to short-term projects of high priority to local agencies, thereby "buying" credibility with the agencies, other funds were channeled to diverse projects that were often not well conceived, duplicative, and designed in such a way that no one could learn from the experience of implementing the project. The proverbial "recreation of the wheel" has been a severe problem with LEAA projects and programs. Some argue that much of this activity has served the useful purpose of educating CJS planners, managers and consultants throughout the United States, and therefore that the funds may not have been wasted. Yet it is difficult to imagine that such education could not be achieved in less expensive ways. The very design of the LEAA--providing three distinct mechanisms for channeling funds to a particular agency (state funds via the SPA, discretionary funds via the regional office, and research funds via the National Institute of Law Enforcement and Criminal Justice

[NILEC])--is conducive to a lack of communication among grant and contract recipients. LEAA staff members themselves often are not aware of other LEAA programs related to or even nearly equivalent to those in their own bailiwick. Even if they are aware of such programs, it is often difficult to obtain a final report of already completed projects--and few final reports contain an adequate unbiased assessment of the outcomes of the program that would be useful to others.

Need for Evaluation

Thus, a strong need is apparent for new mechanisms for appraising or evaluating LEAA programs and for disseminating this information throughout the United States. The activity of evaluation implies a focus on measurable quantities, on system inputs, on measures of process, on system outputs, and on final outcome measures. While quantitative measures do not exclude qualitative analyses--and many process evaluations must use interview and participant-observer techniques to develop a qualitative case study--they do bring about a new accountability in CJS operations. If defined appropriately and used with care, they allow system administrators (or outsiders) to compare the performance of the same system at different times or different systems at the same time. And they facilitate the evaluation of innovative programs and experiments.

While certain quantitatively-oriented measures have been accepted by CJS administrators for some time (e.g., the FBI Index Crime Rates, clearance rates, rates of recidivism from correctional programs), many

of these have been subject to abuse and manipulation or at least misinterpretation. For instance, police have well-known discretion in categorizing crimes and reporting clearance rates. And citizens have a difficult time relating rates of index crimes per hundred thousand to their own personal risk situation. Yet "crime clocks" ("5 rapes every x minutes") and other popular crime reporting mechanisms do not clarify the risk situation, and often only contribute to undue alarm. Recidivism rates are traditionally different by a factor of two or more, depending on one's point of view--police (at one end of the CJS) or corrections (at the other). This confusion--which is primarily definitional--is compounded by the fact that few people--CJS administrators or others--can project the consequences of an 80 percent recidivism rate, say, over the lifetime crime career of an individual. Thus, there are needs to portray the crime picture in a better way to citizens and administrators in order to reveal personal risks more accurately and to link certain system performance measures to other (perhaps more fundamental) measures.

A focus on quantitative measures in evaluation suggests an ability to predict the consequences (in quantitative terms) of alternative program designs prior to implementation. This implies the use of models of system behavior. Strategic models such as Professor Blumstein's JUSSIM can be used at a CJS-wide level to assess (before the fact) the system consequences of increased workload or more personnel at a particular stage or decreased recidivism rates from a particular correctional program.

Likewise, at an agency operational level, tactical models of operation--such as models of police patrol deployment--can be used to project the consequences of alternative patrol allocations. However, from the CJS administrator's point of view, many such models are now clouded behind the mystique of Greek symbols, long equations, and computers. Thus, tools which would be useful in evaluations are perceived to be inaccessible to the typical CJS administrator.

Outline of Paper

The purpose of this paper is to address the problem of CJS and LEAA evaluation from primarily three different points of view. Given the time and length constraints under which the paper was written, our approach has been to lay out these three areas for potential further investigation, but without fully detailing any particular one.

We start at the macroscopic level. How can a particular component of the CJS (say police, courts, or corrections) use expenditure, employment, and resource allocation data to compare its own efficiency and effectiveness over time and against that of equivalent agencies in other jurisdictions? Focusing on police departments as a case example, we discuss ways to study these departments on the basis of data such as those contained in the annual volumes on "Expenditure and Employment Data for the Criminal Justice System" issued by LEAA and the annual "Survey of Municipal Police Departments" issued by the Kansas City Police Department. Questions about near-term (15-year) trends will be addressed and ways will be suggested

for checking the validity of several common assumptions regarding the Criminal Justice System through the use of data of this type. Certain possible pitfalls in the processing of such information will be pointed out. These pitfalls are due primarily to two potential sources of error: the excessive degree of aggregation of the data as they appear in the surveys; and the differences from place to place in the statutes regarding the functions of police departments. This part of the paper will attempt to make a case for standardization of data gathering procedures around the United States and for cross-sectional statistical studies of Criminal Justice System expenditures and employment data as an aid to a variety of decision-makers in this area.

Three families of indicators of "output" for the CJS are then identified. It is believed that the more meaningful of these indicators--the ones which could truly measure the reduction of crime and the overall performance of the CJS--are either too difficult to quantify or are only obliquely related to specific decisions by policy makers with regard to the allocation of resources in the CJS. For this reason, it may be preferable, at least in the short run, to concentrate on using more readily measurable indicators that measure such items as productivity, efficiency, and sub-system performance as gauges in helping evaluate the overall performance and effectiveness of alternative CJS programs.

Continuing at an aggregate level, we next explore certain system-level performance measures that deal with crime, victimization, and recidivism. The focus is on improved methods for collecting, processing, and

interpreting data related to these key issues. This includes more careful attention to controls in LEAA-sponsored victimization surveys. It also includes projecting easy-to-collect annual statistics over a multi-year horizon. For instance, in auto accidents, the annual rate of accidents per 100,000 provides one number for a person to gauge his risk of being in an auto accident. But typically this figure can be extrapolated via models (or in some cases, actual recorded data) to reveal that the "average" motorist becomes involved in an auto accident once about every ten years. In a similar manner we wish to explore certain CJS performance measures that extend beyond the one-year time horizon and allow one to project into the indefinite future. Particularly when dealing with small probabilities, people are notoriously bad at perceiving the implication of low probability events. Pedestrians and automobile drivers in some urban centers have a one-in-10^6 chance of being killed in any particular day due to an auto accident; in some cities the risk of being murdered is even greater. Yet what does daily reporting imply about a person's perception of this risk? Projected over a lifetime, the probability of being killed in an automobile accident or by a murderer can be one in fifty or higher. Similar projections apply in the area of corrections (when dealing with recidivism) and courts (when evaluating the long-range consequences of alternative sentencing policies). So, much of the second section of this paper addresses the issue of long-range projected performance measures and how they could be used more directly in evaluating both CJS and LEAA programs. One important (and novel) evaluative use of

such data could be in monitoring the effects of new societal policies dealing with the CJS. For instance, examples are outlined describing the potential evaluation of the crime-reduction effects of new gun control laws and/or new laws dealing with convicted murderers (e.g., limited use of the death penalty). The section concludes with several specific recommendations to LEAA in the area of crime occurrence, victimization, and recidivism data.

The third section of the paper focuses on less aggregated questions, namely those dealing with operationally-defined performance measures and their use in evaluation. Two different types of evaluation are identified at the state or local level--evaluation of experimental programs and evaluation of routine day-to-day operations. For the first, it is proposed that LEAA should support the development of a formal CJS-focused evaluation methodology, including the sequence of experimental steps ranging from identification of a problem area; to generation of hypotheses; to experimental design, execution, and monitoring; and finally to after-the-fact evaluation. Special emphasis is given to the role of formal models in each of the various phases of the evaluation process. For the second, it is argued that several quantitative measures, models, and methods currently exist that, if properly packaged, could be utilized by numerous CJS agencies in evaluation of very specific day-to-day operational problems. It is recommended that easily-understood handbooks be written for this purpose. This section also concludes with several specific recommendations for LEAA.

The fourth and final section of the paper details some specific recommendations to LEAA in the area of evaluation. While most of the recommendations are derived from the preceding three sections, several are more general, pertaining to the entire process of national evaluation and dissemination which LEAA would like to influence.

II. SYSTEM-LEVEL AGGREGATE MEASURES OF PERFORMANCE

The massive efforts, in terms of expenditures and manpower, that federal and local agencies have undertaken during the last few years for the purpose of containing crime in the United States have made at least one thing clear: there is a need for careful reappraisal of the ways in which resources are allocated among and within the various local, state, and federal criminal justice agencies.

Central to such an appraisal is the capability to address the issue of relative effectiveness versus cost of allocative decisions and hopefully, thereby, to perform meaningful comparisons among alternative programs. Development of such a capability is the goal of most responsible public administrators. Unfortunately, in the case of almost every single aspect of the criminal justice system (CJS) no such capability exists today. This is not due to a lack of trying in this direction, but rather to the enormous difficulties inherent to the subject. Our understanding of cause-and-effect relationships in the area of crime--an understanding which must precede the performance of effectiveness versus cost appraisals--is sorely deficient. As a consequence of this state of affairs, decisions concerning the allocation of CJS resources are made mostly on the basis of past experience, intuition, and wishful thinking (in varying proportions).

This chapter proposes several modest, yet feasible steps that could be taken under LEAA supervision to begin the process of assembling CJS-level data that could eventually be used to assist the resource allocation process. We plan to describe a family of readily obtainable measures of inputs and, what we shall call, "intermediate outputs" of the CJS. These measures, in turn, provide the basis for a few meaningful, albeit highly imperfect, indicators of effectiveness.

Specifically, subsequent sections will cover the following main points:

a) Insufficient attention has been devoted in the past to the establishment of an information basis for understanding current patterns and historical trends regarding the allocation of inputs (resources) in the CJS. Although several private and governmental organizations conduct a variety of periodic surveys of criminal justice agencies, this information often suffers from lack of adequate detail or analysis to assist system-level decision makers.

b) The information already available from concluded surveys has not been analyzed in a way to shed light on system-level resource allocations. Nonetheless, a preliminary examination of such data might reveal several consistent patterns and provide valuable perspectives with regard to the composition of the various costs associated with the CJS.

c) The known relationships between, on the one hand, the inputs (resources) to the CJS and, on the other, its true outputs (i.e., its preventive, deterrent, and corrective effects on unlawful behavior) can best be described as tenuous. In the short run, and until this state of knowledge improves, we may be better off by measuring the effectiveness of the various CJS sub-components in terms of relationships between inputs and more tangible intermediate outputs.

d) A study and analysis program to deal with items a) through c) above is a worthwhile activity for the LEAA to undertake.

Orientation

Before proceeding to a detailed discussion of these points, it is worth clarifying the aims and orientation of the programs to be proposed here. The emphasis is not on sophisticated analysis, but on the development of information that may be useful to several target audiences and of measures of effectiveness that may provide additional focus to their deliberations. Such target audiences may include:

a) LEAA-State Planning Agencies and other regional agencies that administer the distribution of federal funds among regional, state and local levels. The primary concern of these agencies is the effective utilization of such funds. The proposed information and measures may prove of assistance to these groups, especially

during the review phase of existing programs (in anticipation of refunding) in coming to a judgement on whether resources are being effectively utilized.

b) Management and union representatives in the Criminal Justice area who, under today's conditions, must often negotiate in a virtual vacuum with respect to information on wages, benefits, and productivity standards both on an absolute and a comparative scale.

c) Department of Justice and LEAA planners who must decide, at the federal level, on the proper allocation of resources among and within the various components of the CJS in the short and in the long run. It would be utterly presumptuous even to suggest that anything proposed here will allow any such decisions to be made in a less subjective (or more "mechanistic") way in the future. It is, however, plausible to assume that a better perspective on current trends in expenditures and performance in the CJS will provide an improved environment for decision-making. The identification of trends is a major objective of the programs to be proposed here.

d) Elected and appointed administrators at the state and local level (city managers, attorneys general, governors, etc.) entrusted by the public with monitoring the efficient operation of CJS agencies. The program advocated here will facilitate the performance of these duties by such officials. For instance, a city manager could take advantage of newly available information

to inquire as to why his city "spends more money per capita on police protection than city X, while also obtaining inferior service (according to a given measure of performance) than city X."

e) Researchers on the CJS: drawing from the greatly expanded information bank to be advocated here will, in many cases, free these researchers from the onerous task of collecting survey data on the CJS on every occasion when such data are needed. Studies of a national rather than of a local or regional scope will also be encouraged through the availability of such information.

f) Last but not least, the public at large will benefit through the availability of regularly published and readily comprehensible information on the costs and, perhaps, on some indicators of performance for the CJS. The Department of Justice has long ago recognized the public's "right to know" in this area through the annual publication of the FBI Uniform Crime Report and, more recently, with the annual issuing of statistics on Expenditure and Employment Data for the Criminal Justice System [1].

Resources Allocated to the Criminal Justice System

Existing Information Sources

The resources allocated to the CJS can best be measured in terms of funding commitments and of manpower. The sole Federal Government

publication which is specifically oriented toward the compilation and presentation of data along these lines is the annual volume [1] on Expenditure and Employment Data for the Criminal Justice System, issued jointly by the U. S. Department of Justice (LEAA, National Criminal Justice Information and Statistics Service) and the U.S. Department of Commerce (Bureau of the Census, Social and Economics Administration). This annual survey covers all facets of the CJS (police protection, courts and other legal services, and corrections). Data are itemized along the usual federal-state-local lines of authority, as well as by state, county, standard metropolitan statistical area (SMSA), and municipal government. Unfortunately, this information suffers from an excessive degree of aggregation: the details of the allocation of resources within each of the CJS subsystems (police departments, courts, corrections systems) are not dealt with. Consequently, the value of the data for resource allocating decisions is limited, except from the case of "global" decisions of the type that a top-level Federal Government official might make. An examination of the questionnaire used for the compilation of this survey shows that the data collected are of limited value for an in-depth analysis of the underlying causes of changes in CJS expenditures and employment.

Other annual surveys of CJS manpower and expenditures at the local level are available through the following publications: City Government

Finances (U.S. Department of Commerce, Bureau of the Census) [2], The Municipal Yearbook (International City Managers Association - ICMA) [3], and the Survey of Municipal Police Departments (Kansas City Police Department) [4]. The Census of Governments (U.S. Department of Commerce, Bureau of the Census) [5] conducted at five-year intervals provides data similar to those contained in the City Government Finances but in greater detail and itemized along several types of governmental or statistical-purpose jurisdictions.

The value of each of the aforementioned statistical compilations to the analyst can best be assessed by reference to the breadth and depth of their coverage. At one extreme, the Census of Governments and the City Government Finances cover all aspects of the CJS on a nation-wide basis but in a highly aggregative manner. At the other end of the spectrum, the Survey of Municipal Police Departments, as its name implies, is limited to police departments and, in fact, surveys only the cities with a population of 300,000 to 1,000,000 in 1970 (43 cities are covered by the latest surveys). This latter survey provides a very large amount of information for each of the police departments that it covers: numerous items are listed including a breakdown of expenditures by function, employment and salary figures by category of employee, benefit and pension data, information on mechanical equipment in use, some demographic data on the city in question, etc. In this connection, an examination of the latest questionnaire being used by the Kansas City Police Department shows that while some of the questions included in this highly detailed form may be of uncertain informational value or may be inadequately

defined, it is clear that data obtained through this survey may be highly useful in monitoring and understanding the process of change with regard to police department resource utilization.

The type of cost-effectiveness analysis work which will be advocated in this chapter would require information of a similar level of detail as that presented in the Kansas City Survey of Municipal Police Departments. Studies of a nation-wide scope and dealing with the entire CJS imply an information base with depth and breadth comparable, respectively, to those of the Kansas City surveys and of the Bureau of the Census quoted publications. Although development of such a data base may appear to be an ambitious task, it represents in practice only an incremental change in the effort required to procure and compile the present Expenditure and Employment Data for the Criminal Justice System.

Admittedly, however, careful work is needed in order to identify and select all the items of information to be collected and in order to prepare an unambiguous questionnaire which properly defines the requested items for the benefit of the responding local officals. With respect to this latter subject, it is worth mentioning that the unambiguous description of the data sought is by no means a trivial matter. It is often done carelessly at the present time and leads to (occasionally large) differences among the figures reported for the same item in the various surveys. For instance, the figures reported for the total annual cost of local "police protection" in New York City vary, at present, by as much as 30% (or a range of $200 million) depending on the respondent's

interpretation of what should be included under this item. (Transit Authority Police expenditures and Housing Authority Police expenditures in New York City, for instance, are not included in the funds allocated to the New York City Police Department. Some respondents include the former expenditures as parts of "police protection costs" and others do not.)

Analysis of Resource Inputs to the CJS

As noted in the last section several sources are already available with information concerning the resource inputs (economic and otherwise) to the CJS. Despite their serious deficiencies, which were also noted above, these information sources contain sufficient material to provide the basis for serious study of the composition of the various CJS costs. Unfortunately, this particular subject has apparently attracted only limited attention. Part of the explanation for this phenomenon may lie with the (justified) preoccupation of criminologists and other analysts with the investigation of trends and statistical patterns contained in annual crime statistics, case disposition statistics in the courts, recidivism statistics, etc.

A recent review by the authors of existing work on the composition of city police department costs turned up surprisingly little material of interest. Perhaps the most worthwhile is the recently published analysis of Bahl, Campbell and Greytak [6] of expenditure and revenue patterns in New York City. As part of this analysis, the authors have examined in some detail the costs of the New York Police Department and,

through a simple mathematical model have attributed cost increases to the factors of increased employment in the police department, inflation, and real wage gains. They have also performed some simple comparisons between per capita police protection costs in New York City and similar costs in nine other major cities.

Kakalik and Wildhorn [7] in another interesting earlier investigation collected a considerable amount of information on cost breakdowns, manpower, manpower allocation, mechanization, etc. for a few large police departments and, by using simple calculations and regression analysis, arrived at several conclusions and conjectures regarding resource allocation in police work.

Several studies in this area have also been performed in the past in connection with the annual and periodic surveys of the International City Managers Association (ICMA)--often conducted in cooperation with other private or governmental organizations. Among the most recent, Lewin [8] has examined trends in salaries and manpower for several categories of cities, while Anderson [9] in 1973, analyzed the data from a survey of fringe benefit provisions for municipal employees taken in late 1971.

Another set of writings has originated from the Bureau of Labor Statistics of the U.S. Labor Department. For instance, Davis [10], has published summary reports on pay scales and salary trends for policemen and firemen. Data in these reports are aggregated by groups of cities and the percent distribution of pay scales is also provided.

The aforementioned ICMA- and BLS-related reports can be criticized primarily for providing a minimum amount of perspective. Medium- and long-term

trends (i.e., covering a time span of five or more years) are very seldom provided. Similarly, police data are not compared with similar data for other services (except possibly for municipal fire departments which, almost by definition, exhibit similar characteristics) and with data from the private sectors of the economy. Yet, it is only in the long run and by comparison to those other sectors of the economy that changes in police expenditures and employment in the United States can be appreciated and understood.

Several investigations have concentrated on areas which are related to our main area of interest: A lengthy study by Shoup and Mehay [11] has attempted to demonstrate the merits of the Program Budgeting system through application to the case of police services in the Los Angeles area. By inferring the costs of the various types of crimes and of police activities, they have advocated adoption of a cost/benefit (or cost/effectiveness) approach to the allocation of police resources. On the subject of cost/effectiveness, a most thoughtful short paper by Blumstein [12] also deserves special mention.

Finally, an extensive amount of work has been done on multi-variate regression analyses that attempt to identify statistical relationships among police inputs (mostly police costs per capita), crime statistics (or victimization rates) and a host of environmental variables (such as demographic data, street mileage, geographical location, etc.). These studies usually concentrate on groups of cities within a given state or cities located within one or a few metropolitan areas. Among the most recent ones are the analyses of Beaton [13] (New Jersey cities), Walzer [14] (31 cities in Illinois), Hirsch [15] (64 St. Louis metropolitan area police departments), Shoup and Mehay [11] (52 cities in Southern California), and Sunley [16] (selected cities in the

metropolitan areas of Detroit, Cleveland, Minneapolis-St. Paul and Pittsburgh). None of these works examines in any detail the composition of police costs.

Data Collection and Study Recommendations

The brief review of existing literature on police expenditures and resource usage illustrates what are believed here to be common characteristics of similar past work on other aspects of the CJS as well:

a) Studies have to rely on incomplete data bases in terms of both depth and breadth.

b) The focus of attention has been on aggregate measures of cost and resource utilization, while questions related to the detailed composition of costs, cost increases and employment figures have been largely ignored.

c) The few studies that exist, with rare exceptions, offer little in terms of long-term perspectives on developments in the CJS area.

d) Similarly, few comparisons are provided with parallel developments in other areas of public or private activity, especially in the public safety and welfare sectors.

In view of these deficiencies, two principal recommendations can be made:

i) The Law Enforcement Assistance Administration should assume responsibility for the collection and dissemination of information on the various types of resources utilized by the CJS. The collection of the data should be on an annual basis. In terms of breadth, the survey should be of a scope similar to that of the LEAA-Census Bureau annual volume on Expenditure and Employment Data for the Criminal Justice

System. In terms of depth, however, the survey should be considerably expanded and seek much more detailed information than that which is obtained currently. The LEAA, moreover, should publish on an annual basis and in easily comprehensible form the results of these surveys. Emphases in these annual publications should be on exhibiting trends in CJS expenditures and employment and in discussing the probable underlying reasons for these trends. The format of the presentation should be consistent with the needs and backgrounds of the various potential audiences for these surveys which were identified earlier.

The LEAA annual publications should draw their material and support from work done under the second program recommended here:

ii) The Law Enforcement Assistance Administration should undertake or support a series of studies to analyze CJS expenditures and manpower data. These studies should include the examination of time-series trends for individual locations or for groups of jurisdictions and the performance of cross-sectional comparisons among individual (or groups of) municipalities, SMSA's, states, or regions with varying or similar characteristics.

Recent research performed by Odoni [17] on the subject of near-term (1959-1973) trends in police department resource utilization explored issues similar to those suggested under item ii) above. On the basis of that experience, the following (far from exhaustive) list of promising topics, each concerning one or more aspects of the CJS, can be suggested:

a) An analysis of the make-up and composition of the well-known momentous increases that police department budgets (and CJS expenditures, in general) experienced during the last decade.

b) The major trends in salaries and benefits for CJS personnel and the relationship--if any--between wage gains and the various unionization movements that have taken place among police and among correctional system employees.

c) Internal trends in CJS employment such as changes in the relative proportions of uniformed and civilian employees and of ranking officers and regular patrolmen in police departments, or of clerical and professional personnel in the court system.

d) The relative allocation of local and state resources to the CJS as measured by the proportion of local and state budgets spent for this purpose. A review of trends in this respect will provide an indication as to the position that the crime problem occupies in the list of national priorities in view of widespread public concern.

e) The details of the allocation of police manpower among various possible functions such as preventive patrol, response to calls for assistance, investigations, clerical tasks, enforcement of traffic regulations, etc.

f) The marginal cost on an annual and present value basis of additional CJS employees by function and specialty based on the current status of salaries and benefits. In a similar manner, the marginal cost of manning additional functional positions, e.g., of fielding an around-the-clock foot patrolman or of an around-the-clock one- or two-person police car.

g) The relationship of "size" to CJS expenditures. This is the important question of whether any economies of scale may exist in the CJS. While the problem has been partially investigated with respect to police departments (see previous section) evidence to date has been far from conclusive.

h) The budget fractions allocated to capital investments and to current expenditures (the latter consisting mostly of personnel costs) in different parts of the CJS. Also the historical role, if any, that recent funding from the Law Enforcement Administration has played in modifying these fractions. These federal funds, intended as they are to promote "innovation" in the CJS, would--at least theoretically--be more likely to increase capital investments in the form of faciilties and, especially, such expensive equipment as computers, communications apparatus, car locator systems, etc.

It is believed that the additional expenditures for the creation of an expanded data base, as recommended here, would be well justified in

the long run, even if that base would only serve to help clarify some of the questions that were just posed. In addition, however, the data base will also support one of the two aspects (namely, the one related to costs) of the cost-effectiveness studies that will be discussed in the next sections.

Measuring the Outputs of the CJS

Types of Measures for the CJS

Figure 1 presents one possible conceptualization of the various types of measures that could be used to describe the different parts of the CJS. Four categories of measures are identified: input measures, intermediate measures of output, crime-related measures of output and "true" measures of output.

We have already discussed various input (or "resource allocation") measures in earlier sections. It is feasible to collect information about the manpower and expenditure costs of the CJS to practically any desirable level of detail, including the specific allocation of funds and personnel to distinct identifiable tasks. In most cases, this information is readily available--at least at the local level--and can be obtained through the use of appropriate surveys.

At the opposite end of the spectrum, one can theoretically describe measures of the true effectiveness of the CJS. It would be erroneous to use such measures as changes in crime indices for this purpose. The

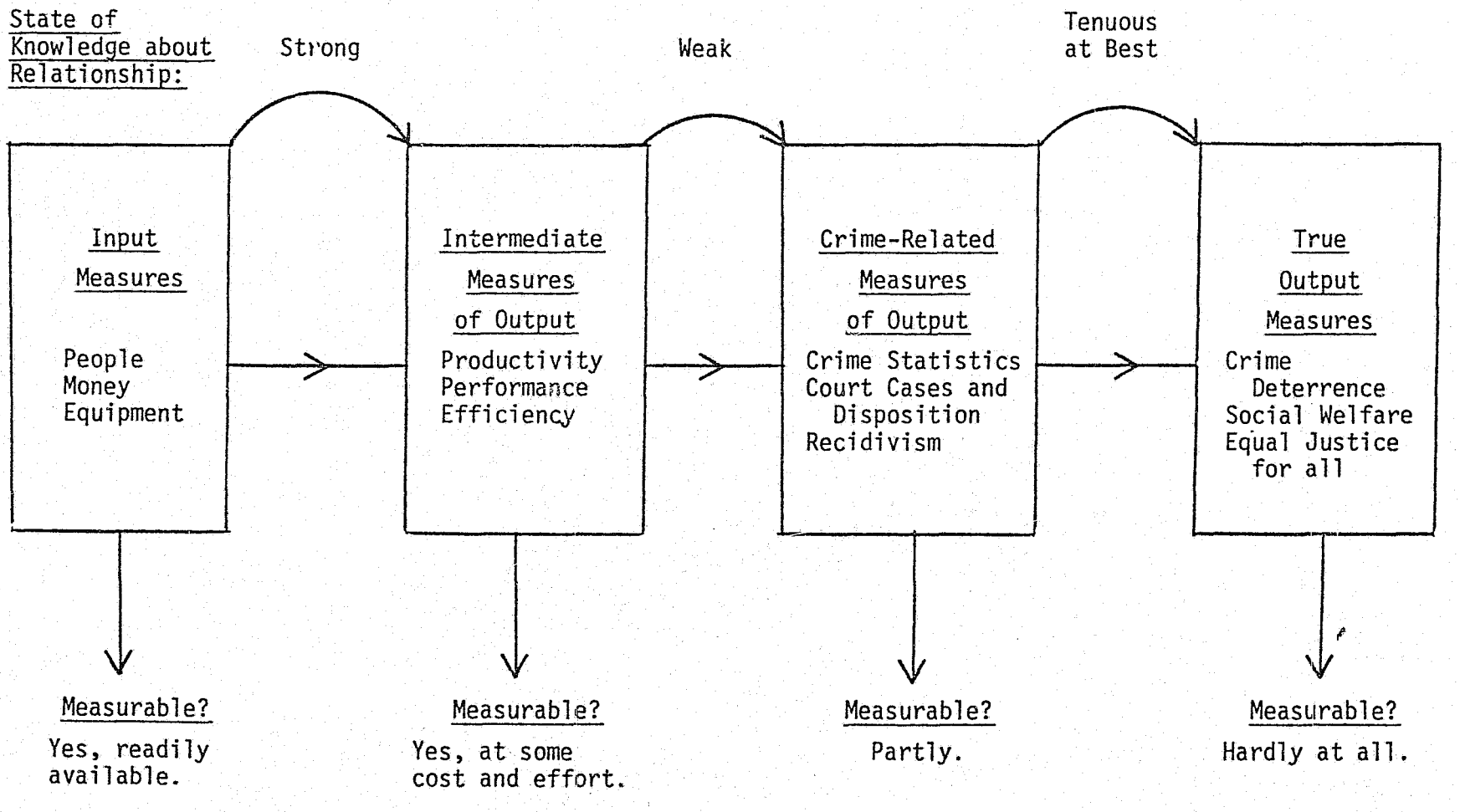

State of Knowledge about Relationship:
Strong
Weak
Tenuous at Best
Input Measures
People
Money
Equipment
Intermediate Measures of Output
Productivity
Performance
Efficiency
Crime-Related Measures of Output
Crime Statistics
Court Cases and Disposition
Recidivism
True Output Measures
Crime Deterrence
Social Welfare
Equal Justice for all
Measurable?
Yes, readily available.
Measurable?
Yes, at some cost and effort.
Measurable?
Partly.
Measurable?
Hardly at all.

Figure 1

number of crimes and other similar measures are insufficient indicators of the true outputs of the CJS. Higher level measures of effectiveness are desirable: as with most large-scale social systems, so with the CJS, too, it can be argued that its true objective is the "maximization of social welfare," in this particular case through the prevention and deterrence of crime and through the provision of a fair and equitable system of justice for all.

Unfortunately, it is practically impossible to obtain any quantitative measurements of outputs which, to begin with, are as "hazily" defined as, for instance, the terms "social welfare" or "crime prevention" are. Even after attempting to express these global outputs in more specific terms, experience has shown that the situation with regard to quantitative measurement does not improve perceptibly. For example, Blumstein [12] has suggested a measure called "social disruption" as a high-level indicator of output for the CJS. After describing this measure, however, Blumstein concludes as follows:

> Having thus identified this higher measure of "social disruption," its complexity and its many unquantifiable features preclude operating with it analytically at this time. Nevertheless, considerable value derives from considering these issues, for they pervade many later considerations, such as those relating to technological approaches to improving the effectiveness of crime control, perhaps at the expense of privacy or due process considerations.

The last sentence in this quote coincides with the authors' perception with regard to the potential usefulness of high-level output indicators in planning for the CJS: <u>they should be viewed as providing</u>

general qualitative guidelines for setting of policy goals. However, their use in effectiveness comparisons between specific alternatives should be ruled out at this time as impracticable. Inability to measure true outputs and our present complete lack of knowledge on how to predict the effect of alternative allocative decisions (CJS inputs) on the true outputs--as indicated at the top portion of Figure 1--make such a conclusion inevitable.

One, then, is forced to turn to the other two types of measures identified on Figure 1. Of those, the crime-related measures of output have probably attracted most of the attention in the past. They include such quantities as crime indices (crime-specific or aggregate), statistics on the disposition of cases in the court system, indicators of the effectiveness of the corrections system such as recidivism rates, etc. The major advantages of these measures are that they are readily understandable and, probably more importantly, are perceived to bear a more or less direct relationship to the true outputs of the CJS. For instance, most would agree that a reduction in the crime rate (if not achieved at the expense of civil liberties) also implies an increase in social welfare. Unfortunately, crime-related measures of output also suffer from two major deficiencies:

i) They are only partly measurable in many instances. The recent surveys of the LEAA, for instance, on the true incidence of crime in major cities indicates that reporting of crimes depends on a complex set of factors not the least

of which is the citizens' own perception of how effective the CJS is. In any case, the surveys clearly indicate that crime is grossly under-reported but it is difficult to determine just by how much. This point is discussed in considerably more detail elsewhere in this report.

ii) Even more crucial with regard to the performance of cost-effectiveness comparisons, is the great difficulty, in view of our present state of knowledge, that is involved in predicting the effects of resource allocative decisions on crime-related statistics. A specific example, which has often been used before, will help clarify this point: Assume that city X has decided to increase the number of police vehicles it fields during the peak crime-period of the day. It is, of course, quite simple to compute the costs of this decision in terms of increased manpower, equipment and funding requirements. (These are the input measures.) Given a description of the spatial distribution of calls for police assistance, of city X's geography, and of the dispatching policy followed by the police department in question, it is also possible, using the methodology developed in recent years by Larson [18] and others, to predict quite accurately what effect this decision will have in terms of a reduction in police response times, i.e., the delay between the time a call is received and the time when a police car first arrives on the scene. Police response time is one measure of CJS

performance or efficiency, (i.e., a measure that can be classified in the category of "intermediate measures of output" in Figure 1.) However, is is very difficult, with our present understanding of the "physics" of crime, to predict the implications, if any, of the reduced response time on the probability, say, of apprehending criminals and, consequently, on the chances for reducing the number of committed crimes. (The latter is a crime-related measure of output.) About all that is known is that reduced response times usually increase the chances of a successful arrest, but the exact relationship is far from clear and it is highly improbable that whatever functional relationship exists is a simple one.

This whole chain of associations is implied by the terms "strong" and "weak" on Figure 1 with regard to the state of knowledge on the relationships between input measures, intermediate measures of output and crime-related measures of output for the CJS.

Usefulness of the Intermediate Measures of Output

We turn now to what have been termed "intermediate measures of output" on Figure 1. Under this category we have included a large class of indicators of performance, efficiency, and productivity in the CJS. These indicators are connected to the true intended outputs of the CJS only through a series of logical inductions of the type already described

in the previous section. If one is willing to accept the validity of these logical inductions--that is, the argument that there is a direct connection between an efficient and smoothly functioning CJS and an effective CJS--then the intermediate output measures can be highly useful in monitoring the performance of the CJS and in providing guidelines for a desirable allocation of resources within and among its various components.

The last example in the previous section illustrates one of the reasons for recommending increased attention to intermediate measures of CJS output: The state-of-the-art in CJS analysis has reached the point where, in many cases, it is possible to predict what the effects on intermediate output measures will be of specific allocative decisions.

A second advantage of intermediate output indicators is that they are measurable--at some cost and after some effort, as indicated on Figure 1. The precise amount of effort required to obtain measurements varies from one indicator to another. For instance, certain indicators of police department performance such as "response times to calls for assistance" or "personnel turnover during a year" are relatively easy to deal with, either through observation of police operations (e.g., response times) or directly from departmental records (e.g., personnel turnover). On the other hand, intermediate output indicators which to some extent also incorporate the notion of "quality" (e.g., indicators of "smoothness" of operations in the court system) are clearly much more difficult to measure. Even in these latter cases, however, one can

devise simpler surrogate indicators which can be quantified and measured more easily.

A bewildering number of indicators related to performance, efficiency and productivity in the CJS can be (and have been) suggested. One cannot hope to deal with all or, even, with a substantial fraction of such indicators in practice. The correct approach is to select judiciously only a small number of indicators (= measures of intermediate outputs) making sure that each one of them is representative of a large "family" of other possible indicators and then deal with only the small sample at hand the rest of the way.

Some very thoughtful suggestions, very much along these lines, are provided--with specific reference to police departments--in the recent publication of the National Commission on Productivity entitled Opportunities for Improving Productivity in Police Services [19]. The authors of this report (Advisory Group on Productivity in Law Enforcement) have identified and described several measures of performance and productivity, a small group for each facet of police work. They, then, recommended that a national program be instituted for measurement and data collection with regard to these indicators. It is strongly suggested, in these recommendations, that the LEAA should act as the focus for this type of activity. A similar set of recommendations will be made at the conclusion of the next section, with regard to measurement, collection and dissemination of information about intermediate output measures for the entire CJS.

Effectiveness Comparisons Using Intermediate Output Measures

In addition to serving as descriptors of performance for the CJS, intermediate output measures in combination with input measures can be used as the basis for performing limited cost-effectiveness comparisons among alternative ways of allocating resources in the CJS.

We have already alluded several times to this particular type of cost-effectiveness analysis. Its main attractiveness is that, because of our ability to predict or measure the changes in the intermediate output indicators that result from specific changes in the allocation of resources, this type of analysis is both doable and believable. In fact, this seems to be the only type of cost-effectiveness analysis which is possible at all, given our present state of knowledge about the CJS. It is, therefore, altogether surprising that so little has been written or done in this area to date.

To the authors' knowledge, the only report which has explicitly addressed this problem is a brief but well-written study reported by Blumstein [12] in 1969, on a cost-effectiveness analysis in the allocation of police resources. As an example, Blumstein has used the case in which a number of alternatives aimed at increasing the probability of apprehension on the scene are compared for a given police department. These alternatives may include the installation of more public call boxes to reduce the delay in calling the police, the assignment of additional "complaint clerks" to the police dispatching center, the use

of a computer to accelerate the dispatching process, the introduction of additional police cars, etc.

In terms of our classification scheme, the "probability of apprehension on the scene" is a crime-related measure of output. Due to the practical impossibility of obtaining an explicit quantitative relationship between each of the alternatives under consideration and the resultant change in the probability of apprehension, Blumstein suggests that the measure of effectiveness to be used could be "delay saved per dollar allocated" (i.e., reduction in response time per dollar allocated, using our earlier terminology). Thus, the measure of effectiveness in this example is an intermediate output measure (response time) and the measure of cost an input measure (dollars invested). A multitude of other similar examples can be offered.

It is believed here that the methodological foundation is already available and that the data base can be created for successful applications of cost-effectiveness analyses of the type described to the CJS. Moreover, these comparisons of effectiveness need not be confined to the allocation of resources within any particular component of the CJS; in fact, a most useful application of the approach is as an aid to making decisions on how to allocate resources among the different subsystems of the CJS. In other words, ideally one should be able to determine the relative effectiveness of an extra dollar spent on, say, expediting the processing of cases through the courts as opposed to a car locator system for a police department. (The state-of-the-art, of course,

CONTINUED

3 OF 5

makes it highly unlikely that such comparisons can be performed at the present time at such a level of detail, although more gross comparisons, e.g., money spent on the courts versus money spent on police, are not beyond the realm of possibility.)

A model of the CJS that is particularly well suited to the conduct of cost-effectiveness analyses has been suggested by Blumstein and Larson [20]. The detailed breakdown of the CJS into a number of interconnected constituent parts makes this model an extremely convenient tool for this purpose. However, the details of how the model could be utilized are clearly beyond the scope of the present discussion.

On the basis of the discussion in the last three sections, the following additional recommendations can now be made:

i) The Law Enforcement Assistance Administration should sponsor studies and activities aimed at:

a) Identifying useful intermediate output measures (indicators of performance, productivity, and efficiency) for all aspects of the CJS.

b) Conducting carefully planned surveys and attempts at field measurement for the purpose of determining the current values of these indicators and the effects of various allocative decisions on the values of these indicators.

c) Exploring the use of intermediate output measures in combination with measures of input for the purpose of

performing comparisons among alternative allocations of resources in the CJS.

ii) The Law Enforcement Assistance Administration should act as a clearinghouse for the dissemination of information obtained from the activities described above to local and state CJS agencies and to the scientific community.

References

1. U.S. Department of Justice, Law Enforcement Assistance Administration and U.S. Department of Commerce, Bureau of the Census, Expenditure and Employment Data for the Criminal Justice System, Washington, D.C. (annual publication).

2. U.S. Department of Commerce, Bureau of the Census, City Government Finances, Washington, D.C. (annual publication).

3. International City Managers Association, The Municipal Yearbook, Washington, D.C. (annual publication).

4. Kansas City Police Department, Survey of Municipal Police Departments, Kansas City, Missouri (annual publication).

5. U.S. Department of Commerce, Bureau of the Census, Census of Governments, Washington, D.C., (years 1957, 1962, 1967, and 1972).

6. Bahl, Roy W., A.K. Campbell and D. Greytak, Taxes, Expenditures, and the Economic Base: Case Study of New York City, Praeger Publishers (New York: 1974).

7. Kakalik, James S. and S. Wildhorn, Aids to Decision Making in Police Patrol, Volumes I and II, R-593-HUD-RC, The RAND Corporation, Santa Monica, California (February 1971).

8. Lewin, David, "Expenditure, Compensation, and Employment Data in Police, Fire, and Refuse Collection and Disposal Departments," pp. 39-45, in The Municipal Year Book: 1975, International City Management Association, Washington, D.C. (January, 1975).

9. Anderson, Eric, "Fringe Benefits for Municipal Employees," The Municipal Yearbook: 1973, International City Managers Association, Washington, D.C.

10. Davis, William M., "Pay Scales of Police Patrolmen and Firefighters Rose More Than 6 Percent in 1973," Current Wage Developments, February-March, 1975, pp. 42-54.

11. Shoup, Donald C. and S.L. Mehay, Program Budgeting for Urban Police Services, MR-54, University of California, Los Angeles, Institute of Government and Public Affairs, Los Angeles, California (1971).

12. Blumstein, Alfred, "Cost-Effectiveness Analysis in the Allocation of Police Resources" in Cost-Benefit Analysis, M.G. Kendall, Editor, American Elsevier Publishing Company, Inc. (New York: 1971).

References (continued)

13. Beaton, W. Patrick, "The Determinants of Police Protection Expenditures," National Tax Journal, XXVII, No 2 (1974), pp. 335-349.

14. Walzer, Norman, "Economies of Scale and Municipal Police Services: The Illinois Experience," The Review of Economics and Statistics, LIV (November 1972), pp. 431-438.

15. Hirsch, Werner Z., The Economics of State and Local Government, McGraw-Hill (New York: 1970).

16. Sunley, Emil M., Jr., "Some Determinants of Government Expenditures within Metropolitan Areas," The American Journal of Economics and Sociology (October, 1971) pp. 345-363.

17. Odoni, Amedeo R., Employment and Expenditure Trends in U.S. City Police Departments, Technical Report TR-16-75, Innovative Resource Planning Project, M.I.T. Operations Research Center, Cambridge, Massachusetts, November 1975.

18. Larson, Richard C., Urban Police Patrol Analysis, The MIT Press (Cambridge, Massachusetts: 1972).

19. The National Commission on Productivity, Opportunities for Improving Productivity in Police Services, NCP73013, Washington, D.C. (1973).

20. Blumstein, Alfred and R.C. Larson, "Analysis of a Total Criminal Justice System," chapter in Analysis of Public Systems, A.W. Drake, R.L. Keeney, and P.M. Morse, editors, The MIT Press (Cambridge, Massachusetts: 1972).

III. IMPROVED ANALYSIS AND INTERPRETATION OF CERTAIN CJS STATISTICS

In this second section our focus is on system-level performance measures and statistics, many of which have multi-year consequences. They aim at improving a citizen's perceptions of the magnitude of crime victimization risk, of the long-term consequences of alternative correctional programs, and of other CJS-related issues which can be clarified by using quantitative techniques that are somewhat different from those in common use today. They are useful not only in system-level evaluation, but in other system-level considerations as well, for instance the public's view of the total magnitude of the crime problem and the amount of national resources devoted to its solution.

Public views about the dimensions of a problem, while not shaped solely by official statistics, are clearly not uninfluenced by them. Thus criminal justice planners and, indeed, all other citizens should take interest in the question: do statistics currently released about the amount of crime--and the response to it by police, courts, and corrections systems--give a graphic yet accurate view of the true state of affairs? And when the answer is "no," what changes should be made to make the numbers more informative?

We believe that many crime statistics now in circulation do not serve their ostensible purpose, which is primarily to evaluate the effectiveness of the CJS and its constituent parts. This happens largely because of three problems. One, quite widely noticed, concerns

the accuracy of much of the data; skepticism is often so great as to recall the old saw "lies, damn lies, and statistics." A second difficulty is that analyses of the data are often incomplete or inaccurate; sometimes, one suspects, the work of ideologues posing as statisticians. The third problem, somewhat more subtle, is that the popular indices about crime calculated from raw data are often inherently inappropriate--opaque at best and misleading at worst. We believe, however, that all of these problems could be greatly reduced if certain changes--all entirely feasible--were made in the ways information about crime is reported and analyzed. We attempt below to justify this premise by examining in detail some prototypical cases. Then we discuss some steps LEAA can take to ensure that accurate yet comprehensible crime statistics are freely available.

Murder Victimization

As the first example of a situation where current statistics may be inadequate, we consider the incidence of murder, clearly the most serious of the seven index crimes. For murder, the accuracy of raw data is not greatly questioned; there is widespread agreement that reported numbers of willful homicides are generally close to the mark. But the accuracy of the data has not prevented confusion about the nature and magnitude of the American homicide problem. We give two illustrations of this confusion below; we first examine a fashionable but erroneous theory about current murder trends and then discuss misconceptions caused by the widely-used murder rate statistics.

"The experts who have precisely studied homicide patterns in the United States say that the real cause of the increase is demographic rather than social." So reported The New York Times on January 1, 1972, in an article discussing the great rise in murder levels in the previous decade. The Times amplified its statement by quoting a leading criminologist as saying "the statistics are a reflection of the high birth rate in the U.S. following World War II." A logical implication of the Times analysis is that the high murder levels are a transient phenomenon, destined to be reversed because of lower birth rates in the 1950's and 1960's. A policy-related corollary of this viewpoint might be "this will pass on its own, so just sit tight."

The only problem with the Times analysis is that it is not consistent with the data [1]. Changes in the age and ethnic make-up of the U.S. population explain less then one-tenth the rise in murder that actually occurred between the early 1960's and early 1970's. We will not discuss the issue further here, except to stress that the Times' identification of demography as the "real cause" of murder growth is, to say the least, bewildering. This situation exemplifies a problem of numerous crime-related analyses: the conclusions are inconsistent with the data that supposedly spawned them.

Another problem with murder is the way its prevalence is described to the public. To express murder tolls more meaningfully, the number killed in a geographic area is often compared to its total population. The standard statistic--used by the FBI and others--is the number of murders per 100,000 residents per year. Thus, for instance, in Baltimore in 1972, the murder rate was given as 36 per 100,000.

Let us examine that statistic for a moment. Its complement is the observation that for each 100,000 residents of Baltimore, 99,964 were not murdered in 1972. This might give comfort to local residents for they might infer that, not only are they very safe over one-year periods, but their chances of ever becoming murder victims are exceedingly small.

There is, unfortunately, a potential fallacy in such reasoning. If one considered all single Americans over 18, the fraction who get married in a particular month is very small. Yet, sooner or later, the vast majority do get married. Small probabilities accumulated over long periods may yield large probabilities that events ultimately occur. And, indeed, a randomly-chosen person who lives in Baltimore from birth, confronting homicide at the 1972 level every year, will eventually die of murder with probability 1 in 38.

How many people, aware of the 36-per-100,000 figure realize the 1 in 38 statistic it implies? Evidence the authors have seen suggests the number is very small. Casual inquiries to many people in the Boston area revealed a peculiar pattern: people tended, if anything, to overestimate Boston's annual murder rate, yet to underestimate very greatly (usually by 95% or more) the corresponding murder probability for a lifelong Bostonian. An official in the crime anlaysis bureau of a large American city who spoke to one of the authors, found 1 in 1,000 plausible as an estimate of lifetime murder-risk in one of that city's most dangerous parts; the actual risk is over 40 times as high. A most fascinating misunderstanding of the annual murder rates appeared recently in, of all places,

---

* By "homicide" we mean only willful homicide, and exclude manslaughter by negligence.

The New York Times (4/28/74). In an otherwise solemn report about murder in Detroit--one of the most homicidal cities in the world--its reporter stated "if you live in Detroit you have a better than 2,000 to 1 chance of not being killed by one of your fellow citizens." He went on to note that "optimists searching for persepctive in the statistics of murder insist those odds are pretty good." [Emphasis added.]

Unfortunately, however, the estimate 1 in 2,000 for a citizen's murder probability is just Detroit's annual murder rate in the form 1 in X. It would be correct only if a Detroiter's life expectancy were exactly one year. A randomly selected Detroiter's lifetime chances of becoming a victim of willful homicide are actually about 1 in 28.

When misunderstanding of the statistic is apparently widespread, we submit that murders per 100,000 per year is not the appropriate figure for the FBI (and others) to publish. What statistic would be better? One obvious possibility is the answer to the question: assuming current patterns persist forever, what is the probability that a randomly-chosen baby, born now in region X, will eventually die of murder if (s)he lives there all his(her) life? People tend, we believe, to ask themselves an inchoate version of this question when they hear the current figures; the problem is that now they often answer very wrongly. Alternatively (or additionally) one could calculate the drop in such a baby's life expectancy because of homicide; this figure would reflect the special

tragedy that murder victims are usually young. Either of these statistics can be obtained easily from raw data; both would describe clearly and accurately the danger murder poses now.

Certain criticisms can be (and have been) raised against providing murder statistics in the form suggested above. The homicide probabilities obtained would indicate only a macroscopic "average" risk, and not reflect the great dependence of murder risk on race, sex, income, life style, etc. People who note this sometimes forget that the same criticism applies to the region-wide statistics currently in use. In any case, there is no problem in breaking populations down to as many subgroups as desired, and making the probabilistic calculation for each group.

Another line of complaint is that these descriptions of murder risk have no policy implications. This is true, in a short-term day-to-day sense, but making clear to the public how much murder there is might well raise the intensity of attempts to reduce killing.* In any case, other statistics about murder could be more directly valuable in evaluating public policies; some are described later in this chapter.

Victimization for Nonlethal Violence

For homicide, the key problem is not the accuracy of statistics, but of understanding their implications. With other violent felonies,

---

* Analogous statistics greatly changed the American public's attitude toward cigarette smoking.

however, the situation is somewhat reversed. People who learn that over 115,000 robberies, rapes, and aggravated assaults--one for each 70 residents, took place in New York City in 1972 tend to sense that, at that rate, it would not be surprising for New Yorkers to fall victim to such violence roughly once in their lives on the average. But many people apparently consider the official statistics about nonlethal violence gross underestimates, because of underreporting by the public and, sometimes, deliberate distortion by authorities for political reasons. Even former Attorney General Richardson, releasing the 1972 FBI figures, was openly skeptical of their accuracy. In such circumstances, reports of a 10% drop in assaults, for instance, are often greeted with derision.

The crime victimization surveys begun recently by the LEAA are potentially of great value in this connection. They may make the problems of underreporting by the public and manipulation of data by officials far more tractable in the future. Underreporting is dealt with quite explicitly in the surveys; those who declare themselves victims of recent crime are asked at once whether they informed the police of their trouble. And the very existence of such independent estimates of crime levels may reduce the incentive to distort official statistics, lest every new survey produce a new scandal.

But to fulfill this important role, it is necessary that the accuracy of LEAA surveys be beyond serious question. At the moment, it is not clear that this is the case. Chart 1 below compares two estimates

Chart 1

AGGRAVATED ASSAULT 1972

City	Expected # of Crimes Reported to Police (LEAA)	Actual # of Crimes on Police Blotter (FBI)	Ratio of Expected/Actual
Atlanta	2,100	2,100	1.00
Baltimore	4,100	6,400	.64
Chicago	12,700	11,200	1.13
Cleveland	2,900	2,000	1.45
Dallas	3,400	4,600	.74
Denver	2,900	1,900	1.53
Detroit	7,800	6,100	1.28
Los Angeles	13,400	15,100	.89
Newark	800	2,600	.31
New York	11,700	37,100	.32
Philadelphia	9,300	4,600	2.02
Portland	1,700	1,300	1.31
St. Louis	2,300	3,200	.72

Sources: FBI Uniform Crime Reports (1972)
LEAA Crime in Five Largest Cities and Crime in Eight American Cities (1974)

of the number of aggravated assaults reported to the police in 13 major cities in 1972. The first estimate is based on LEAA surveys in those cities in which an extrapolation was made from the sample data to estimate the number actually reported to police; the second is the official police figure reported to the FBI. It should be stressed that the LEAA figures listed have already been corrected for estimated under-reporting; in other words, one might expect the two numbers for each city to be roughly equal.

Differences between LEAA and police reports on aggravated assault in a city might be anticipated for the following reasons: (1) the LEAA estimates are based on random sampling, with its inherent potential for sampling errors, (2) people surveyed by the LEAA were asked about their experiences in the last twelve months which, in some cases, overlapped some of late 1971 or early 1973, and (3) only city residents were questioned by the LEAA, so crimes against suburbanites and out-of-towners presumably did not enter LEAA totals. But even taking these factors carefully into account, a liberal estimate of the 95% confidence interval for each ratio in the chart's final column is .80 to 1.15 (i.e., the probability a ratio falls outside the range (.80, 1.15) is about 1 in 20). But the actual percentage of ratios in the range (.80, 1.15) is not 95 but 23 (!). Similarly, there is in theory only about a 1 in 40 chance that <u>any</u> of the 13 ratios falls outside (.70, 1.22); in fact; fully 8 of them do. We have, therefore, disagreements of great statistical significance.*

---

* Points in this section are elaborated in "The NCP Victimization Surveys: A First Look" by A. Barnett and D. Kleitman, available on request.

What are the reasons for these disagreements? Peculiarities in police department records? Lack of randomness in LEAA sampling? Misunderstandings or errors of recollection by LEAA survey respondents? These and other possibilities come to mind at once. While the causes of these inconsistencies (and others that arose in data for other crimes) may never be completely understood, their existence should prompt the LEAA to perform a complete scrutiny of its surveying procedures. Controlled experiments in surveying (e.g., with a subgroup of respondents more totally assured of anonymity than respondents are now) and other possible modifications of present practices should not be excluded. The LEAA should realize that, unless its methods of polling inspire general confidence, discrepancies between its results and those of other sources will only lead to shrugged shoulders, and the relegation of LEAA data to the potpourri of other numbers widely regarded as useless.

Hopefully, these "growing pains" of the LEAA victimization reports will soon diminish. But the investigations cost roughly $2,000,000 per city, and thus clearly cannot be performed in every community every year. The most realistic approach to getting accurate crime figures on a year-to-year basis in a given community is to synthesize sensibly the insights of LEAA survey and traditional police statistics.

How should such a synthesis be made? The appropriate procedure may well vary from crime to crime; we consider below one specific offense--aggravated assault with injury. The actual results of the 1972 LEAA survey will serve as the basis for the discussion. As noted, the

accuracy of this aggravated assault data is somewhat uncertain; our purpose here is to illustrate a general procedure rather than obtain exact results about this crime.

The LEAA estimated that in 1972, 55,100 aggravated assaults with injury (hereafter AAI's) took place in the thirteen cities it examined, and that 60.1% of these felonies were reported to the police. There were some variations in reporting rates between different cities but they were not statistically significant because of sampling error. Indeed, the observed pattern of variation greatly resembled what one would anticipate from sampling randomness alone. Under the hypothesis of a true 60.1% reporting rate in each city, the expected numbers of rates more than one standard deviation and two standard deviations from 60.1% are, to the nearest integer, 4 and 1 respectively (out of a total of 13). The actual numbers were precisely 4 (Chicago, Los Angeles, New York, Portland) and 1 (Chicago).

It thus appears that, based on the survey results, the fraction of AAI's reported was quite constant over the different cities. Great disparities in city sizes and locations, in AAI rates, in ethnic compositions, and in police effectiveness (under certain criteria) were virtually irrelevant in the actual reporting patterns. Two hypotheses would seem plausible in consequence of this data: (1) the reporting rate for this felony was about 60% in 1972 in all large American cities; including those not surveyed and (2) the rate of reporting in any given city has not changed appreciably since then. Hypothesis 2 seems reasonable because changes in

any city since 1972 are probably small compared to its differences in 1972 with members of the group of surveyed cities, all of which had about the same reporting rates. The two hypotheses above would imply the approximation:

$$\begin{array}{l}\#\text{ of AAI's reported to police} \\ \text{in City X in Year Y}\end{array} \approx .6 \times \text{actual } \#$$

or, equivalently,

$$\begin{array}{l}\text{Actual \# of AAI's in City X} \\ \text{in Year Y}\end{array} \approx 1.67 \times \text{Reported } \#$$

Obviously, follow-up surveys should be performed every several years to validate (and, if necessary, revise) such hypotheses and their resulting estimates. But until contradictory results emerge, the approximation above would seem appropriate if the surveys are correct.

As noted earlier, it is not at all clear that the number of AAI's actually reported is the same as the number officially recorded.* Thus just multiplying an AAI figure from, say, the Uniform Crime Reports by 1.67 need not yield an accurate estimate of the true AAI total. But, as we said earlier, differences between the number reported and the number on police blotters may well "wither away" because of the existence of independent procedures to estimate crime levels. This salubrious result may be achieved even if surveys are conducted as rarely as once a decade in a particular locale. (We should stress, by the way, that we are discussing falsification of data as a theoretical problem; we do not believe the surveys sufficiently exact at this time to demonstrate clearly that

---

* Here we are referring to deliberate police underreporting.

distortion exists.) Hence, if present patterns continue and have been identified correctly, we foresee a time when 6,000 AAI's on police blotters can be taken confidently to reflect close to 10,000 AAI's overall. Again, the discussion above should be viewed as a somewhat hypothetical simple example of how general insights can be drawn from particular survey results.

We are hopeful that raw police figures on each crime type and the corresponding estimates of true totals including unreported crime will appear in reports released in the future. The "correction factors" used should be justified carefully; they may often be far more complicated than the simple constant 1.67 that worked so well for the AAI data.

Once these corrections are made, estimates can be prepared, for instance, of the expected total number of victimizations a randomly-chosen baby, who lives his life in region X at the current risk level, will sustain. Specific breakdowns by race and sex, income, etc., can be made. We have noted that annual rates per 100,000 people are not as potentially misleading for other felonies as for murder, but the lifetime statistic above still seems a desirable quantity for explicit calculation.

Evaluating Deterrence Policies

So far we have discussed statistics about the prevalence of crime. But transparent statistics are also needed to describe and evaluate society's response to crime: the efficiency of the police, the effectiveness of the courts, the corrective effect of correctional programs, the

deterrent effect of the laws. The problems in this area are varied and complex; we will not attempt an exhaustive discussion here. We will consider instead a "random sampling" of topics and some new statistics that are potentially illuminating. Our suggestions on the subjects below, however, often imply what we would recommend in areas not directly mentioned.

The deterrent effect of particular measures against crime is a subject of bitter and widespread controversy. It would be fatuous to suggest that statistical measures can end such debate, but they can help clarify the raw data that is available. We turn once again to the particular crime murder. Public interest is currently high about possible steps to reduce killing, and might well increase greatly as true victimization chances become known. An imminent Supreme Court decision may restore the death penalty for first degree murder in over half the states in the Union. Gun control legislation is now pending at many levels of government. Before any such measures come into being--when there could not be accusations of "ex post facto" criteria--it would be desirable to create a statistical framework to help evaluate their deterrent effects.

Many laws designed to reduce willful homicide may arise at the state level. And while many states seem ready to adopt new regulations, many others do not. Thus comparing changes in murder levels in different states may give some indication whether new laws are working. A challenging question is how such comparisons should be made.

A crude statistic for comparing two states is the ratio of their annual homicide rates; such a statistic, however, may not be sensitive enough to evaluate new laws in one of them. FBI homicide data makes clear that, over the nation, murder victimization rates are unusually high in cities, among blacks, and among the poor. (These are hardly distinct categories; many blacks are poor and live in cities.) Thus if, between two points in time, two states change differently with respect to degree of urbanization, ethnic composition, or economic conditions, one would anticipate changes in their murder rates ratio quite independent of changes in laws or forms of penalty. One desires that any comparative statistic in use should automatically correct for such changes (and others).

With appropriate care, one can introduce an "adjusted" ratio of murder rates between states that "weeds out" the effects of overall national patterns. This ratio will still often differ greatly from the number 1 because of nonquantifiable local trends (e.g., "tradition of violence") superimposed on the national. Consider, for instance, the states Indiana and New Hampshire. In 1972, the per capita murder rate in Indiana was 3.53 times that of New Hampshire; when, however, one gives proper weight to the fact that Indiana has six cities (Indianapolis, Gary, Hammond, Fort Wayne, South Bend, and Evansville) with over 100,000 people compared to New Hampshire's zero, and the (somewhat correlated) fact that Indiana's ethnic distribution differs substantially from New Hampshire's, one obtains as a first estimate of an adjusted homicide ratio the number 1.49. This number (actually its more precise counterpart)

might merit watching if Indiana restores the death penalty but New Hampshire does not.

There are several pitfalls one must avoid in calculating these adjusted ratios. Beyond noting changes in urbanization, economics, age, race, and sex distributions, one should specifically correct for migrations from other states with different "traditions of violence." And while one wants to extirpate interstate differences reflecting national patterns, one must not forget that national patterns may themselves change because of the laws whose effects are being investigated. (E.g., if urban states ban handguns in greater proportion than rural ones, and if the ban actually does diminish murder, the relatively higher homicide risk of city dwellers may be reduced. If, because of this trend, one used a smaller correction for urbanization in the adjusted murder ratio, one could erroneously wind up concluding that the ban was ineffective.) One must also be careful to distinguish meaningful changes in the ratios from the effects of random fluctuations. (New Hampshire had only 13 murders in 1972; random fluctuation about this figure could change the ratio with Indiana 20% or more between two consecutive years.) But if prepared carefully by statisticians, these ratios (and charts describing their evolution) could graphically inform the people if any antimurder measures are actually working.

Evaluating Correctional Measures

A panoply of statistics about the operations of police and courts are floating about; perhaps too many for citizens to absorb. For each crime, we hear numbers about the fraction of offenses cleared by arrest, about bail policies and pretrial detention, about sentencing policies for those convicted. One major question these numbers concern is: how much time incarcerated is actually meted out to people who commit a given crime?

A statistic particularly suited to answering the last question is the average amount of time incarcerated, distributed for each offense of type X. Unsolved crimes would contribute zeroes to this average; prison terms and pretrial jailings would enter the statistic through their average lengths. Thus, for example, if 40% of type X crimes are solved (i.e., end with convictions) and the expected time incarcerated is two years per offender, the average time incarcerated per offense is $2 \times .4 = .8$ years. This is the type of expected value statistic calculated at numerous stages of the Blumstein-Larson CJS model [2], which is now made operational as the JUSSIM model.

This statistic combines sensibly the probability of arrest and the expected consequences of arrest, two factors presumably considered by many would-be criminals. If this quantity were calculated for several different localities, interesting new correlations of incarceration levels and crime levels might emerge. One could also calculate an

average cost level for each victim of a type-X crime, where cost is measured in time spent recovering from injuries, time spent at work recouping financial losses, time spent testifying at trials, etc.

We would like at this point to propose a new statistic R_x defined by

$$R_x = \frac{\text{average cost per criminal for each offense of crime x}}{\text{average cost per victim for each offense of crime x}}$$

As described above, the costs can be computed in such common units as time. From available data, such a ratio can be computed for each felony. One could argue that--for reasons related to justice quite apart from deterrence--such a ratio should not be allowed to fall below one (except, perhaps, for murder). Thus if a calculated R_x is smaller than one, this might be an argument for a change in public policy, such as public compensation of crime victims, longer average sentences for convicted offenders or greater resources devoted to raising arrest probabilities. Great differences in the ratios for different crimes might suggest inequities in the sanction structure. Other possible applications of such statistics come to mind.

We do not wish to suggest that these ratios have mystical properties. But they summarize vast amounts of data from throughout the criminal justice system and provide one digestible index of how the system is working. Devising and calculating such system-wide performance measures should be a high priority for the LEAA.

Recidivism

One word seems to pop up in every discussion of correctional programs: recidivism. And the statistic that seems to dominate discourse is the probability that a given offender will commit crime again. While of obvious interest to behaviorists and the offender's parents, it is not clear that this statistic is very useful for the public-at-large. Recidivism probabilities are inherently ambiguous quantities in that the same number can describe vastly different situations, as we illustrate below.

We define recidivism rate for crime x as the probability that a randomly-chosen person who has committed one or more offenses of x will commit one or more additional offenses in the future. Under this definition, each of the two situations below is characterized by a 40% recidivism rate.

Case 1:* Each past offender for crime x has a .4 probability of committing one or more additional crimes.

Case 2: Each first offender has a .8 probability of committing a second crime. No second offenders ever recidivate. A first offender who recidivates is equally likely to commit his second crime at all times between his first offense and his death. First offenders are generated at a constant rate.

In Case 1, the average number of recidivist crimes per offender is 1.67 and 63% of the incidents of crime x are committed by recidivists. The corresponding figures for Case 2 are .8 and 44%. The difference between these two cases have direct policy implications. Measures that reduced the

---

* See, for example, Ref. [3].

recidivism rate by half (i.e., to 20%) would reduce overall offense levels by 22% in Case 2, but only 16% in Case 1. Yet all of these differences are ignored by the "one-dimensional" recidivism rate statistics in current use. (Other definitions of recidivism rate exist, but they suffer from similar problems.)

We pause now to reflect on something ironic. For murder, probabilities of victimization seem more useful than annual rates, yet the popular statistic is the annual rate. For recidivism, an annual rate seems more useful than a repeating-probability, yet the probability is the widely-quoted figure. This curious phenomenon makes sense if we consider that, given the way data are collected, it is easier to find annual murder rates than victimization chances while recidivism probabilities are more readily obtained than other indices on the subject. Still, for something as serious as crime, ease of calculation might not be the best criterion for choosing statistical measures.

Recidivism rates, however, are but a means to a statistical end. In comparing two correctional programs, a good question to answer is: given the current pattern for generating first-offenders, and given the recidivism patterns for each program, what is the expected difference in total crime over the next (say) 50 years? The difference should be expressed both in absolute magnitude and as a percentage. If the more successful program is also somehow the more costly, this statistic would give some reasonably precise idea what the extra expense might buy.

Conclusions

The previous sections of this chapter have attempted to argue the need for certain programs. Specifically, we believe that LEAA should adopt as long-range goals:

1) The development and dissemination of new statistical indices to illuminate raw data about crime levels and to help evaluate the effectiveness of the police, courts, and corrections systems.

The "crime clocks" already included in the FBI Uniform Crime Reports set a precedent for efforts of this kind. The LEAA should work with the FBI and other relevant agencies to provide for implementation of the new measures in official reports, not generally as replacements for current statistics but in addition to them. In the shorter run, we feel the LEAA should support research efforts to develop appropriate statistical indicators in all areas and to prepare "computer packages" to allow their calculation by government agencies.

2) The improvement of the conduct of LEAA victimization surveys to the point where their accuracy cannot seriously be challenged.

An intense effort should be made to identify and avoid the sometimes very subtle problems in the massive survey program. While statisticians, political scientists, computer experts, sociologists, etc., are clearly relevant to such an effort, so are police chiefs, elected officials and not a few individuals who are genuinely "street smart."

3) The development of a set of statistical testing procedures to measure quantifiable effects of innovations in the area of criminal justice.

Many of the changes in penalty structures, in corrections policies, and in court operations that have been proposed are intended to have specific quantifiable effects. LEAA should encourage the systematic recording and reporting of certain statistics that, while not tied specifically to any given program, will be of direct help in assessing the program's effectiveness. [Unlike the indices described in 1) above, such statistics may not be inherently "illuminating," but watching directions in which they move over time may be useful.] The adjusted annual reporting of interstate homicide ratios is a specific example of such a de facto statistical test. We believe the LEAA should support efforts by researchers to devise such procedures, and should work to ensure their systematic use in appropriate annual reports.

References

1. A. Barnett, D. Kleitman, R. Larson, "On Urban Homicide, A Statistical Analysis," Journal of Criminal Justice, Summer, 1975.

2. A. Blumstein, R. Larson, "Models of A Total Criminal Justice System," Operations Research, Vol. 17, March-April, 1969, pp. 199-232.

3. A. Blumstein, R. Larson, "Problems in Modeling and Measuring Recidivism," Journal of Research in Crime and Delinquency, Vol. 8, No. 2, 1971, pp. 124-132.

IV. OPERATIONAL PERFORMANCE MEASURES FOR EVALUATION

In this third section, we shift focus from system-level and multi-year considerations to problems of evaluation at the agency level. In particular, we address the role of quantitative models in the evaluation process--with particular reference to police departments, but also LEAA state planning agencies and programs they may undertake in any component of the CJS.

We identify two different types of evaluation at the local or state level--evaluation of experimental programs and evaluation of on-going (routine) day-to-day operations. First we discuss experimental evaluations, then day-to-day evaluation, and finally some steps that could be taken to develop these ideas toward implementation.

Evaluation of CJS Experiments

In recent years, we have seen a trend developing toward the acceptance and use of the scientific method in acquiring knowledge about the CJS and its component parts. This entails the identification of a problem area, the listing of conjectures or hypotheses regarding system structure and operation, the design of an experiment to test these hypotheses, the execution of the experiment, and the evaluation of results. While the term "evaluation" is often only identified with the last of these steps, we will utilize a broader meaning referring to the entire process of experimental design, execution, monitoring, and formal (final) evaluation.

In performing an experiment in part of the CJS, the entire jurisdiction under consideration (e.g., city, county, state) serves, in effect, as a laboratory and the number of "actors" and resources used is usually large. Thus, the experiments tend to be extremely expensive and time-consuming. This, plus the fact that the experiments are non-repeatable (at least under identical conditions), make it very important that the design of the experiments be extremely well thought out. Besides, the publicity that such experiments receive--just by virtue of being performed in the field--and the consequent major impact that they are likely to have on local agencies throughout the country, raise the stakes of assuring against serious experimental errors to unusually high levels.

It is our strong belief that the quantitative models developed in recent years at MIT, Rand, Carnegie Mellon and elsewhere can play an important role in assisting in every phase of the aforementioned experiments, from initial planning to the final evaluation of results. To implement this concept, it would be necessary to perform research on the use of quantitative models in experimental design and to report the results in nontechnical handbooks for CJS personnel. The emphasis would be on the description of a normative experimental approach in which models would continually be used by the experimenters to design, monitor, evaluate, and revise the experimental procedures in the field. It would be useful to illustrate this approach as well as to make a clear case for the usefulness of the quantitative models in this respect, through a detailed

review of some recently performed and highly important experiments, such as the Kansas City Preventive Patrol experiment [1].

One type of experiment that could provide a focus for the work would typically deal with reallocation of resources of a police department and/or with a revision of the operating policies of these resources. Listed below, as an illustration of the types of issues which could be investigated, are a few examples of the use of quantitative models in such experimental contexts:

(a) Simple, rule-of-thumb techniques can be used to check some of the initial basic premises of the experiment. For instance, whether four police cars in a district make available four times as much preventive patrol time as a single car.

(b) More sophisticated analyses based on the newly developed hypercube queuing model [2] can be employed to try to assess in detail whether the conditions under which an experiment would be conducted actually conform to the conditions that the experimenters have in mind.

(c) During the performance of the experiment itself, a parallel use of quantitative models and tools is helpful

in monitoring progress (through comparison with field results). In fact, one of the primary uses of these tools is an "adaptive feedback" mechanisms, that is, as means of adjusting experimental design parameters in the desired direction.

(d) Finally, and very importantly, the significance of the results of an experiment should not be evaluated (as, unfortunately, they are all too often) solely on the basis of "before" and "after" baseline data. Instead, a more proper method of evaluation is a comparison between the results predicted a priori (on the basis of whatever theories or beliefs prevail prior to the conduct of the experiment) and the actual measurements in the field. Quantitative models provide a unique tool for producing in detail the a priori predictions necessary for the successful conduct of the experiment.

As an example of the use of models to produce causal relationships in evaluative studies, consider a simple example from the corrections component of the criminal justice system (CJS). It has been documented (by Wolfgang and Figlio [3] and others) that multiple offenses (measured by arrest or some other formal contact with the criminal justice system) can be modeled as a Markov process to predict the future criminal career profiles of offenders. Thus, if we say that an offender leaving the CJS

has a recidivism (repeat) probability of X, then the average number of repeats over the criminal career of the offender can be shown (using Markov analysis) to be $1/(1 - X)$. Now consider an experimental corrections program that is to be evaluated. Suppose that the "failure" rate from the corrections system prior to the experiment was 90% (a high value, but one found in some correctional institutions). After the experiment, the "failure" rate was only reduced to 80%. An evaluator comparing the before and after "baseline data" may be tempted to say that the Y millions of dollars spent on the experiment could hardly be justified since there is only an 11% decrease (from 90 to 80%) in the key outcome variable. However, using the model for future criminal career behavior, one finds that the average number of future repeats per offender is reduced from $1/(1 - 0.9) = 10$ to $1/(1 - 0.8) = 5$, fully a 50% decrease. Here the model suggests a simple manipulation of the baseline data to predict a more fundamental outcome measure--the number of future contacts with the CJS. In the same way, models can be used in the evaluation of experiments in other parts of the CJS.

Returning to the police area, in our interactions with police personnel we have found that even the simplest of quantitative models of police patrol can provide insight into operational behavior of the system not readily available from other means. In several instances these insights have revealed the inadequacies of long-held rules of thumb or points of view. We give four examples:

Example 1: Statement: "Doubling the amount of patrol resources doubles the amount of preventive patrol."

The reasoning behind this statement is widely accepted in police circles today. Yet it is clearly incorrect as we can show by a simple counterexample: Consider an isolated patrol sector with one unit assigned to it. Suppose the average call-for-service workload per eight-hour tour is four hours and the time on preventive patrol is four hours. Doubling the patrol resources (i.e., adding a second unit) would result in four hours (total) of call-for-service time and 12 hours of preventive patrol time. Here a doubling of resources yields a tripling of preventive patrol effort. Other examples can be presented in which doubling the patrol resources increases the amount of patrol by a factor of 4, 5 or even 10. The simple traditional argument is incorrect because it neglects the effect of units spending time on calls for service.

Example 2: Statement: "Doubling the amount of patrol resources halves average travel time."

This "linear" argument is incorrect because average travel time varies as the square root of the resources allocated per square mile. A doubling of resources decreases average travel time by about 30%, not 50%.

Example 3: Statement: "The fraction of dispatches which are intersector (cross-sector) dispatches is usually small enough to ignore in most cases."

In police circles it is preferable to keep intersector dispatches at a minimum in order to maximize an officer's contact with his "own" sector, thereby enhancing "sector identity." In early traditional police administration texts (such as O.W. Wilson's), intersector dispatches were often ignored, largely because at the first time of writing of these texts (usually the 1930's or 1940's) the workloads of urban police departments were nowhere near the workloads experienced today; hence, the sector unit was most often available to respond to emergencies that arose in "its" sector. Nowadays, if a police precinct has a call-for-service workload causing its units to be busy, say, 55% of the time, then at least 55% of all dispatches are intersector dispatches (hardly an insignificant amount). The correct argument is simple: Consider a randomly selected call for service. With probability 0.55 it will occur when the sector unit is busy on a previous call; thus, with probability 0.55, it will require an out-of-sector unit. But, this applies to all sectors and all calls, and thus the statement is true.

Example 4: Statement: "The workloads of units will be balanced (i.e., equalized) if the workloads of their respective sectors are balanced."

This statement or its equivalent appears in nearly all of the classic police administration texts. And, due to intersector dispatches, it is wrong. In our recent MIT work, we have derived useful sector configurations in which the unit assigned to the least busy sector had the greatest workload (among all units) and (simultaneously) the unit assigned to the busiest sector had the least workload. This type of behavior can be modelled very well with hypercube-type models.

To place the required new work in the context of traditional evaluative research, we refer to E.A. Suchman's book, Evaluative Research [4]:

> Many of the newer techniques and research designs, such as...operations research, have not yet been adequately incorporated into the planning and conduct of evaluation studies. The valid interpretation and successful application of findings, while the sine qua non of evaluation, is often grossly neglected and misunderstood. These are only some of the important problems and needs in the field of evaluation. Undoubtedly one of the reasons that many of the current attempts at evaluation have seemed weak and invalid is the lack of any clear-cut theory or method to support the research.

We are postulating the need to develop a quantitatively based conceptual framework (and details of the conceptualization necessary to implement it

in practice) for evaluating experiments in the CJS. The required quantitative framework would focus on three distinct phases of the experiment: 1) design, 2) execution, and 3) evaluation. The methods employed should be useful for any of the standard experimental designs: the "one-shot" case study ("XO" in Campbell's [5] notation); the one-group, pre-test, post-test design (O_1XO_2); the static group comparison ($\begin{matrix} XO_1 \\ O_2 \end{matrix}$); and the pre-test, post-test control group design ($\begin{matrix} O_1 X O_2 \\ O_3 \quad O_4 \end{matrix}$).

The design and evaluation phases would be constructed so as to use quantitative models (wherever relevant) to predict causal relationships among the variables. The need for this is overwhelming, as pointed out by Suchman [4]:

> The primary reliance of the evaluation guides upon existing records discourages the utilization of research for the collection and analysis of data. This means that in most cases one deals with statistics obtained from samples of biased or unknown representativeness, with available rather than pertinent data, with unreliable and invalid measures, and with relationships whose causal connections are not at all clear. [Emphasis added.]

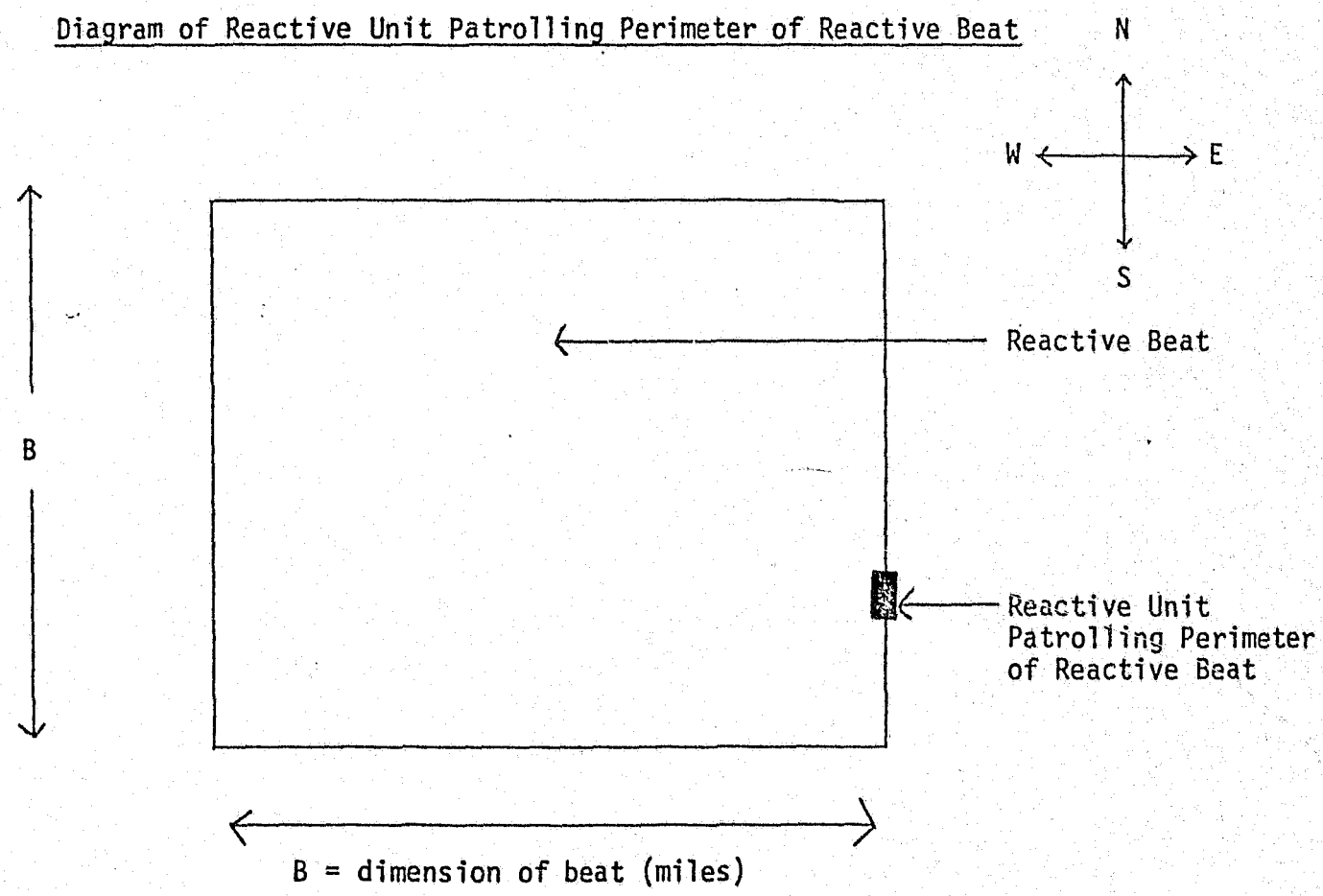
Simple "back-of-the-envelope" reasoning, such as that illustrated with the four examples above, would first be used in setting up the experimental and control environments. In the context of the Kansas City patrol experiment, such reasoning could be used to predict a priori the response

times, patrol levels, and patrol workloads in each of the three experimental areas. Then, once a reasonable design is selected on the basis of crude models, more precise models (such as the hypercube model or the JUSSIM Model) would be used to fine-tune or calibrate the experimental design. So, in the design phase, quantitative models would be used to predict causal relationships between control variables (resources and deployment techniques) and empirical performance measures (response time, workloads, number of cross sector dispatches, patrol frequencies, etc.).

As an example of using such models in the design phase of an experiment, Figure 4.1 depicts an idealized square "reactive beat" in the Kansas City Preventive Patrol Experiment. Each such beat (there were five in the Experiment) receives no conventional preventive patrol coverage; the closest unit, called the "reactive car," patrols the perimeter of the beat and is dispatched into the beat to respond to calls for service. Back-of-the-envelope models can be used to predict the average distance travelled within the beat by the reactive car as a result of responding to a call for service there. If B is the length (in miles) of a side of the beat, the result is that the average distance $\overline{d}$ is the sum of (5/6)B, representing average travel distance to the incident, and (1/2)B, representing average travel distance from the incident. Thus, $\overline{d}$ = (4/3)B. This type of analysis can be continued to predict the total mileage travelled within the reactive beats both before and during the Experiment. The analysis suggests that the experimental design utilized in Kansas City did not greatly reduce total patrol

Figure 4.1

Diagram of Reactive Unit Patrolling Perimeter of Reactive Beat



visibility in the reactive beats, especially during summer evening hours, when citizen opportunities for judging police presence are high [6]. We are suggesting that such analyses should be done with this type of experiment before the fact, in order to assure that the experimental design yields the type of operations desired.

This proposed evaluation method appears to offer considerable advantage over one that does not incorporate such causal relationships. Without them, the initial design is not likely to produce the experimental environment desired. Furthermore, one would have no way of predicting the likely amount of increase or decrease of a particular performance measure as a result of the experimental conditions.

In a directly parallel manner, the evaluation phase (the third phase) would use quantitative models to compare the observed values of performance measures to those predicted by the models. Discrepancies could quickly point the way to aspects of the experiment that did not operate in the way intended. Or they could indicate phenomena that were heretofore unknown and therefore not incorporated in the models. At the least, the models would provide a rigorous basis for considering the quantitative outcomes of the experiment. The qualitative outcomes would still have to be examined carefully, utilizing the standard techniques of evaluation research.

It is possible that some of the insights gained from the quantitative modeling component of the evaluation could shed light on results of the qualitative evaluation. As an example, models predict that dispatching the closest police vehicle (using an automatic vehicle location system) greatly increases the amount of cross-sector dispatching. This, in turn, could help explain the negative attitudes of police officers interviewed as part of the qualitative evaluation.

The quantitative framework to be developed for the second phase--the execution of the experiment--would recognize that uncertainties in the real world (due to unanticipated responses from agency personnel, or from citizens, or due to mechanical limitations of experimental hardware, or due to lack of consideration of one or more important environmental factors) usually cause one to change the "operating rules" of an experiment during its execution. In a patrol experiment, for instance, it would be very unwise for the planner to ignore cries from citizens' groups who claim that they are receiving inadequate police protection as a result of the experiment. Yet, the planner faces a dilemma--if he changes the operation of the experiment, he runs the risk of destroying any chance of successful evaluation.

Thus, in general, when one is planning the sequence of events in an experiment, the planning should take into account unexpected (or only partially expected) events which may arise during the course of the experiment and which, if ignored, could drastically reduce the chance of successfully evaluating the experiment. In the planning literature, this

dilemma focuses around the issue of master planning vs. contingency planning. As another example, in a Police experiment, a planner following a master plan would lay out the schedule and accomplishments of the experiment in a hard and fast way, not allowing for information learned during an early part of the experiment to influence the conduct of a later part. An experiment allowing for contingencies--such as adverse responses from neighborhood groups, suggestions from patrolmen, an increase in the workload due to publicity--would consider and incorporate these inputs throughout the program, thereby having a greater chance to complete the experiment successfully and to evaluate the outcome.

The need for such a flexible environment for the execution of an experiment has been noted by several authors [7], [8], [9], [10]. McCaskey, while arguing for a contingency approach to planning, incorrectly identifies such an approach with "informal" or "intuitive" procedures, as contrasted to the "formal"procedures associated with a master planning approach. Weiss identifies the "tendency of the program to change while it is being evaluated" (p. 93) as one of the big problems in evaluating the experiment. Following Parsell, Weiss recommends that the evaluator "...develop a dynamic rather that a static model of the program to categorize it in terms of its movement as well as its conceptual location." (p. 94) However, no formal procedures are offered for carrying out this process.

The required work in this area would develop a methodology to incorporate feedback from the experimental environment to affect the experimental design. It would use tools from operations research--particularly those associated with short-term and longer-term planning (such as dynamic programming, decision tree analysis, and Markov decision processes) to plan for such contingencies prior to implementation of the experiment and to assist in structuring an evaluation plan that anticipates adaptive changes in the experimental design.

Formally, the resulting experimental design would be a matrix of contingency plans, columns corresponding to discrete time periods and rows corresponding to the state of the feedback process. Each entry would be of the form, "What I (the experimental designer) would do if I were here (here determined by the particular feedback received)." Informally, one would hope to extract some general properties from the solutions to this type of problem to provide useful guidelines to evaluators who do not have the time or resources to apply the formal method.

Evaluation of Routine Operations

While the previous section has discussed the need for formal, quantitative tools for experimental evaluations, it is perhaps even more

important to monitor and evaluate routine, day-to-day operations. Here, too, quantitative models can play a useful role.

Unlike the area of experimental evaluation, we see no need for further methodological development prior to incorporating a significant number of quantitative models in day-to-day evaluation.* The major tools are available. The key impediment to their implementation up to this time is, we believe, their perceived inaccessibility due to mathematical notation, overly formalized and technical presentations, and the frequent need for computer assistance.

As illustrative uses of quantitative models in day-to-day evaluation, consider the following three situations: 1) evaluating the performance of telephone operators at a police 911 facility; 2) predicting the recidivism profile of a convicted offender as a function of sentence type (from the judge's point of view); 3) evaluating the utilization of the jury pool called up in a particular month.

In a 911 emergency call handling facility, the key type of quantitative model that is relevant is a queuing model. Such a model of the 911 system predicts, as a function of the number of operators assigned and their skills, the delays that can be anticipated by incoming callers and the average workloads (i.e., fractions of time busy) of the operators. A standard queuing model was applied to this problem in the New York City

---

* This is not to say that all relevant methodologies have been developed. Rather, a number now exist that are implementable.

Police Department in 1968 [11], and has been used for many years since then as the primary basis for scheduling operators and evaluating their performance (and the performance of their supervisors). This same procedure could be utilized by any other police department if a nontechnical handbook on its implementation were widely available.

In the second example, a judge is repeatedly confronted with the problem of which sentence (or "correctional alternative") to select for a convicted offender. Here, for the classes of offenders which occur often in a statistical sense, it would be useful to provide a judge contemplating a sentencing decision an estimate of the expected recidivism profile of the offender for each alternative being considered. In this way the judge can eliminate inconsistencies in his sentencing practices, greatly reduce inequities, and (hopefully) develop a store of knowledge of the effects of various correctional alternatives so as to choose the most reasonable one for each type of offender. The quantitative modeling required here entails relatively simple categorization of offender classes and detailed statistical analysis of former sentencing decisions (along with the recidivism behavior of those sentenced). The resulting probabilistic model is a simple "expected value" model, one that could be implemented on computer systems containing criminal career profiles. Such a capability enhances a judge's ability to evaluate his own alternative sentencing policies.

In the third and final example, consider the problem of calling up a jury pool for a particular month. Usually, from a larger sample of eligible citizens (selected for that year), a number N_i is selected to be in the pool for the i^{th} month. The magnitude of N_i depends critically on the estimated court workload to be incurred during the i^{th} month. Too often the number is made too large, so scores of citizens sit around the court house all day with very little to do, incurring significant social costs (in terms of lost time on the job, jury pay, and disillusionment with the judicial system). Occasionally the number is made too small, resulting in delayed trials and jury screenings. However , here too, rather simple mathematical models can be employed to estimate better the monthly demand for jurors, thereby balancing the costs of oversupply and undersupply of jurors. We have already demonstrated the feasibility of this approach in a court in the Boston area, but further documentation is required to make the technique generally available.

Recommendations

For evaluations of experimental programs and/or day-to-day operations that require quantitative performance measures and models, the following recommendations appear appropriate:

(1) LEAA should support research that would lead to improved methods for conducting and evaluating LEAA-sponsored experiments. This research would include methods for experimental program design, execution, monitoring, and

after-the-fact evaluation. The product of this work, in addition to being useful in general social science applications, would focus particularly on the unique problems one encounters in CJS and LEAA experiments. Hopefully, in addition to research reports, the product would include a handbook for CJS and LEAA personnel contemplating the conduct of experimental programs.

(2) LEAA should select an already completed CJS experiment (e.g., the Kansas City Preventive Patrol Experiment) and test the developed methodology in an "after-the-fact" manner. As a result of this process, the methodology may be modified or changed in several ways. At that point LEAA might consider employing the methodology in an on-going test case to discover its usefulness in real-time applications.

(3) To facilitate evaluations of day-to-day operations, LEAA should select a small number of common operational problem areas (e.g., scheduling of 911 personnel, jury pool selections) that are conducive to improvement through the use of quantitative performance measures, methods, and models. For each such area LEAA should support the writing of a handbook to be used by the relevant agency personnel in implementing the technique in their own agency. A limited number of such implementations should be evaluated, and if successful according to the evaluation criteria, then other common operational problem areas should be tackled in the same way.

References

1. G. Kelling, T. Pate, D. Dieckman, and C. Brown, "The Kansas City Preventive Patrol Experiment: A Summary Report," Police Foundation, Washington, D.C., 1974.

2. R.C. Larson, "A Hypercube Queuing Model for Facility Location and Redistricting in Urban Emergency Services," Computers and Operations Research, Vol. 1, pp. 67-95, 1974.

3. M. Wolfgang and R. Figlio, Delinquency in a Birth Cohort, University of Chicago Press, 1972.

4. E. A. Suchman, Evaluative Research, Basic Books, New York, 1967 (reprinted through 1973).

5. D.T. Campbell, and J.C. Stanley, "Experimental and Quasi-Experimental Designs for Research and Teaching," in Gage, N.L., ed. Handbook of Research on Teaching, Rand McNally and Co., Chicago, 1963, pp. 171-246.

6. R.C. Larson, "What Happened to Patrol Operations in Kansas City? A Review of the Kansas City Preventive Patrol Experiment," to appear in the Journal of Criminal Justice, Winter, 1975-1976.

7. Michael B. McCaskey, "A Contingency Approach to Planning: Planning with Goals and Planning without Goals." Academy of Management Journal, Vol. 17, No. 2., 1974, pp. 281-291.

8. Alfred P. Parsell, "Dynamic Evaluation: The Systems Approach to Action Research," SP-2423, Systems Development Corporation, Santa Monica, California, 1966.

9. Michael D. Maltz, "Evaluation of Crime Control Programs." U.S. Department of Justice, Law Enforcement Assistance Administration, National Institute of Law Enforcement and Criminal Justice, April, 1972.

10. Carol H. Weiss, Evaluation Research: Methods of Assessing Program Effectiveness. Prentice-Hall, New Jersey, 1972.

11. R.C. Larson, "Improving the Effectiveness of New York City's 911," in Analysis of Public Systems, A. W. Drake, R. L. Keeney, P. M. Morse, eds. The MIT Press, 1972, pp. 151-180.

V. GENERAL RECOMMENDATIONS IN THE AREA OF EVALUATION

While the previous three sections have focused on a variety of specific evaluation areas of relevance to LEAA and CJS planners, researchers and managers, we offer here several additional suggestions that also bear on the evaluation issue.

Dissemination

Despite recent and current attempts at disseminating the results of "successful" projects and at pressuring agencies to include evaluation design in their overall program design, there is still much to be done in communicating the results of LEAA-funded programs to the potential user community. Evaluation now is often thought of as a 3% to 6% "add-on" or "surtax"--a price to be paid in order to obtain an LEAA grant. And many evaluation designs are shoddy and eventually never carried out. Why not require dissemination of all evaluations of LEAA-funded projects through some yet-to-be-created national medium? This could be an LEAA newsletter or, preferably, a popularized research journal with its own board of editors. Requiring such dissemination would make each LEAA project very visible to the user community--with regard to design, execution, outcome, and evaluation. Such a step may bring about an accountability for the outcome of LEAA projects which is now largely nonexistent.

Evaluation Workshops

Most people--within the CJS or LEAA or any other governmental agency--have not been exposed to a formal presentation of evaluation methodologies. No wonder, then, that many "forced" evaluations are poor in design and execution. Perhaps the LEAA should consider running a series of evaluation workshops around the country--probably at least one in each of the LEAA regions. Each workshop could last anywhere from one or two days to two weeks, depending on the audience and depth of material to be covered.

It might also be appropriate to encourage universities receiving LEAA funds in a criminal justice training program to offer courses in various facets of criminal justice evaluation. Perhaps even the IACP could be encouraged to present concepts of evaluation in several of their workshops that are presented yearly around the United States.

Interfacing Evaluation Data with New Technologies

The LEAA has been funding for some time now various types of new information processing technologies to improve the operational effectiveness of parts of the CJS. These include computer-assisted dispatch (CAD) systems, automatic vehicle monitoring (AVM) systems, computerized criminal history files, in-the-field inquiry systems (for stolen automobiles, wanted persons, etc.), computer-assisted court scheduling systems, etc. Yet it is only the exceptional implementation of one of

these technologies that processes the data in the system that is relevant to ongoing management and evaluation of activities. For instance, a CAD system includes within its computerized files the most up-to-date information possible relating to management, allocation, and evaluation of patrol forces in the field. Yet few (if any) CAD systems to date utilize this information by reducing it to understandable form and then feeding it back in a timely manner to relevant agency decision makers for routine day-to-day management and evaluation.

Perhaps NILEC's Office of Evaluation should take steps to encourage CJS technology designers and consumers to utilize management- and evaluation-oriented data which are now usually viewed as a (neglected) by-product of the technology. These steps could take the form of grant requirements, of educating both the designers and consumers, and of funding exemplary projects that propose to utilize the evaluation-oriented data.

AN SPA EVALUATION KIT

While we are generally opposed to adding to the already-too-numerous bureaucratic requirements associated with LEAA grants (or grants of many other federal agencies, for that matter), it might be reasonable to suppose that evaluation will not be carried out on the majority of SPA-funded projects unless an easily implemented mechanism is set up. One such mechanism could be a format for describing the evaluation plan

for each proposed project. The format, which might vary by type of program (say, police, courts, corrections), could take the form of two or three additional "boiler-plate" pages to fill out in the formal grant application papers. While we are convinced that this step in itself will not bring about usable evaluations of LEAA programs, perhaps without such a formal requirement "evaluation" will be taken as this year's fad word (hopefully to be replaced by something else next year).

The formats of these pages could be motivated and presented in a NILEC-produced "SPA Evaluation Kit," which would discuss many of the general issues of evaluation as well as illustrate detailed versions of the forms.

A PROGRAM OF RESEARCH ON PERFORMANCE MEASUREMENT AND EVALUATION FOR THE CRIMINAL JUSTICE SYSTEM

prepared by

Robert C. Lind

in collaboration with
G. Robert Blakey and Richard H. Blum

for the

Office of Evaluation
National Institute of Law Enforcement
and Criminal Justice
Under Contract 5-0968-J-LEAA

TABLE OF CONTENTS

Preface

This Report was produced for the National Institute of Law Enforcement and Criminal Justice under contract order No. 5-0968-J-LEAA by Professor Robert Lind in collaboration with Professors G. Robert Blakey and Richard Blum. The work and ideas that have gone into this Report represent the joint efforts of these three men, however, the Report was prepared by Professor Robert Lind and is his responsibility alone. While he has drawn heavily on the ideas and comments he received from Professors Blakey and Blum, no attempt was made to obtain a concensus on the wording or content of the main body of the Report.

Professor Lind received written comments from both Professors Blakey and Blum which constitute substantial documents in their own right. The paper from Professor Blum in fact constitutes a complete report in itself. These papers are included as appendices to this Report and should be read with care as many of the insights, perspectives and recommendations they contain are not fully captured or reproduced in the main body of the Report.

The work plan was as follows: Professor Lind asked Professors Blakey and Blum to think about the statement of work and develop some preliminary ideas for discussion. On the weekend of April 12 and 13, Professor Blakey, Blum and Lind met in Ithaca and spent two days discussing the role of measurement and evaluation as it pertains to the criminal justice system. Professor Blakey and Blum agreed to prepare separate documents to be submitted to Professor Lind for inclusion in the final report. These documents were to reflect their own thinking on the topics that had been discussed at the April meeting. The Report is based on the discussion of this meeting, on the written input from Professors Blakey and Blum, and on considerable independent research and thought by Professor Lind on the problem of measurement for criminal justice evaluation. A report of several hundred pages could easily have been produced; however, we felt a more concise document with specific program recommendations would be more useful to the Office of Evaluation.

As further background, Professor Lind, who is an economist, chose as collaborators Professor Blakey, an expert in the criminal law, and Professor Blum, a social psychologist and criminologist with a long history and experience in the criminal justice field, to get varied perspectives. As was predictable, the participants presented different points of view and these points of view sometimes conflicted, although the three participants did find themselves in agreement much of the time. Perhaps the biggest difference among the participants was that, while Professor Blakey was interested in questions pertaining to the objectives of criminal justice and how their achievement might be measured, he felt that in many places the system of criminal justice was in such a state of disarray that we should focus on more process-oriented measures of how it is functioning, with the hope of making immediate improvements. He was skeptical of the return from money and research time spent analyzing the overall goals of the criminal system and of trying to measure its performance with respect to these broad goals. While Professor Blum and Lind agreed that there was a real need for looking at basic indicators to determine whether the system is performing those basic tasks that we think it should be performing, e.g., providing fair trials without long delays, minimum standards of decency in the prisons, etc., they both felt that much is to be gained from more long-run research designed to describe and measure the basic processes and output of the criminal justice system and from the use of such measurement to evaluate how the system is functioning and how it can be improved. As a result, they would recommend a larger scope of work than Professor Blakey, but one which would encompass the type of analysis that he supports.

Introduction

In this Report, evaluation is considered to be a management tool to improve decisions with regard to the planning and operation of our criminal justice institutions. Implicit in the notion of improvement is the concept of a scale of measurement by which we can determine how well the system as a whole, a given institution within that system, or some sub-unit of an institution is performing with respect to some task or to some objective. Based on an appraisal of how well it is performing, we may decide to increase or decrease the resources allocated to that activity, to reorganize it, to change it management, etc., in an effort to improve its effectiveness and the effectiveness of the overall system. Therefore, evaluation is essential to decisions with regard to the allocation of resources within the system and to the internal management of our criminal justice institutions.

At the same time the process of evaluation poses both a threat and an opportunity to the individuals within the institutions being evaluated. It is a threat in that an unfavorable evaluation or one that would lead to disruptive changes in the organization may adversely affect the members of the organization; it is an opportunity in that a favorable evaluation may result in more resources, greater power, etc. Therefore, evaluation cannot be viewed simply as a neutral tool for making better decisions with regard to the planning and management of our criminal justice institutions, but must also be viewed as an instrument for control and power within the system that will be resisted, coopted, and manipulated by participants for their own purposes. [See Appendix A by Richard Blum, pp. 1-3, for an insightful discussion of this issue]. While the major part of this Report will focus on evaluation and measurement as a technical tool for better management and resource allocation, it must be emphasized that the development of measures of performance should be part of an overall evaluation strategy that includes incentives for the adoption of sound evaluation practices. Further, the evaluation procedures and performance measures that one developes must be consistent with the objectives of those who will use them; otherwise, they will be discarded. More will be said, subsequently, on this in the context of a program of research for the Office of Evaluation.

Given that evaluation is a management tool to promote better decisions, then the value of improved evaluation must be measured in terms of the value we place on the improved decisions that it facilitates. The same principle also applies to the development of improved measurement techniques. The value of refinement of a measurement or the elimination of sources of error should be assessed in terms of the improvements it facilitates in decisions it is used to support. Thus, a first cut at measuring some dimension of the system that heretofore had not been measured may be of much greater value in terms of decision making than the sophisticated refinement of some existing measure.

This point also relates to the previous one that evaluation, and measurement related to evaluation, may meet significant resistance within the agencies responsible for implementing its use. A simple, well understood, and established method of evaluation that provides roughly correct information and that is used is superior to one that it more sophisticated and precise, but that it not. This is not to say that a program of evaluation research should not undertake to develop new and more sophisticated evaluation and measurement tools that will impact only the future planning and management of our criminal justice institutions. In most cases the adoption of new methods takes time, but in allocating our limited research funds, we should look for those areas where the potential return from better management as a result of new or improved evaluation and measurement tools is highest. In short, just as the principle of cost-effectiveness should guide decisions with regard to criminal justice programs so it should guide our research strategy.

Defining Objectives as a Basis for Measurement and Evaluation

The very concept of evaluation of the criminal justice system and its component activities implies that we have objectives or reasons for the existence of those activities, and that we can assess whether these activities are performing in a way that more or less promotes the achievement of these objectives. Thus, implicit in evaluation is an objective or set of objectives and rule of determining whether or not one situation is better than another with regard to these objectives. At a very minimum, evaluation presupposes that given any two

situations we can determine which one is preferred given our basic objectives. Put differently, evaluation presupposes the ability to establish an ordinal ranking of alternative situations or that we can rank alternatives on an ordinal scale.

There are several ways that one can approach the problem of evaluation and each has different implications for the role of measurement. The first approach is to define the objective or the objectives that one wishes to achieve and then to develop a procedure for assigning numbers to alternative situations so that a situation that provides for a higher level of objective achievement is assigned a higher number than all other situations corresponding to lower levels of achievement. In cases where there are several objectives, a measure of achievement must be developed for each objective, and in addition, we must define an objective function that maps vectors comprised of the scores with regard to each of the individual objectives, into a single real number. In theory, the objective function incorporates one's value system as it pertains to how various objectives should be weighed. This approach is classic and is familiar to every student of operations research or decision sciences.

Given this approach, the decision problem or the problem of evaluation requires that one consider each of the alternative courses of action available to the decision maker, predict what the effect of each will be on the level of achievement of each objective and therefore on the value of objective function, and choose that course of action that maximizes the value of the objective function. Given that you can define and measure the objectives, define an objective function, and predict the effects of alternative courses of action, the objective function provides the evaluation and the problem of choice becomes that of finding the alternative that maximizes the objective function.

This approach has much to recommend it in that it presents a complete representation of the elements that are involved in any decision or evaluation. One can see clearly both the elements of prediction or assessment of the outcome of alternative courses of action and the elements of valuation. The problem, however, is that for many reasons (which we shall not go into here) it is almost never possible to define an objective function that would command any degree of

acceptance. In almost all cases, the most that one can do is to define various objectives and measure their achievement. This information can then be given to a decision maker with the weighing of alternative objectives left to be balanced as part of his internal decision calculus.

However, even this is not a clearly satisfactory procedure for evaluation of the criminal justice system and its component activities. It is frequently pointed out that there are many objectives of the criminal law and of the system of criminal justice institutions that has developed to enforce it. Most men will agree that there is more than one objective and that different men will differ on what they believe the objectives of the system to be. This has led expert observers such as Professor Blakey to conclude that it is impossible to develop an objective for the criminal justice system and to measure the performance of the system as a whole. He would argue that it is much more cost-effective to concentrate on more process-oriented measures of performance at the level of operating units within the system such as clearance rates for police, case loads processed or delay times for the courts, etc.

While most people would grant that the definition of a single, widely-accepted goal of criminal justice and the development of a procedure for measuring the achievement of that goal will probably always elude us, it does not necessarily follow that the only alternative for evaluating the performance of the system as a whole and its component parts is to look at process-oriented measures of performance alone. While different people will differ on what the objectives of the criminal justice are, we can get wide agreement that certain types of information are relevant to an evaluation. For example, whether one views the purpose of the criminal law as crime control, keeping the King's peace, or of teaching and reinforcing certain moral values, one would probably agree that the level of crime in the community and changes in the level of crime are relevant data for any evaluation of how well the system is working. Similarly, if we found that doubling the clearance rate for a police department had no effect whatsoever on the level of crime in the community either because the underlying

incentives or reasons for crime were unaffected or because some part of the system was not functioning, e.g., backlogs in the courts make obtaining convictions impossible except in a few cases, then most men would seriously question the value (in terms of what they would be willing to pay) of obtaining higher clearance rates.

While there may be a number of goals or objectives of the criminal justice system and while different men may subscribe to different goals, most people think about the system and behave as if crime control were one of the major goals of our criminal justice system. Our preoccupation and use of crime statistics supports this contention. Further, as we watch individual citizens adjust to the perceived threat of crime by moving to safer neighborhoods, staying in at night, installing locks, etc., it is absolutely clear that their objective is to reduce the threat of being victimized, to increase their security. It would be strange indeed if these same people did not see the reduction of the threat of crime as a major function of our criminal justice institutions. Therefore, in evaluating the system, we will not go far wrong if we proceed as if crime control were a major objective shared by most citizens.

Similarly, there appears to be wide agreement that one of the roles of the courts is to protect the rights of every citizen and to guaranty due process. Others would also argue that the courts serve a fact-finding role, and that one of our objectives should be to improve the accuracy of the fact-finding process both to assure that the innocent are not found guilty and that the guilty are. Whether one subscribes to incarceration of the guilty as a means of rehabilitation, as retribution, as paying one's debt to society, or simply as a way of keeping criminals out of action, one can agree that conviction of the guilty is an objective.

Two conclusions emerge from this discussion. First, despite differing philosophies and models of our criminal law and criminal justice institutions, we can get wide agreement that these institutions should have an impact on things such as the amount and type of crime. Second, even if one does not get agreement on a single goal or objective, it is important to ask the questions: What do we hope to achieve by undertaking action A? If we do achieve what we set out to achieve,

why do we care? When we have pushed this line of questioning to the point where the answer is, "Because that is something that we value in and of itself," then we have identified a basic objective of the system or a reason for action. This process of questioning is itself of critical importance for the planning and management of any institution as it forces a searching analysis of the activities of that organization.

In the case of the criminal justice system and its components, one can through such a process of questioning and analysis identify higher level goals and develop procedures for measuring the achievement of these goals. The limitations to the use of such measures in evaluation arise largely because of limitations in our ability to predict or assess the effect of particular courses of action on the achievement of higher level goals rather than from an inability to define such goals.

To illustrate this, consider the following examples. Programs by the police to increase patrols, to reduce response times, to improve detection work are all undertaken with the idea that they will reduce crime either by increasing arrests which, assuming that this leads to increased convictions, gets criminals off the street and acts as a deterrent to others, or by the deterrent effect of a greater presence. Crime reduction is a stated and accepted goal of these programs. The problem of using the impact on crime to measure the effectiveness of these police actions is that it is exceptionally difficult to determine what crime levels would be with and without them. This is what is required for evaluation. We can observe crime levels before and after the programs were put into effect, but unless we can reasonably assume that all other factors that influence the level of crime including random fluctuations remained constant, we cannot attribute these changes to the actions being evaluated. Therefore, to evaluate criminal justice programs with respect to their effect on the level of crime, one must have a model or theory of crime that allows us to control for changes in other factors affecting crime.

The broader the questions we ask as part of evaluation, the more difficult this problem becomes. For example, let us suppose that one goal that led to the creation of LEAA was to stem the rising tide

of crime in America. One would like to be able to answer the question: Has the existence of LEAA and the programs it has funded made any difference on the level of crime in America and, if so, how much of a difference? This is a question we cannot now answer. The problem is not that we cannot agree on what it is we would like to measure or how to measure it, but rather that we don't sufficiently understand the factors that determine the level of crime to separate the effects of LEAA from the effects of a multitude of other changing factors that are known to influence crime.

What this forces us to do in evaluating criminal justice programs is to measure effectiveness in terms of some more limited objective or task on the assumption that the performance of this task will contribute to our higher level goal. We may use some behavioral or judicial theory to justify such an assumption. For example, one would predict that a higher arrest rate would both deter crime and reduce crime by reducing the number of criminals at large, therefore we might measure the performance of a police department using the clearance rate because we feel we can better assess the effect of police action on the clearance rate than on the level of crime. Similarly, in the case of the courts one might be willing to assume that, other things being equal, justice would be better served without long delays. Further, one might be able to predict the effect on the speed of the judicial process of making procedural changes, adding more judges, changing the size of the juries, etc. These alternatives could be evaluated with regard to this dimension of the judicial process in terms of their effect on waiting times. It is critical, however, to keep in mind what the connection is between the operational, process oriented measure of performance and the higher level objective. Clearly speed is not itself the goal and if pursued, without some concept of the higher goal in mind, would lead to the adoption of procedures that are totally inconsistent with our concept of justice.

To summarize one approach to performance measurement and evaluation of the criminal justice system and of its component institutions and activities is to ask what are we trying to achieve and to push the question why until the answer is either that "that is something we value

in itself" or "whatever my ultimate goal, I am sure that this is something I want to increase or decrease." The level of crime is a case in point. Whatever our theory of justice most of us agree that we would like to decrease crime or that decreases in crime will correspond to more effective justice, greater security, or whatever. If we can develop a procedure for assigning numbers of situations so that higher numbers correspond to higher levels of achievement, then we have developed an ordinal scale on which to evaluate performance with regard to that objective. The primary limitation to using such measures to evaluate particular policies, programs, or actions is our inability to assess what impact they will have on the measured objective. In such circumstances we generally choose some intermediate objective that we believe is positively related to the more fundamental objective but in terms of which we can assess the impact of the actions we wish to evaluate. This will lead to the development of an array of performance measures that are appropriate for evaluation in different circumstances, for different kinds of decisions, at different levels in the system.

Many measures for evaluative purposes do exist and are being used in some form. While in many cases these measures need refinement and should be based on better data, they do exist and are being used. Therefore, one job for research is to look at what measures we are now using, how they are computed, and the data on which they are based, and to suggest improvements. The much larger job will be to develop the capability to determine the effect of policy options on measured performance.

Public Opinion as a Measure of Performance

Another approach to performance measurement and evaluation is to go directly to people in the community and to try to ascertain and measure how they think that the system is working and to determine what dimensions of its performance are important to them. We will refer to this as the marketing or survey approach to determining what people want and how they evaluate what they are getting. One can legitimately raise the question, since its the taxpayers' money that is paying for the criminal justice system, shouldn't we be evaluating its performance and making decisions about whether it is worth the cost

on the basis of his preferences and his perceptions of what he is getting for his criminal justice dollar? This line of argument certainly has a strong and appealing logic.

However, one might object that people's perception of how our criminal justice institutions are working is not based on first hand knowledge and is largely shaped by such things as items that appear in the press. Also, it might be argued that a police department could change its performance rating simply by undertaking a vigorous public relations campaign. All these things are undoubtedly true, but do they matter? We allow consumers to judge the quality of cars, appliances, and many other complex products and to make expenditures decisions even though they may not have engineering data on various aspects of performance. Further, we accept individual judgements in the face of intense and continuous advertising designed to color and change their perceptions of various products. Should we not be willing to make allocation decisions for criminal justice on the same basis as for other products?

There are some differences between the situation with respect to criminal justice and other products. First, the degree of ignorance by most citizens as to how the system and its component agencies are operating is probably greater than in the case of most products. This is largely because most people have very little contact with the system. Most citizens never have been in a jail or prison nor have seen how they operate. The situation is almost the same for the criminal courts. To the extent that citizens have contact with the police, it is generally in connection with traffic control or the social service functions of police work. Therefore, the ability of citizens to make informed decisions with regard to expenditures on criminal justice will probably be significantly less than on the family automobile. A person at least knows if the engine will start, if it will go over thirty, and if it rides like a truck; most people do not know such rudimentary facts about the police or the county jail. [It should be noted that there are two groups of citizens who do have first hand knowledge of the system, namely victims of crime and those who have been arrested. These groups are a potential source of valuable information on how the system is performing and their opinions should be surveyed].

Besides the consumer's lack of first hand knowledge, there are several other differences between justice and other products. Justice is not sold in a market so that individuals cannot comparison shop thereby sharpening their understanding of alternative products and their prices. Further, people do not actually buy criminal justice services in the market so that we cannot verify their stated preferences with hard market evidence. We are left to rely on our survey results without the benefit of observing actual sales behavior.

With all these problems in mind, it appears that using survey and associated scaling techniques to measure how people feel the system is performing has promise and should be pursued as part of a program for criminal justice evaluation. This approach is complimentary in a number of ways to the previous one of beginning with the definition of objectives and proceeding to develop measures of objective achievement. First, survey techniques can be used to determine the attributes that people value from the service they receive from their criminal justice institutions. This is another way of saying that survey techniques can help us identify the people's objectives for the system. Such surveys are likely to show that while people are concerned with crime control there are many other important dimensions of, say, police service and police behavior. Thus, survey techniques are likely to help us identify new objectives, in particular practical ones which may be very important in how people assess their local criminal justice agencies.

Second, survey data combined with scaling techniques may help us around the problem of defining an objective function in multi-attribute decision problems. For example, such techniques can be used to weight crimes by their seriousness (as for example the Sellin and Wolfgang crime seriousness index) or show how various police activities are weighted in value by citizens in a community. Thus, not only can such techniques be used to identify objectives, but to weight them as well.

Finally, it provides another method of getting a crude measure of how the criminal justice system and its major components are performing. It provides an alternative barometer to the standard barometer, namely, the crime rate. Like the crime based measure of system performance, measures based on public opinion are subject to many of the same difficulties. There are many factors other than what our criminal justice agencies do that will affect the ratings; these factors and the

extent of their influence are not well understood; we can not separate the effects of changes in the system from the effects of changes in these other factors. In other words, we have all the same problems that we have with crime statistics. Yet, the public's evaluation of the system is important just as is the level of crime and must be considered. In fact, it is more important as in the final analysis it is the taxpayer's money that supports the system and he is the beneficiary or casualty of its performance. In the end it should be his judgement that counts.

Performance Measurement and Evaluation for Resource Allocation and Management

Certainly two of the primary reasons for measuring performance and evaluating programs on the basis of performance are to enable us to better allocate our resources and to better manage the criminal justice system. Without claiming that these are the only purposes for measurement and evaluation, they are important ones and they are the ones that we focus on in this Report. It is important to discuss measurement and evaluation in the context of resource allocation and management because the decisions with regard to allocation and management determine the requirements for measurement and evaluation.

Consider first the question of resource allocation within the justice producing system. It can be analyzed in terms of three separate but mutually interdependent decisions: (1) decisions about how much the society's resources are to be devoted to the justice system, that is, what should be the total budget allocated to the criminal justice institutions, (2) decisions about what proportion of the total expenditure will go to each of the production units, and (3) decisions within production units about the allocation of funds among specific production tasks.

The first decision involves contrasting the incremental increases in justice against the incremental costs of obtaining these increases. These costs, measured in dollars, represent the foregone opportunities of the use of goods and services that are inputs. The optimal expenditure of justice will be reached when the total spent has been increased

up to the point where the incremental increase in justice is valued equally to the incremental costs of obtaining it. In other words, the optimum is reached at the point where the value of the last increment of justice produced just equals the last increment of cost required to obtain it. How much a community should spend on criminal justice activities depends on how justice is valued with respect to other goods and services, the relative cost of these other goods and services, and the total wealth of the community.

Because of the fragmentation of the criminal justice system and because of the multiple sources of funding, it is unlikely that the total amount spent on criminal justice will be the result of the balancing of the gains measured in terms of some overall measure of performance for the total system against the costs. At the same time, each jurisdiction that funds criminal justice activities will weigh the value of those activities against the costs. While it is unlikely that a single performance measure will be used by policy makers in making trade offs that go into the budgeting process at the state and local level where most of our criminal justice institutions are funded, it is likely that if local officials can see promising programs, and if their performance can be documented, performance measures are likely to play a major role in budget determination, e.g., if police can show that increased patrols will lower burglary or if the state correctional department can show that more parole officers will lead to a lower rate of parole violations.

There will be very different effects on the production of justice depending on where funds are allocated within the system. Similarly, the amount of justice that can be obtained by any additional expenditure will depend on how these funds are used within the agencies to which they are allocated. Suppose that resources were allocated to the police for increasing the arrest rate. This increased capability would probably result in an increased number of cases for trial. If the resources allocated to the courts were not sufficient to handle this increase, then overall effect on crime control, and in the broader sense of justice, of increased police effectiveness would be minimal. What we see here is that the contribution to justice by any one agency is critically

dependent on the operation of the other agencies within the system. Therefore, it is critically important that whatever we spend on the system as a whole, we allocate those funds in a way that will maximize the total effectiveness of the criminal justice system as a whole. This means that balancing expenditures among agencies is an important allocation problem, one that must be seen in terms of the production of the total system.

This particular point is one that we discussed in some detail at our meeting and one that greatly concerns Professor Blakey. It turns out that in many situations there seems to be a real imbalance between what the police on one hand can achieve in terms of making arrests, albeit not as good as we might desire, and what the courts and corrections are capable of handling. Professor Blakey points out that we don't need refined measures of performance and refined techniques of evaluation to know that if an offender cannot be brought to trial and must instead be released because of a bottleneck in the courts, better police work is going to have little effect. Similarly, if crowding in our prisons means that people are given lighter sentences, given alternatives to prison, or if as a result of putting one criminal in prison another one has to be released, then our system of deterants is not going to work as planned. Good police work will not have the desired effect even in cases where convictions can be obtained.

Finally, at the lowest level, each agency has to make decisions with regard to how it spends its funds on men and material and to which tasks it assigns these men and this material. It is at this level that measures of performance may have their most profound impact on the day-to-day operation of the system. The manager at the agency level needs to know several things: (1) Are the tasks to which he has assigned men and material being performed well and in a technically competent manner? (2) Is the performance of these tasks having an impact on higher level objectives? For example, the rationale for increasing traffic controls may be that by giving more tickets one will cut down on the traffic offenders and thereby reduce traffic accidents. From a management point of view, the police administrator wants to know:

(1) Were his men performing their task, namely, were they giving more tickets to traffic offenders which is a measure of performance that shows that they were doing their job?, and (2) Did this have any effect on the basic objective of the action? Thus, he needs measures of performance that are both task-oriented and measures of effectiveness related to basic objectives.

Suppose that the police administrator found that when the program was in effect that there was no decrease in the number of traffic accidents in the area. Given this information, he still needs to determine whether this was because the program of increasing patrols and citations did not have an effect on the accident rate, or whether his men were simply not carrying out the task that was assigned to them. Therefore, to evaluate programs he needs a measure to tell him if the task was performed well and one to tell him if it made a difference. He also needs the task measure as a means of monitoring and controlling the operation of the organization.

This discussion suggests several conclusions. First, even if we had an acceptable measure of the total performance of the criminal justice system, it does not appear that this measure would be used in balancing the gains from greater expenditure for the system as a whole against the added cost. The primary reason for this is that the funding of our criminal justice institutions is fragmented between local, state, and federal agencies with the vast majority funded at the state and local levels. Also different units of government have responsibility for different institutions and activities within the criminal justice. Therefore it is likely that the total budget for the criminal justice system will be determined on the basis of a large number units. Each will trade off the gains from better performance against the cost, but will do so in the limited context of its own programs and environment. For this reason, it is likely that more limited measures of performance will play a role in the budgetary process to the extent that they are used at all.

Second, if we are to obtain any balance at all between the activities of the various parts of the system within this fragmented system of criminal justice institutions and governmental units, we must be

able to coordinate the activities of each unit with the workload in all the other units upon which that unit has a significant impact. To obtain an effectively working interaction between components of the system, we should at a minimum develop the capability to measure the impact of increases in the output of one sector of the system on the workload of other sectors of the system. Beyond this it would be desirable to know what the effects on performance will be of heavier workloads in those institutions that are affected. The thrust of this is that for the purpose of allocating resources among different branches of the criminal justice system, it is vitally important that we develop measures of workloads and of performance for each of the component parts and that we develop the capacity to show the connection between changes in one part of the system and workload and performance in other parts of the system. In other words, we need more of the kind of the system modelling done by Blumstein, Larson, and others in order to affect a better balance between different parts of the criminal justice system.

Again, going back to Professor Blakey's concern, he sees the system as breaking down largely because of an imbalance in the allocation of resources to the courts and to corrections. His point is that we don't need a sophisticated measure of the output of the courts or corrections to know that the system is in a state of disarray when those institutions do not have the capacity to perform those minimal functions for which they were designed.

At the same time, it is important in analyzing such allocation decisions to keep the higher level objectives of the system in mind. For example, if we found that by releasing a large fraction of those who are accused and who would otherwise go to trial or by increased plea-bargaining, there was no effect at all on the crime rate as opposed to what it would have been if these people had been processed by the system, then we might question whether in fact we would want to commit more resources to the courts. This might suggest that we approach the problem in a different way.

The remainder of this Report proposes a long term program of research for the Office of Evaluation in the National Institute. This

program was prepared by Professor Lind based on his research for this project and his experience as a researcher and consultant in the field of criminal justice evaluation. It draws heavily from the contributions of Professors Blum and Blakey, but makes no attempt to present comprehensive coverage of their views. For this reason, their inputs are presented in their entirety as Appendices. It is strongly recommended that they be read. In particular, Professor Blum has outlined a complete program for research in Appendix A that should be carefully studied.

In the course of the work on this project several points of emphasis developed that were not fully anticipated at the beginning of the project. First, the effective use of performance measurement in evaluation is critical to justifying that a program of evaluation that supports the development of performance measures is cost-effective. Therefore, such a program should address not only questions of how to measure performance for purposes of evaluation, but also to ask where and how can these techniques be most useful and how can we implement their use in the decision process.

Second, it became clear that while better measurement is important, probably the single biggest obstacle to evaluating programs on the basis of their contribution to achieving various objectives is not our inability to define objectives and to develop reasonable measures for them, but rather our inability to determine what the effect of a given policy action will be (or what in retrospect it has been) on measured performance. This leads to the conclusion that basic research on the criminal justice system and how it operates, and on the social processes that cause criminal behavior central to better evaluation. The development of performance measures should be an integral part of this research, because for purposes of evaluation, the performance measure serves as the dependent variable.

Third, crime rates and the level of crime by type are and will remain an important element in criminal justice evaluation. We need to develop tools so that we can get more reliable crime data and collect such data on a basis that is both appropriate for evaluation of a wide range of programs in the criminal justice system and appropriate for

research on the factors that determine the level of crime in a given community. We need better models that allow us to identify and isolate the causes of changes in crime rates so that we can better measure the effect on crime of criminal justice programs and policies.

The following items are descriptions of the proposed program elements:

I. Research to Promote the More Effective Use of Evaluation by Performance

The purpose of this element of the program is to provide LEAA and researchers on criminal justice evaluation with guidance as to how to make criminal justice evaluation more effective. Further it might also serve as a guide to legislators and other professionals in the criminal justice field as to how to implement better decision making based on performance measurement within various parts of the criminal justice system. If implemented, it would enable us to identify those agencies that would be receptive to the use of performance measurement in evaluation and to provide them with the tools that would both be helpful to them and that would make a real difference in the decision making process. In cases where there is resistance, this resistance and its causes would be identified and analyzed. This would help us think about how to structure incentives to overcome these obstacles. This program element consists of several parts.

A. A number of studies of our major criminal justice institutions such as the police, the courts, the prosecutor's office, etc., would be undertaken to analyze what decisions they make, how they make those decisions, and whether they use performance measures. The point here is to assess current practice and whether it could be improved by improved measurement or by the use of better evaluation techniques. Further, if the answer is yes, then would the agency be receptive to technical assistance and evaluation and what types of tools and technical help would be best suited to its needs? If the answer is no, a study should determine why not and how the agency might be made more receptive.

To do this, part of each study would include an analysis of the private objectives and reward structures within agencies that create resistance or receptivity to evaluation. The output of these studies

would be used to identify both classes of agencies and in particular agencies within a class that would be receptive to better evaluation and measurement and also used to identify particular areas where better evaluation might have a significant impact. As part of this work, one would review the use of traditional measures in evaluation and assess the usefulness of these measures to determine whether modification of the measurement techniques or better data would significantly increase their usefulness in the decision making process.

The work that is envisioned is not totally different from the projects being carried out on the police by the American Justice Institute or on the courts by the Rand Corporation. In all likelihood these studies would constitute a strong first step in the program of research envisioned for this program element. At the same time, the emphasis on implementation and on the characteristics of the institutions and their personnel may take the proposed programs beyond the present studies. Given that there is work underway in this field, certainly the first step before pursuing further studies would be to review the work on these topics. What is envisioned is not the reinvention of the wheel. Instead it is a drawing together of our existing knowledge about how our criminal justice institutions operate and a focusing of this knowledge on the problem of implementing better evaluation techniques and better use of performance measures within criminal justice institutions.

B. The second item under this program element is addressed to the question of getting a better balance of expenditure between the various branches of the criminal justice system. One clear concern is that there is an apparent significant misallocation of resources among different criminal justice institutions in some instances where the system appears to have broken down in that one part can not handle the work generated by the other parts. It is proposed that research be undertaken to study ways in which we could evaluate the relative effectiveness of interrelated parts of the system and based on such an appraisal find ways of reallocating resources, or at least providing supplemental resources, to help those parts of the system that constitute a bottleneck. Such work should address the technical problems of assessing the situation and the development of models to assess the impact of each part of the system as it relates to the workloads of other parts of the system. It should analyze the institutional network that results in the existing funding decisions to see how it might be influenced and how a better allocation of our criminal justice resources might be effected. This activity seems to be highly appropriate for LEAA because we are talking about making funding decisions at the margin, and while LEAA does not provide the major source of funding for our criminal justice institutions, it can have a very big impact at the margin. Thus, an analysis of this type might show LEAA where it should spend its resources if it wants to have the maximum impact in the short run on the criminal justice system as it now exists.

II. Studies of Crime for Evaluative Purposes

It appears that crime is certainly one of the major concerns of the criminal justice system and that many specific programs within the system are designed to reduce crime in one way or another. Regardless of one's theory or model of justice, the fact of the matter is that both the public and our criminal justice agencies are interested in the rate of crime and how various activities of the system affect that rate. Crime data are important for evaluating criminal justice institutions and it is worth making a major investment to make crime statistics and crime data a useful tool for evaluation. There are several major problems with the crime data that we now have that makes it almost useless for evaluative purposes and for the purpose of doing basic research on the factors that influence crime. First, most of our crime statistics are based on data collected and prepared by police departments on the basis of reports to them. The inaccuracies and biases in this data are familiar to us all, yet, because we have nothing else, we continue to use it. For purposes of evaluation it is simply unacceptable to have agencies that may be evaluated contol the source of data. The incentives and potential for manipulation are great. Second, but as serious as the problems of manipulation or bad record keeping by the police, are the problems of underreporting which are well known. Third, much of our present data on crime cannot be broken down by geographical location, by type of crime, or by time period in a way that makes it useful for either evaluation or for basic research on the causes of crime. For example, when one is trying to analyze the effect of a new patrol program on a particular area, one would like crime statistics broken down on the basis of a region that may not correspond to one of the geographical record-keeping units of the police. One would also like to have selective information on what has been happening to crime rates in adjacent areas where displacement effects may be felt, etc. In general, it is almost impossible, or at least very costly, to get this kind of data out of police records. Similarly, with regard to basic research on the factors that influence the rate of crime, one would like to have one's crime data developed

in such a way that it could be put next to other sources of data such as basic demographic and economic data to analyze the interaction among these influences. The point is that our existing data is not adequate for the job. Furthermore, the answer is not to try to bludgeon the police into keeping more and better records. We still will have all of the biases that come and are inherent in reported crime statistics and from having the police control the data. Further, it is very costly to develop data in this way and it burdens the police with an enormous record-keeping task.

One additional problem is that there are a number of large areas of crime where we have few if any, records at all, and very little understanding of the entire process of crime in these areas. They are organized crime, white collar crime, and transnational crime.

Therefore, the following program elements are recommended:

A. A study would be undertaken to design a research methodology and the supporting data base that would be required to assess the degree to which various socio-economic and criminal justice system factors influence the crime rate. What is recommended is not that the entire study be commissioned, but that some group be commissioned to put together the strategy and the models that would be used in such a study. Most of the work on the causes of crime has been deficient in the sense that it has not taken into account the interaction between the variables that determine the level of crime. Until the Votey and Phillips work, little attempt was made to take into account the fact that communities with higher crime rates were also likely to have higher levels of police activity in response to those crime rates. The relation between police activity and crime is a two-way relationship and this must be accounted for in the statistical models that are used. What is needed is for someone with strong methodological skills to develop a simultaneous equation model of crime including both socio-economic variables and criminal justice variables. Further, this person or team should assess the data requirements of using such a model and develop a strategy for collecting the data. In other words, such a study would be to design the ideal analysis of the factors that influence the rate of crime. Further, this would be developed in such a way that it would be useful for the purposes of criminal justice evaluation in trying to separate out the effects of various basic criminal justice actions from other factors in the community.

The purpose would be to develop a plan that could be subjected to criticism and that could be refined before embarking on a major effort to collect the necessary data and to implement the model. Since this work is not only important for the evaluation of the effects

of criminal justice activities on the level of crime, but would also constitute a basic step forward to our understanding of those things that affect the level of crime in our society, it might be funded collaboratively with some other office within LEAA or be funded jointly with NSF or some other agency.

B. It is proposed that studies be undertaken to determine whether it would be possible to develop survey instruments that would be used in assessing the rate of specific crimes, in certain areas, during specific periods of time in an effort to facilitate the analysis of the impact of particular criminal justice activities, mostly police, in various localities. On many occasions evaluations are requested where several million dollars has been allocated to police agencies to undertake programs with the primary objective of crime reduction. Many such programs have been funded by LEAA as experimental programs. In most cases, there has been an evaluation budget of up to $150,000 to analyze the impact on crime. The problem for the evaluator is to get the relevant crime data for the areas that were affected by the program and for the relevant areas where displacement effects may have been felt over the relevant time period. The problem is not of controlling for changing socio-economic conditions, but of looking at what was happening to crime in the area in the short run as a result of the program.

Perhaps the only way to get the necessary data at a reasonable cost is to develop a survey instrument that could be applied cheaply to get data within the relevant areas. The budget for most single evaluations is too small to develop and test the necessary survey instrument. In almost all cases the evaluator has no choice but to use whatever data he can get from the police.

Therefore, it is recommended that the Office of Evaluation support the development of crime and victimization survey techniques that can be used to assess crime levels both on a continuing basis and in particular situations. As an exploratory part of this program the possibility should be investigated of having the police keep only those records on crime that are necessary for the internal operation of the system, e.g., evidence on which to base prosecution and data requisite to internal police management, and of using survey methods for assessing crime levels and trends and for evaluating criminal justice programs. This would free police resources for other things and could provide a more reliable data base both for evaluating the performance of the criminal justice system and its institutions and for scientific study on the causes of crime.

C. Most crime data is related to violent crimes and to commmon crimes against property, particularly those that are physical in nature. Basically, we are keeping data on violent crimes and on street crimes. Yet there are a number of areas of criminal activity that are largely unexplored and not well understood. These are white collar crimes, organized crime, and transnational crime. Obviously, there are many cases in which these forms of crime are related. The

criminal organiztation that operates across national borders supplying heroin, arms, stolen documents, etc., is organized and operates transnationally, and the swindlers who mastermind and implement major international stock frauds fall in both the white collar and transnational categories. It is proposed that in each of these areas, The Office of Evaluation fund pilot studies to assess what is known in these areas of crime, including surveying what data we have, what capabilities we have for dealing with such crimes, how we would measure the impact of such crimes on society, and how we would evaluate and measure the success of programs to reduce such criminal activity. It appears that the magnitude of these forms of crime and the resources that are currently being spent to cope with them are drastically out of balance and that we are not allocating the resources to enforcement in these areas that is justified. It would be the purpose of these pilot studies to take a first cut at assessing the situation in each of these areas and of evaluating the magnitude of the problem and the degree of success of our response.

III. Exploration of the Potential of Survey and Attitudinal Scaling Techniques for the Measurement of Criminal Justice Outputs

There may be significant potential not only from using survey techniques to find out what various groups such as victims and offenders think about the operation of the criminal justice system, but also in the use of modern techniques of attitudinal scaling in order to develop measures of performance that are based on citizen perception and citizen values. There are many such techniques that have been developed and they have come primarily from social psychology and have found their application in such areas as marketing research. It is their marketing related applications that appear to have the most promise for making a substantial contribution to criminal justice measurement and evaluation. These techniques have promise for developing overall measures of performance for the system, of helping weight various objectives of the system, for identifying new objectives or new sources of concern among citizens, and for providing institutions like the police with feedback about how people feel about their operation which may in fact influence the performance of those operations. Although there is potential, it is not proposed that a major program of research and development in this area be funded immediately.

A. Rather, it is proposed that the Office of Evaluation commission a study of the potential of these techniques for criminal justice measurement and evaluation. The members of the team for this project should include people who have a firm knowledge of our criminal justice institutions, a knowledge of criminal justice evaluation and the state of its art, and individuals that have a strong capability in the theory and application of survey and scaling techniques. This is a job that could very well be done at the university, or in any case, by a team that includes consultants who have the specific techniques in the areas of scaling and evaluation research that are required. The functions of this study would be to provide the Office of Evaluation with a careful analysis of the potential and the details of this program element if this line of research appears promising.

IV. Methodological Developments for Criminal Justice Research and Evaluation

There are a number of methodological issues that should be addressed. In most cases, this methodological work will be relatively inexpensive and will have a wide application. Some of the basic methodological work has been discussed in connection with other programs elements. Nevertheless, there are some additional methodological issues that should be addressed.

A. It is proposed that a program of a basic methodological nature should be funded to review the types of criminal justice performance measures that are being developed and used and to analyze the mathematical and statistical properties that these measures should have for different kinds of uses. In general the development of measures of performance in all areas of criminal justice, and in fact other public programs areas as well, have been divorced from the theory of measurement. Yet the uses to which some of these measures have been put, such as in cost effectiveness analysis, require that they have special properties, e.g., certain scaling properties. Further, many statistics on crime and the performance of criminal justice institutions are misleading because they do not in fact represent what their users claim they represent. A program of research is proposed to consider the uses of the measures of performance in the criminal justice field to determine the properties that measurements should have given how they are used. This might include a number of relatively small studies of a methodological nature reviewing the properties of existing and proposed measures of performance.

B. Wherever it is possible to measure the benefits of a program in dollar terms, benefit-cost analysis can then be applied. This frames allocation decisions in terms of a common unit measurement, namely, dollars. It is important that we explore in a systematic way the potential for using benefit-cost analysis in the criminal justice field

and in particular that we invest money in developing ways to measure benefits in those cases where it may be possible. Therefore, it is recommended that a program element support first a survey of the areas in which it may be possible to measure the benefits from criminal justice activities, and then to fund further methodological studies for development of benefit measurement methodologies in these specific areas.

V. Continued Basic Research on the Operation of Criminal Justice System

It became clear in the course of our analysis and deliberation, that our ability to analyze the impact of programs on measures of performance was perhaps the greatest obstacle to better evaluation. There is ongoing research that is helpful in making the connection between particular programs and overall criminal justice performance. The purpose of adding this element is to suggest that a certain part of the budget go to funding a continuation of such basic studies.

A. The Office of Evaluation should support systems studies of the criminal justice system that show the interrelation between the operation of various parts of the system and their effects of the other parts of the system and how they operate. This work is basic in making predictions about what the effects of programs in various parts of the system will be on other parts of the system.

B. There are many possible influences or alternatives to criminal justice action that lie in the private sector. If we are to evaluate criminal justice actions we should take into account the alternatives that lie on the boundary or outside of the system itself. Research should explore potential solutions and their impact that lie at the boundary of the formal criminal justice system.

C. Basic work on such things as record keeping, basic indicators of performance with respect to tasks within the system or basic conditions within the system should all continue to be supported where good proposals by good people come to the Office of Evaluation.

VI. There is a major opportunity to do an evaluation study of LEAA's own program, both retrospectively and prospectively. Because LEAA provides funding, it has leverage to get the agencies which it funds to do evaluations. It would be exceptionally important for the LEAA, in fact for our knowledge of evaluation in general, to analyze what evaluation we should require or would desire from agencies receiving LEAA funds.

A. It is proposed that a project be funded to design an evaluation strategy for LEAA. Such a project would analyze alternative sets of objectives and specify what one would need to know in order to evaluate the impact of LEAA programs individually and as a whole. Further, it would make recommendations regarding how evaluation might be improved given limited data, uncertainty regarding basic relationships, and limited budgets.

B. It is proposed that the Office of Evaluation support a post audit of LEAA projects both to evaluate project evaluation and to evaluate projects. This work and that proposed in "A" above should provide the basis for LEAA to develop a set of procedures for evaluating, selecting, and monitoring its program in the future.

Concluding Comment

While the subject matter of any research program is important, good research is critically dependent on having first rate researchers to do the work. A high priority research topic treated by second rate researchers is likely to produce a product of little value. Therefore, in building a program of research related to criminal justice evaluation, a part of the strategy should be to build a community of first rate people in this field. Continued support of such individuals or groups should be an important part of the Office of Evaluation's strategy while it, at the same time, uses its leverage to steer their research to areas on which the Office places high priority. Only with the development of a strong research community and the development of a receptive clientele of practitioners can the state of criminal justice evaluation be advanced.

APPENDIX A

EVALUATION OF THE PERFORMANCE OF THE CRIMINAL JUSTICE SYSTEM: IMPACT OF THE SYSTEM AND THE OPERATIONS AND OUTCOME OF THE SYSTEM

SOCIAL SCIENCE CONSIDERATIONS

by

Richard H. Blum

INTRODUCTION: Evaluation Management

We shall consider the evaluation of the criminal justice system in several ways. First, we shall consider the strategy and politics of evaluation. Secondly, we shall consider various criteria on which it rests and methods of evaluation. Thirdly, we address implications and proposals. The perspective of this section is drawn from the social/behavioral sciences. Its primary assumption is that evaluation is a necessary activity to accomplish rational public administration and public policy formulation. It is also a useful device for assessing the fairness or equity of the system, viewed from the standpoint of law, or democratic principles, morality, social science and administrative. Evaluation also contributes to knowledge as such.

In spite of its contributions, it must be recognized that there is much resistance to evaluation; individuals and organizations resist and reject appraisals, for these necessarily imply that there may be discrepancies between what is done and what should be done or that power is being given to others to. Insofar as evaluation is conducted by "outsiders," over whom system members can exert no administrative or social controls (via quid pro quo, censorship, punishment, etc.) resistance can take the form of extreme hostility and defensiveness, especially if it is anticipated that either the methods of evaluation will themselves interfere with existing functions and satisfactions or the findings will threaten equilibrious arrangements. Evaluators, as "experts",are often seen, therefore, as socially "illegitimate" competitors for or arbiters of power or prestige with system members. A further threat arises insofar as evaluation findings may imply the need for specific changes which will displace existing personnel, institutions or budget arrangements or will require efforts which are feared to be burdensome and "unrealistic" from the standpoint of the existing values or procedures of system participants. As studies have shown (Blum and Downing) resistance both to evaluation and to innovation in public governmental institutions can be extreme.

The fact that evaluation is generally unwelcome in any program is fundamental to considering its uses. It follows that a basic aspect of evaluation planning, almost without regard to the approach taken, is the examination of the likely sources of resistance, the development of means for incorporating, neutralizing or overwhelming these, and provision for facilitating innovations once recommended. Implicit here is what might (pompously) be called "meta-evaluation" or the evaluation of evaluation itself. We must affirm the importance of recognizing that

evaluation is not a means by which to guide or accomplish change unless the phenomena of competition for the privilege of judging performance and setting goals are understood.

Evaluation implies that justifications for the existence of criminal justice institutions, procedures, objectives and endorsements are being questioned. A justification is, essentially, any approving answer to the questions "why is A there, why does B do what he does, why is the achievement C sought or accomplished or why does D finance, support or otherwise sanction or legitimate A, B, C or D himself? It is immediately seen that the criteria set forth by the evaluator may not be those employed by those participating in any aspect of the criminal justice system. As noted, there is likely to be conflict over the "right" of an evaluator, be he within or without the system, to decide what shall be the standard, how it shall be examined, what shall be done with the results, and what inferences are to be drawn from the findings. The variety of criteria, methods and inferences, is evidence for the non-homogeneity of beliefs, interests and values among observers of and participants in the criminal justice system. That diversity, in turn, guarantees conflict based on these differing interests and perspectives and, necessarily, competition for the power to influence the system - its components participants and objectives - one way or another.

One aspect of evaluation management is to appreciate what factors contribute to a given position on an issue, as for example a person's education, vocation, criminal history, political philosophy, institutional affiliation, etc. An important approach to evaluation is to determine what these positions are vis a vis any single component, process, function or goal in or of the system. By determining through research the characteristics of those supporting and opposing such positions, one can construct a social geography of sorts, that being a map of positions, issues and the kinds of people affiliated with these positions. By determing the importance of a position, that is to say, the strength and intensity to which a person adheres to one or another criteria as the "correct" way to judge a criminal justice system component, one can also map the politics of evaluation and the intensity of conflict.

Insofar as there are disparate positions strongly held as to what standards are proper to use in appraisals, then evaluation undertaken by commitment - that is adhering to one criterion rather than another - is an exercise in, if not an

instrument of, social or political power. It follows that evaluation based on standards which are or are seen as threatening to others can hardly expect to resolve the problems of the criminal justice system if the diversity of interests and views is conceded to be a part of the problem. If, on the other hand, the value position from which evaluation proceeds is not only broadly based, but has a potential appeal even to those in opposition (here one can assume that individual and institutional beliefs and values while intercorrelated are simultaneously multiple so that, for example, an individual can be in favor of "punishing criminals" but at the same time oppose "expensive long term sentences") then evaluation findings may be used to implement the positions from which that same evaluation proceeded. The conduct of evaluation requires then that it be guided by a strategy. As a strategy it means that alternatives will be considered, costs and benefits weighed, value positions considered and resources allocated.

That strategy should acknowledge the interests of those authorizing the evaluation. It requires that those supporting an evaluation have at least some kind of map of the beliefs and interests of those who are to observe or be affected by the evaluation. It is necessary that the choice of methods be appropriate to the criteria which are to be employed and that these methods be capable of use in the social/institutional setting where other observers and participants live, work and may be adjudicated.

What Is to be Evaluated?

When asking what is to be evaluated, three quite different considerations apply. The easy choice is of that component or process or outcome which is known and of interest, as for example the arrest and clearance rates for a given police department for burglaries, the recidivism rate for parolees from a particular state prison, or the costs of a municipal court apparatus. Known institutions, forms of conduct or records data constitute the traditional targets or measures in for evaluation from which come conclusions such as "the burglary clearance rate is such and so" or "a five-year follow-up shows that parolee recidivism is..." or "the municipal court costs taxpayers X dollars a year."

It is much more difficult, however, when considering what is being evaluated, to develop knowledge about the events which comprise the phenomena being appraised. Saying that the police department in Midville has a clearance rate of 5% for burglaries may meet a minimal requirement for administrative reporting, but it is ordinarily not sufficient knowledge. One usually wants to know more: What is it within the Midville police department which affects on the outcome measure burglary clearance rates? One assumes it is not the 7th precinct building itself which

affects the measure, nor the color of uniforms, nor the fuel consumption of the bikes piloted by meter maids, nor the deliberations of the city manager about who will be the next police chief. What is it then, the actions of the detective bureau? The arbitrary practice of the chief records clerk who has always shown a 5% clearance rate in lieu of any records analysis? On-view arrests by one interested policeman out of a force of 30? Unless one refines the focus of the evaluation, moving from the easy semantics of the designation of a facility or gross process to what may initially be the search for and test of hypotheses bearing on influential events, it is quite possible never to know what it is that one evaluated.

The likelihood that relevant processes may be unidentified does not matter if one's interests as an evaluator are limited to the identification of gross features without concern as to antecedants and accuracy. This, in turn, depends upon the strategy for evaluation. If the strategy calls for an understanding of events occurring within a particular context linked, in turn, to measures of these and subsequent events, the latter perhaps considered as "outcomes," then one does want to know what it is that one is investigating. That implies an objective for the evaluation itself. If that objective is at all refined it also implies a concern with accuracy and that, of course, requires an interest in methodology. These three attributes of evaluation; defining purpose, defining the target, selecting a method are essentials in evaluation strategy. Through such definitions alternatives are set forth and decisions made possible.

Attributing to a police department a burglary clearance rate without knowledge of the processes leading to the figures is an example of over-inclusiveness, or just plain not knowing enough about what is going on within an institution and about an evaluation measure. A somewhat similar error of definition and measure - and sometimes logic - constitutes a third consideration in asking what it is that is being evaluated This one, however, is the problem of under-inclusiveness. It occurs when one fails to include enough (e.g. institutions, persons, processes) in defining an interest area or seeking to comprehend the events leading to the results or interpretation of a measure. Two examples will illustrate. The daily newspaper editorial complains that the police department has failed by not stopping crime. They are referring to a recent rash of muggings of the elderly. The editorial claims it is a police responsibility to bring such muggings to a halt and that heads should roll. The editorial "evaluation," for it is certainly that, has a limited view of crime. It does not consider what kinds of young men (it

tells us that much about the criminals in the editorial) engage in muggings. It doesn't consider what events in their lives have influenced their careers so that they engage in this criminal enterprise. The editorial certainly doesn't enlighten us about the attitudes that muggers have toward themselves, their elderly victims or towards the law which enable them to elect such nasty, vile conduct. Nor are we advised about the prior experience of these muggers in being arrested or processed in the criminal courts and corrections which have increased or decreased (if either) the likelihood of their engaging in further muggings. The editorial doesn't tell us about the characteristics of the urban environment which facilitate street crime (street lights, citizen apathy or fear, crowding, etc.) nor does it tell us about the chain of circumstances which lead to the impoverished elderly living in proximity to young hoodlums and having to walk without familial, neighborhood or police companionship. The editorial gives us no aid to understanding the limitations on the police department in its assignment of personnel to the victim's neighborhoods nor does it comment on the extent to which visible police would in fact reduce muggings (where do they occur? at what time? how often have they occurred with police in view?) We shall not continue further with the analysis, for it is clear by now that the editorial evaluation was based on a narrower presentation of the factors influencing mugging than the reader of criminological studies enjoys in his books. The editor was evaluating the police department role in muggings, but he did so in violation of the rule against under-inclusiveness, that is drum-beating over-simplification. Had the editor aimed for a more modest evaluation objective, asking for instance, "can increased police patrol and increased arrest of muggers affect mugging rates?" he could have justified his attention to the police only. Given what we know,he might not, however, have been able to answer ever that relatively simple question.

The "criminal justice system" is a convenient abstraction for an immensely complex series of presumably interrelated events. The terminology invites the attention of systems analysts and organizational specialists who seek to describe how one part links to another, and what internal events constitutes flows of and blocks to the processing of the daily business; people, papers and objects. Not charted in systems studies, but of interest in some kinds of evaluation, are also the flow and counterflood of emotions, ideas and power. One encounters the problem of under-inclusiveness here in defining what it is that is being evaluated when a component or event is missed which, in fact, influences an outcome (or charted progression.) An ordinary evaluator describes what he is asked to see; for example, a type of offense, say heroin sales, then police intervention, referral

to prosecutor, being in jail, a court with its internal events concluding in a trial (with numerous appeals, of course) and disposition of some sort. These major formal operations can be fleshed out by descriptions of plea bargaining, by notations on dispositions in lieu of sentence to a methadone maintenance center, and other realistic commentaries on how the "system" really works.

The "system" is ordinarily conceived according to its formal aspects; official offenders, official police, official judges, the whole official potpourri. No systems investigator is so naive, of course, as to presume that the buildings, badges of office, official documents and public records are the only elements. These are but aspects or emanations of daily work in which people confront and accommodate one another, pursue their personal advantage (be that to avoid jail, enjoy a fat bribe, receive a high legal fee, write a decision that will merit the approbation of judicial peers, find a job as jailor for a drunken brother-in-law, do their professional best in spite of it all, or what-have-you) and try to adjust, grow wise, get out or whatever it may be. Nevertheless, if the focus of evaluation is to be the criminal justice system, that focus is likely to be attuned to the visible and existing formal apparatus, what it processes and with what outcomes of social interest. There is no problem in this providing that the evaluation objectives are only these public events and do not have, whether in full innocence or by guile, hidden expectations and agenda. What if the assignment to the evaluator is, "Tell me about the efficiency of the criminal justice system in Smalltown?" Let us, hypothetically, grant him money, patience, skill and access He can do well at his task, telling us how many reported offenses lead to apprehension, what the proportion of those apprehended are that appear for trial, what dispositions occur there and, for each part and the whole, how long it takes for each participant or each stage, what way stations are passed, what ritual marks are inflicted on documents or biographies and what the outcome at each stage costs for each official participant, each official wrongdoer, and each taxpayer groaning behind them. The evaluator can do even more, of course, given many years, dollars and professional teamwork but this is enough for the illustration.

What, on the other hand, if the evaluator's instructions are, "Tell me how effective the system is?" In the past, typically, one has counted reported crimes cleared or alleged wrongdoers arrested, cases processed by prosecutors, bodies held in jails, cases going to trial, decisions not reversed on appeal, offenders coming out the other end having experienced justice, parole revocations or finally, recidivism. None of these measures is as simple as it seems, but measures, at

least, there have been. Effectiveness, it is seen here, depends on the official mandate to each component of the system and relies on the component's own traditional measure of its product. No one is surprised that each component can claim that its part of the system "works" even though each process need not be in harmony with the other and the first input and final output measures are as yet but dimly comprehended in terms of any principles of overall effectiveness.

When one wishes to go beyond these traditional measures, by expanding the criteria of effectiveness and the subsequent accuracy of measurement, one encounters greater challenge in the search for greater knowledge and the potential ability to improve public policy. One aspect of such desires leads us to that same problem of under-inclusiveness in saying what it is we are evaluating. It occurs first when one decides to classify, as being outside the system of justice those operations and institutions which interrelate to it but are not traditionally considered as components. Here one could include all of the social institutions through which alleged offenders pass on their way to the formal system (usually thought of as beginning with arrest) and without which they would not have appeared as suspects had not some prior formal action been taken. These are the institutions which serve as "gatekeepers" for referrals to the police, bridges if you will between the "dark number" of unreported crime and the official bookkeeping of the criminal justice system. They include most, if not all, of the social groupings in which humans live, and where they exhibit such foibles as to lead to conflicts and disapprovals. These are the institutions which have the power to invoke a police response. They are the settings in which citizens make the judgment that crime has occurred and set in motion the wheels of justice. Such settings include the family, work groups, schools, leisure associations and the like. In these settings expectations operate as to how people should behave, what constitutes an offense, how the criminal justice system should respond to a call or complaint, and what should happen to the offender committed to justice. Remarkably little is known about the circumstances under which these definitions and decisions arise, except perhaps that there is marked variation in how acts which might be judged similar legally are, in fact, judged socially. On such judgments, as with the magistral functions of the cop on the beat, entrance into the formal legal system depends. The exclusion of these "entrance level" institutions from the purview of the criminal justice system can be understood in terms of existing classifications in the minds of public administrators or perhaps the public itself, for the formal system is the governmental one which is, essentially, reactive. Yet, if one is interested in how the system operates, one can contend that operations

at the intake side are as consequential as mid-point and outcome activities.

In the same manner, one may append to the existing components of the system all those institutions which act as reception centers for those processed sideways out of the traditional stages (police, courts, corrections.) When a policeman, acting as curbstone magistrate, "books" a neighborhood hooligan before his priest or family instead of the station house sergeant, that is a sideways referral. Other examples occur when the sergeant reprimands the alleged wife beater at the station house but releases him without record, when plea bargaining results in charges dropped rather than reduced, when the prosecutor accepts the defense attorney's argument that the hurried self-referral of the junkie to Synanon satisfies the interests of justice or when the judge in chambers agrees that the exhibitionist church elder can elect psychiatric care in lieu of trial and a sentence. These are all normal practices, but insofar as sideways reception centers are dismissed from evaluation because traditional classifications of the justice system exclude them, then the evaluator may be missing not just pieces of the system, but influences on outcome measures as well as constituent processes.

There is a strong tradition in limiting definition of the criminal justice system to the public sector, that is what taxpayers pay for directly and what legislation defines as the justice system. Understandably, public agencies which wish to have an impact on that system limit their intervention to these agreed-upon components. Insofar as evaluation is primarily financed by these supporting or supervising public agencies, evaluation research focuses on those traditional public operations. One might illustrate this by considering the expenditure of all LEAA funds since its inception, comparing monies given for support of and research in or on public agencies compared to monies given to criminological studies which take place outside public agencies, as for example on the post-prison job adjustment of offenders, the possible genetic features of the "schizopath" (See Snyder, MADNESS AND THE BRAIN) or (Robins DEVIANT CHILDREN GROWN UP) which place him/her at risk of criminality, or the family values which predispose one minority adolescent to want to join the police force but which turn others away from such vocational aspirations.

One could take as criteria for the evaluation of the system as a whole some of the more general expectations which are offered as reasons for its existence. These include, for example, maintaining the (king's) peace, preventing crime, general deterrence (educating the public, contributing to group norms of lawlessness and propriety, creating risk for unlawful conduct, etc. (see Andenaes,

PUNISHMENT AND DETERRENCE) reinforcing the individual conscience through public displays of punishment to wrongdoers, etc.) satisfying vengeance, substituting controlled community retaliation for the vagaries of individual or family retaliation (blood feuds, etc.) as part of maintaining social order, reassuring citizens by the display of lawful authority, or even that lofty correctional goal of penance, reform and rehabilitation. Each of these may be seen as goals of social control or the satisfaction of individual or group wishes or emotions. While the formal criminal justice system can trace its historical origins and present justification to these objectives, there is no a priori reason to assume either that the formal system is the only means by which these crime prevention and control objectives are sought or achieved nor that the existing formal system best operates to achieve these goals under its current programs. It is quite likely, for instance, that most "crime prevention" occurs when families of well adjusted parents rear law-abiding children who grow up in wholesome peer groups where neighbors exercise social controls over children generally, and where the opportunities for private illicit conduct by children are few. It is likely that general deterrence operates as part of the general fabric of morality, reinforced by individual conscience and social norms. One could go on, but the point is that given the earlier general objectives (and we grant here that the research to say just who has these goals is not at hand) it is evident that the criminal justice system is by no means the only force which society harnesses to achieve them.

Many would contend that the formal justice system exists only because the other forces have failed, and that the entire apparatus of the law is a third line of defense defending against those who have already proved uneducatable or incorrigible under the normal rearing and restraints of primary and secondary groups; the family, peers, church, work, school, etc. (This view ignores the public service functions of the police and the civil dispute mediating functions of the courts.) Be this the case, then the evaluation of the objectives for the formal system might be more charitably seen as tests of its ability to reinforce the major influences which, we assume, account for civilized conduct (i.e. general deterrent) or to succeed where all else has failed (specific deterrent and corrections.) This comparative emphasis implies that evaluation of the formal system's achievements of general goals might best be conducted relatively, that is in comparison with other influences operating at the same time on the same people. Such a comparative emphasis in research would not require that the criminal justice system be defined so as to include all child-rearing, moral training, school supervision,

peer social control and the like--clearly an impossible extension--but only that its evaluation be considered in this larger context.

Another implication for the argument for greater inclusiveness in our view of what is to be evaluated--or how to go about it--derives from the observation that we cannot be sure that the present focus of criminal justice operations is optimal. For the most part the direct enterprise of justice is aimed at offenders, their apprehension, adjudication and disposition. Whether one wishes to argue for their correction as well as their specific deterrence (i.e. holding them so they cannot commit crimes and influencing them so that they will not) as an essential current goal of the system depends on the optimism of the discussant. Although there are good traditional reasons for these functions arising from general objectives, the current sentiment about the failures of the current system (apparent public and professional evaluations based on crime statistics, public anxiety, court loads, etc.) allows consideration for alternate emphasis and allocations of resources. Such innovation could take many forms, but would, at least, require the inclusion of altered operations within traditional systems. (Examples of internal reformations currently proposed are mandatory short penalties for any second offense, greatly reduced trial and appeal time through condensed and universal review procedures, restitution as the major emphasis for first property offenses and post release conduct in association with work training and rehabilitation and decriminalization of vice, elimination of the indeterminate sentence, etc.) The likelihood of such revised emphasis points to the need for evaluation not only to be of that which exists now but designed, as part of general strategy, to facilitate experimentation with new procedures.

As far as new efforts outside of existing formal components are concerned, one could design evaluation in support of such experimentation. Let us say the objective is crime prevention and general deterrence. Let us assume that family factors such as lack of family organization, presence of parental criminality, alcoholism and drug abuse, presence of other psychopathology (schizophrenia, etc.), absence of parental supervision, discipline, affection and humour, all create risk of greater than ordinary delinquency. (The data show that is the case.) At present the criminal justice system intrudes only when trouble has occurred (police called for family beefs, report of runaway child, apprehension of juvenile offender) although there are minor preventive programs in police juvenile bureaus. Assume that mandatory incarceration of all second offenders on index crimes works in the sense that less police time is required for apprehending offenders (this assumes

that more of the high risk criminal group will be in jail and off the streets.) Assume that more police effort can be allocated to direct family contacts with children, as for example officers being trained to be Big Brothers for families identified through neighborhood contact as ones where parental traits place children at risk of early delinquency. Evaluation in this instance has been of the existing allocation of police resources, of police training, of criminogenic families in neighborhoods. It them moves to an experimental or demonstration phase testing to see if the police resource freeing and juvenile crime preventive aims so exercised are achieved in controlled experiments.

The foregoing examples illustrate that evaluation strategy is seriously limited if it accepts existing classifications of the criminal justice system as the only appropriate targets for work. If, on the other hand, what is to be evaluated is considered in a broader context, either of related institutions, of events occurring to determine "intake" of suspects into the formal system, or of innovation and demonstration such that evaluation becomes synonymous with experimentation with new methods to meet general objectives at times and in places now excluded from the ordinary interest of the system, then evaluation strategy increases its options. It serves then not only as a means for feedback control of current practices, but helps define unrecognized existing aspects of current systems and, further, assists in innovation to test whether new approaches work better, given particular objectives, than do existing ones.

Criteria and Methods

Evaluation is usually linked to objectives. One asks, what do I intend? One then hopes to learn, what did I achieve? Yet with institutions, as with individuals, goals are often multiple and/or unclear. Furthermore, they may have to be inferred from activities and consequences rather than from statements and programmatic schemes. This introduces the notion that an evaluation cannot always be designed on the basis of given objectives. One important function of evaluation is to determine what real functions (operations, processes, gratifications) exist in the absence of accompanying or in opposition to stated goals. We speak of objectives and functions in the plural; that should serve to underscore the evaluator's attention to the diverse activities of any person, or group. Even for formal systems which may admit to only one goal, diversity is the rule. Diversity, in turn, suggests that objectives vary with persons, time and place. They need not

CONTINUED

4 OF 5

be consistent or in harmony. Indeed, one of the concerns of the evaluation of complex settings is with the array of goals and functions and the degree to which these are compatible. An important feature of the criminal justice system is not only this diversity, and the possibility of incompatibility, but also uncertainty as to whether goals and functions are compatible or not. That assessment is not easily made.

Let us say that we have, as an overall strategy for evaluation of the existing formal justice system, the goal of identifying objectives and finding measures which help us learn to what extent these are being met. How might we begin? It requires that we learn what people expect from the criminal justice system, whether or not they think their goals are being met, and what characteristics people have which allow us to anticipate what their position on this question will be. "People", of course, is too general; we are referring to particular groups or populations. Until the work is done we cannot know how to characterize these biosocially, but one does assume that, as with other social research findings, values and satisfactions will vary with such things as personality, education, politics, age, residence experience with the system, sub-cultural membership and the like. The methods for such work are those typically used to assess the distribution of positions in large populations; intensive interviews as part of pretesting, surveys, follow-up representative group discussions, content analysis of documents, analysis of voting patterns when criminal justice issues are at stake, examination of the actions of interest groups and the like. Our resulting map would tell us which groups in the country, including "insiders" in the criminal justice system itself, share views as to what the objectives of the system are. Their differences will, of course, define the groups more clearly, telling us not just the range of objectives which exist - and the degree of satisfaction with system components - but also how important these functions are to group members.

As a second step, one wants to learn which groups are not "served" by the system as inferred from the discrepancies between their views, those of administrators in system components, and evaluation of what the system is in fact doing. That work requires two steps. The first is to see which groups have objectives and satisfactions akin to the managers of the system and which groups do not. Agreement does not tell us who influences whom; managers and others may come from the same background or, alternatively, one group may control communication processes which reach and persuade the others. That can be learned. The second step is to

compare the stated goals of systems managers with process and product measures. These latter can be derived only in part from observations from diverse sources as to what happens day-to-day and whose interests are served, interests being defined rather narrowly in terms of personal esteem, gain, protection and the like. Measures of product are indeed hard to come by in social and public services and, at the present time, can only be tentatively suggested. One approach is to interview samples of people participating in the system to learn what they say they have gained or lost. Another method is to observe those participating - including, of course, those being served or processed - to see how their careers are affected by their exposure to the system. What is the income and prestige of judges compared to an age and grade matched sample of their law school colleagues? What proportion of offenders diagnosed as alcoholics emerge from prison without an alcohol problem measured two years post release as compared to age, sex and class matched alcoholics measured after the passage of the same period of time, but outside of prison? What if judges make more money and enjoy more esteem than their classmates whereas alcoholic offenders post prison are worse off than an "untreated" sample? Hardly a fair test, of course, but with enough such comparisons one could begin to see which unstated functions the system achieves (here a career ladder to success for attorneys?) and which claimed objectives are not met (the unregenerate alcoholics). That is obviously a narrow example.

What about the conclusions of an IAACP team which spent six months studying a metropolitan police department? Their still unreleased findings are neither narrow or theoretical. They concluded that although this department has a sizeable per capita ratio of police to citizens, that the departments major functions, in terms of unstated operating priorities, was to serve as a massive police benevolent association. It provided ambulance service and health care to officers and families, full pay regardless of disabilities including alcoholism, time off for any family function, disregard of high yield incomes from corruption on the part of most departmental members, rules and procedures which guaranteed against change, and very little work to do for anyone if "work" was defined as being busy apprehending offenders. (Their detective bureau made arrests at per officer rates as low as 1% per year of other departments.) As for the secondary services directed to citizens, there was a real commitment by the detective bureau to apprehending rapists and homicides for which offenses their clearance rates were equal to other metropolises. There was also a symbiotic relationship between citizens and the police, for citizens reportedly were reassured by the visible presence of officers, appreciated their generally kindly interactions, and were grateful for such atten-

tion as the citizens did receive (a visit to homes, for example, where the accumulated value of properties stolen was over $3,000; no visit, otherwise.)

The study team members may have been unduly cynical; unfortunately they did not use pre-tested techniques for rating and quantifying their observations, a common failing in visiting expert appraisals. What they reported did privately reflect differences between what the department said it was doing and what analysis of activities showed it accomplished. But the report also warns us how subjective appraisals are dubious. Inferences about organizational objectives of which come the logic, "this is what they do, this must be what the organization is there for" must be substantiated by the use of reliable observation instruments, ones not only pre-tested but retested so one can be sure that different observers of the same phenomena will report the same thing. This is the requirement for objectivity and replicability of any process study which is to yield data on agency functions and, thereby, inferences as to the objectives of participants. Such studies in the justice system are very rare indeed.

There is probably no dispute about the wisdom of observing big city police at work to find what the system does for them, although taxpayers are not likely to agree that the major purpose of a department is to be a benevolent association for the police. Taxpayers might prefer to see services directed toward more "legitimate" clients. The victims of crime surely are such a population, one that, at least, some part of the justice system is there to give high priority attention. Victims can be a very useful source of performance data about the system.

Ask a random sample of victims in Westown one day after they have reported their victimization what it was they wanted from the police. Immediate response? Courtesy? Sympathetic understanding? Attention to recording the details of the crime? A report on the progress of the investigation? Insurance claim backup? Property recovery? The satisfaction of knowing the offender had been caught and punished? Restitution from the offender? Assurance that they would be protected against further outrage? Then ask these victims how they feel about the police and courts two weeks later, and then six months and a year later. Have their general views about crime or the system changed? Have they influenced the opinions of their friends and families? Have they changed their behavior better to protect themselves? Provide this victim data to the local police department? What do they do with it? If problems in police service are apparent, do the police managers

acknowledge them? Six months and a year later does one find that the department has made any changes seeking to improve upon deficient services? If not, why not?

The foregoing suggestions, the first for a wide-based population survey and the last, victims follow-up and police response to victim data in one local illustrate quite different approaches to evaluation. The former is of broad sociological and political interest. It is the kind of thing on which political parties and legislators base their promises and programs and the stuff of which professorial lectures are made. The latter is local and particular. It is even being done by a few advanced police departments. It is a feed-back procedure useful to attentive departments, probably welcomed by harmed citizens, and exemplary as a product measure for one kind of police service; how are victims handled. The evaluation strategy would tell us which approach met the knowledge need.

The crime victim is a client of the system whose evaluations are often overlooked. Her needs and dignity are,too,if current concern with the cruelty of trial procedures for rape victims is a measure. The victims' financial needs are only of recent concern, as the efforts to provide for compensation to victims testify. And restitution as a measure to aid victims as well as to create a moral relationship in the community for the offender is rarely considered. Each point here sets forth an objective which some would advocate and for which evaluation measures are readily available. A dark number study would tell us (plus or minus an X error rate) how many rape victims do not report their victimization because they say they fear the indignity of police interrogation and or trial. Assessment of the status of compensation laws and the circumstances governing allotments would tell us where the legislatures stand on this issue, while in any court we could determine how often judges require restitution. When the offender is indigent we could also learn how clever judges are, should they state that restitution is one of their values, in seeing to it that probation officers find jobs for offenders so that they are income generating.

This latter comment introduces the reality of job markets and vocational skills as well as the availability of industrious probation officers, work furlough programs and the like. The judge, operating without community resources, is hardly able to require restitution of unemployable offenders. He is also hardly likely to expect such offenders to stay crime trouble free very long. We see that the evaluation of the attainment of particular objectives each requires a probing in context. For rape victims one needs find unreported victims to learn what they

fear from the system and how they came by such fears. For compensation programs to victims by the state one must look to the political arena and the operations of lobbies, budgets and felt priorities. If the objective were restitution, the function of the justice system toward that end could also be understood only if one knows the judicial alternatives, priorities and court and community resources. One concludes that the evaluation of particular objectives force is best done in context. What is required in such instances is knowledge of the larger setting, of the forces determining participant behavior at any given time, and of the relative priorities for given actions.

The appraisal of the criminal justice system to date has benefitted from research which does go behind the obvious, but not sufficiently so to allow us to be sure of what is happening even in major sectors. An example is the study of Canadian judges by Hogarth who determined that the objectives which a judge held for the corrections system did indeed predict what he would do. His background, history on the bench and cognitive styles were also correlated with sentencing, making sentencing idiosyncratic indeed. What predicted his conduct little, within a given range of offenses, were the facts on the offense and offender. Were the abstract objective "justice" and were that assessed in a lawyer-like manner by our non-appellate review here, we would perhaps say that so subjective a system of justice was not working. If the objective were only to hope for consistency not among judges but between one judges own past and present sentencing behavior, then there is evidence of a kind of efficiency. We see in this study how research on sentencing behavior by evaluating judges tells us little of what larger objectives might be met by the system in operation, but how such evaluation does allow insights into the forces which account for what does happen. That is the kind of process understanding for which we argued in the introduction.

Victims are system clients little studied, judges are systems managers poorly understood, whereas offenders are the systems' products whose views are remarkably ignored. There will be objection to the notion that the offender is the product of the criminal justice system. The criminal, himself, may claim it was his mother's fault; his attorney may hold it was an unconstitutional procedure, whereas the social reformer will blame the offender's deplorable plight on poverty, lack of opportunity or oppression. Nevertheless, if one is interested in product measures, that most often counted by the system is the detainee become suspect become charged case become trial case become convicted offender. The system measures itself most often by the number and speed with which it collects, labels,

holds, disposes of and then often recollects these persons deemed perilous to public interest. Why then does it never ask them what they think? One would expect that as consistent clients of the system, some of them on hand to view it on and off and in and out for 50 years, all of them certainly sensitive to its ministrations, that there would be routine inquiries to the mill's grist. Was the apprehension efficient? Was there brutality? How many apprehensions were missed when they could have occurred? Were the police charitable, drunk or simply too busy to make those other missed arrests? What's the jail like compared to others? Is the inmate penitent? What has he learned today about car-theft or forgery? Has he learned enough to counsel his defense on the small points or assist in the selection of a judge? And just between us, when he is released tomorrow will it be back to robbery to make money to pay the lawyer or what? Each query bears on aspects of the system which might be held as partial goals, or at least as values to which the system says it would conform. Why not ask the outcast experts? In a comfortable city, the detective knows where to go to find who is holed up where to snitch on a red hot case; he asks the professionals and habituals working the street. Or in the jail, who is it that's planning the escape or buggering the new boys? Ask the trusty stoolie and find out. The phenomena of not systematically querying offenders - those arrested and those self-admitted still working or now retired from the streets - to appraise the system's functions and goals is as blatant an ostrich act these days, as is the avoidance of victim interviews, as was - and sometimes is - ignoring the dark number of crimes committed but officially unreported. In all instances the significance of not asking is, in itself, a kind of measure. One presumes the system has a very limited capacity to query itself. Probably it has a limited capacity to digest information. Probably it prefers not to learn about more problems and a problem occurs whenever anyone says it does. Whenever there is a discrepancy between goals and achievements, between what the system says it does versus what it really does, between the formal claims and the informal realities between what people want and what they get. Who is in a better position to generate demands and prove discrepancies than are victims, unreported and reported, and offenders, apprehended and unapprehended. That all are ignored has told us something about the politics of evaluation by the system itself.

It would be a mistake to consider that only external clients are excluded as experts whose opinions matter. The official participants rarely are asked their

private opinions about their own organizations and only recently, with Banton, Wilson, Reiss, Bordua, Rubinstein, have observers done the looking at participants, mostly police. There is a caveat. The police in their fraternal organizations, the chiefs with their professional lobby, the prosecutors with theirs, the guards and prison administrators in assembly, the judges with their several associations; they do have voices and strong ones, too. These speak to the issues which trade unions anywhere address, supplemented by moral convictions on crime and such political ideology as is correlated. These voices speak to the rest of society, mostly in the direction of the ears of the legislative and the fourth estate, but they speak about what others should do for them, as unions must. New laws to make their jobs more meaningful or better paid, to make their facilities more efficient or comfortable, new guarantees and benefits to make lives more secure and old views on what and who is wrong with society. But these assemblies do not ordinarily speak to evaluate the institutions in which they work in terms of the objectives which others have set forth. The wardens' group wants newer, bigger prisons with larger, better paid staff (see Milford), the objective perhaps being rehabilitation. But the wardens' assembly does not fund and report on a self-study based on the survey of prison goals set forth by the nearest townspeople, or the prison employees, the prisoners themselves, supplemented by a critical appraisal of recidivism statistics, prison violence data, or prison hospital morbidity figures. It is the measure of performancy by others' standards which is lacking, lacking in the justice system just as it is almost anywhere else. Chairman Mao's self-criticism program is, perhaps, an exception, but a reading of Lifton suggests that not all participants there are enthusiasts either.

It is clearly too much to expect organizations to criticize themselves in any serious way, just as it is unrealistic to anticipate that the managing participants in a system gratuitously distribute their powers by seeking unsettling appraisals from outsiders. Customer satisfaction or consumerism has not yet reached the justice system, although we anticipate that fashion will one day see that it does. In the meantime, it would appear remiss for oversight agencies not to fill this function. As experience comes to show justice agencies that they can live with a wider view of themselves, they will be encouraged routinely themselves to sample consumer goals and appraisals.

Oversight agencies will, of course, demand more than files full of clients good ideas. The uses to which evaluation data are put depend on records utilization

for planning And that is another story. There are several books on police records keeping, including the "green bible" on municipal police administration. Each shares technical excellence about who is to record what on which form filed where. What is missing are studies of what really happens in records keeping. We have ourselves conducted such a study in a public mental health agency (Blum and Ezekiel) and have seen what does happen. Whether its purpose is to keep track of clients, events, personnel activities, costs, complaints or outcomes, there are errors at best and in some settings complete chaos. Insofar as routine records form the mainstay of justice agency self-monitoring and external reporting, one would hope that the emphasis in records keeping would be as much as on error sources and their control as on ideal forms. Yet well-kept records can be worse than poorly kept ones if they take time for no known purpose. One wonders when the ordinary justice agency last reviewed its records system with an eye to determining purposes, errors and utilization. One wonders,too, how many of these unstudied agencies would doubt the accuracy of such records as they do use to report index offenses, count patrol cars, describe offenses and the like. Insofar as administrators rest upon internal records for their kind of evaluation, we believe it is of paramount importance to show how model records studies can be conducted, how error sources can be identified and countered by personnel selection, training and supervision as well as equipment purchase and maintenance (e.g. computers) and how records systems can be revised as no evaluation needs arise.

Our reference to internal records does not indicate that justice agencies need keep no records of external affairs. In police agencies these have typically been limited to often dubious intelligence files, press clippings, equipment catalogs and very small professional libraries. What are missing are those aforementioned studies commissioned to survey dark number crimes whenever evaluation requires that crime reports be checked against crime commissions. Missing, too, might be appraisals of departmental performance, when these are in order, offered by known offenders on the streets, and by samples of citizens - perhaps called on at home by officers as part of community relations. We do not discuss here how other external file data may be both useful to and lead to measures of departmental performance by linking agencies to citizens in ties of work and consultation.

There will be some aspects of agency performance that will not be linked to specific and long-term objectives, but will be considered right in themselves. Unfortunately not all parties will agree on what is right. One of the functions

of administration is to monitor the status of affairs regarding some of the predominant values. One of the functions of external evaluation is to determine if managers share these values with the citizenry and, whether they do or do not, to see if they can assess the state of their organization. Such general values include efficiency, safety, cleanliness, health, honesty, personnel morale, courtesy, and the like. For a jail, they will include nutritious food, cleanliness, segregation by age, sex and trial status, lack of crowding, supervision to prevent staff or inmate violence and to detect illness, lack of oppression and violation of civil rights, etc. For a court one hopes for up-to-date calendars, mentally competent judges working a full day, sober bailiffs, sufficient prosecutorial staff, adequate indigent defense counsel, etc. These are very simple requirements, none of which are the grand objectives of the criminal justice system but all of which are likely to be demands which society imposes for conformity to its general standards. One would think that such demands would be easily met, especially if the longer broader term goals such as crime prevention or offender correction cannot be. Yet as the work of the President's Crime Commission - and numerous committees before and after have shown - many justice agencies do not meet the minimal standards of competency, decency, lawfulness or efficiency. An IAACP national survey team characterized most of America's police departments as "underdeveloped countries too backward to care." Federal jail surveys have found hundreds of jails unfit and unsafe for humans. The American Bar Association study groups have, on occasion, expressed doubts about the efficiency and competency of at least some judicial selection methods and, of course, many reform commissions and most observant citizens have their doubts about much of the sense, consistency or efficacy of existing criminal codes.

One may feel it not succiciently fancy for an evaluation strategy concerned with the major goals and sophisticated methodology to address itself to the nuts and bolts issues of minimal standards which have plagued the justice system probably since Hammurabi. Yet if the system cannot deliver rape free jails, policemen who are at least discreet if they must be corrupt, judges who are not so senile that they relieve themselves in the courtroom corner (not hyperbole but a recent event in California,) probation records which at least get the name and sex of the probationer right, well, what is one to expect in regard to the more spectacular achievements?

Evaluation addressed simply to minimal operating standards has no reason to expect to contribute more to the achievement of affairs right in themselves

than have the dozens of prior and on-going study groups, commissions, accreditation bodies, investigative reporters, MA theses, and the like. This does not exclude its worth; like the criminal law (if one accepts the theory of general deterrent) the description of discrepancies between what is and what should be serves to educate communities and their laggard justice agencies. It serves as a hook on which to hang public demands for improvement or requirements to be met before federal or state aid is granted. Yet it is possible that an evaluation strategy can seek to stimulate achievement through analyzing the reasons for the chronic discrepancies between standards called for and operating status in fact. Necessarily that is research on the social, economic, psychological and political context in which substandard agencies exist and are managed. Necessarily such research will suggest demonstration projects designed to bring about change through altering the fundament.

In this regard one thinks of LEAA itself, with its billions of dollars expended to achieve just such immediate goals; the improvement of the operating characteristics of justice agencies. Have these dollars helped? Which agencies now meet standards that they failed to meet prior to LEAA funded (via State agencies for the most part) assistance? But which police departments still have poor management? Which jails still serve maggoty food? Which courts continue to show sentencing disparity so great as to demonstrate that the justice is but whimsy? Here is an immediate and profound evaluation need: to review the pattern of LEAA funding, to see historical data on before and after local agency characteristics, to see what standards are not met even after funding, to see which substandard local agencies did not improve with funding, which improved without it, and to analyze the reasons for the failures compared to the successes. As an immediate evaluation strategy for LEAA (regrettably made more difficult by its imposition retrospectively rather than as part of prospective research) this analysis of the circumstances under which fiscal assistance does and does not influence agencies to meet minimal standards will be most worthwhile.

An important feature of many evaluation enterprises is to keep in mind the dynamic aspects of the social field in which change takes place. What is particularly important is to realize that it is not sufficient to measure the extent to which objectives are achieved without also measuring what else happens as part of this process. There can be hidden costs and benefits. A good discussion of attention to social costs in the pursuit of the criminal vice law is found in

Packer's THE LIMITS OF THE CRIMINAL SANCTION and, with a sounder data base, in Kaplan's MARIJUANA THE NEW PROHIBITION. Both illustrate that neither the policy-maker nor the evaluator can afford to attend only to the objective and its achievement, for all programs of change bring side effects. These side effects may be of any order of magnitude far greater than the objective itself.

The design of evaluation studies which embrace this awareness of dynamic effects depends on the discipline from which the evaluator comes and the kind of concerns which he, along with administrators and policy-makers have about the operation of an agency or other component. Generally the attention to side effects will require knowledge of social psychology, organizational structure, and economics. The side or dynamic effects first measured will be those known to have occurred elsewhere when similar changes were introduced, but will also include measure of shifts away from or towards additional goals, short and long-term, in the system component. As an example, let us say that concern with pollution by industry has become so great in a community that local legislation is passed assigning to the police a pollution control responsibility. Political interest is high so that men are taken from the traffic bureau and assigned to patrol cars which monitor effluents day and night. They cite many violators which are processed by the courts. In the meantime, the undermanned traffic bureau has not been able to complete its paperwork on accidents with the result that insurance company claims and settlements are blocked. Angry citizens and insurance representatives protest and the city manager, acquiescing, returns the pollution patrol to the traffic division.

The example but demonstrates not only a very simple side effect which would not be noted if one only looked at the short-term enforcement of the pollution law, but it introduces the role of evaluation in the consistently difficult problem of priority setting for resource allocation in systems with limited capabilities. Thus the description of capabilities within a given agency becomes an important evaluation activity, defined not only in terms of the achievement of goals but immediately in terms of such things as funding and equipment, the number, training and competence of personnel, the adjacent resources which are used to supplement normal activities, be these mutual aid agreements in a police department, volunteer visitors in a prison, or lawyers for indigent defendants provided by the rotating list of the local bar association. Once capabilities are described numerically, one can not only calculate various service ratios which may prove to index and correlate with goal achievement through comparative studies (e.g. a recent limited

study concluded that increasing police per capita ratios did result in reductions of index crimes) but one can begin to estimate the efficiency of activities in relationship to an array of objectives. Such an exercise can guide the establishment of priorities for resource allocation. Consider for example a probation service with ten officers and an average case load of 120 parolees per officer. Assume that an evaluation study learned that by characterizing officer characteristics psychologically, let us say degree of personal warmth, involvement and authoritativeness, and linking these to offender age and criminal record, one could predict which officers were best assigned to which offenders so as to minimize the revocation rates in two offender classes; burglary and child molestation. Say there are three such officers out of the ten. The first strategy of assignment is to consider the warm-involved-authoritative officer a special resource for the three offender groups and to assign the three officers to these offenders. Let us say additional research shows that two further sub-groups of other offenders emerge, one which has a high revocation rate linked directly to how many cases the officer supervises. The more cases supervised, the fewer revocations. Another sub-group has greater revocations the higher the case load. Let us assume that car thieves are in the first group and narcotics dealers in the second. Let us also assume that an objective of the department is to minimize revocations and that this objective has been tested in two ways for side effects. One test was of crime rates for these offenses during a period when supervision case loads were varied, the other was systematic interviews with prosecutors, detective bureaus and judges to learn of their informal observations or sentiments. If there were no adverse side effects, then the strategy of assignment based on evaluation findings would be to give the remaining seven officers case loads including both car thieves and addict dealers with instruction not to supervise the addicts but to spend much time with the thieves.

There may be an objection to the strategy if one involves a new objective for the probation department. That could occur if a further study had asked why it was that supervised car thieves didn't get revoked whereas supervised addict/dealers did. Assume we found the car thieves were 19 and 20 year olds who responded to authority and matured out of this delinquency fairly rapidly (the younger majority of car thieves would, of course, not be on parole with an adult parole authority) whereas the addict/dealers were chronic drug dependent delinquents whose continuing criminality would be detected by any parole supervision. This is a typical situation. If the administrator is guided by one goal and his evaluation demonstrates

achievement, the lowest possible revocation rate among those classes of offenders on which it can be shown intervention (personality, case load) make a difference, he should stick to his guns. If, however, he adds an additional objective, a kind of "equity" which increases the risk of revocation for both addict/dealers and car thieves by equalizing supervision, then he would return to the old and "inefficient" pattern.

If another kind of evaluation study determines that the reason the administrator seeks an equal and higher opportunity for revocation for the addict/dealer group is his fear of newspaper censure, then he may be wise to evaluate the views of his local editors about his department functions. Perhaps he will learn that a quid pro quo of an immediate notification by his office to the press for each and every revocation will reduce the newspapers' court and police reporter's work load sufficiently, and provide filler for the presses, so that a modus vivendi is worked out. The editor will not be unduly critical. That stage is one not of evaluation but of political negotiation, yet it is undertaken on the basis of a primitive evaluation of what can influence the newspaper's editor's views of and response to the parole unit.

The example shows initially how evaluation research, in considering interaction effects between such factors as personality, case loads and revocation rates becomes identical with field experimental studies typically done in the social sciences. If an objective, in this case reducing revocation rates of parolees, is sufficiently important then one must expect to do research which isolates the variables which influence rates measuring attainment. If those variables are found to be under any administrative control (as officer assignment is, whereas a variable such as parolee age is not; and local employment conditions may or may not be) then such research can lead to administrative action which will increase success in obtaining the objective. A second conclusion from the example is that one may learn that the costs of achieving a single objective in a dynamic situation reveal the existence of operating restraints and short-term objectives (e.g. avoiding press criticism for ignoring addict/dealers.) Here, too, evaluation can help both in planning action and in follow-up to see if results are as anticipated.

The example led us to conclude that variables which are under administrative control can be acted on so as to improve the achievement of objectives, providing the full array of constraints and goals is understood and examined and that direct research is conducted as part of evaluation. But what about a situation in which

research shows that the objective does not seem to be a function of variables which are under administrative control, that is, the causes or antecedents of the phenomenon estimated by the outcome measure are immune to the operations of the system which, presumably, was designed to act up on them? This is likely to be, or appear to be, the typical situation for many of the formal goals for the criminal justice system. Consider that in Kansas City the presence or absence of patrol has been found to make no difference in criminal activity (although the availability of patrol in response to calls was not reduced nor do we know how much the experiment was advertised to burglars.) Recent data suggest that variations in correctional experience (length of sentence, treatment, vocational training, etc.) makes not the slightest impact on recidivism rates when one controls for offender characteristics. Observation from police strike situations show that while some crimes go up (looting, burglary) some go down (a reporting function?) and some are constant. A study of judges shows that sentencing practices, within a class of offenses, are almost unrelated to the offender and his case, but depend upon the attitudes and background of the judge. Add big city plea bargaining statistics which, when 90% of the cases are settled by the advocates, effectively exclude the courts from any role in dispensing "justice" either in determining guilt or innocence or in setting penalties. Add, too, corruption data arising from city police department reviews (cf Knapp Commission, Reiss) indicating that the majority of officers engage in chronically felonious behavior. Add, for police departments, clearance figures not above 2% in some cases for such class one index offenses as burglary, or the estimate (Bradon, New York City) of the detection of homicide at a rate of no more than 10%. If arrest rates are 70% and convictions 50%, then for every 100 murders one may expect about four convictions. With conditions even approximating the foregoing one would conclude that if major goals of the criminal justice system include crime prevention and detection by the police, effective correction by prisons, or the careful consideration of all cases being prosecuted so that justice can be judicially assured, sufficient evidence of failure exists to suggest that the existing criminal justice system in at least some locales is not in administrative control of the variables leading to the outcome measures (crime, recidivism, justice assured by judicial hearings, prevention, detection and subsequent apprehension of offenders.)

The Grand Questions

Assuming now that there will be agreement that the foregoing do constitute some but not all major objectives, we now face the question as to what impact the

criminal justice system does have on their achievement? The policy-maker, legal scholar administrator and researcher also ask what impact change in the system introduced from the outside (legislation, reform, LEAA support, etc.) may have on the attainment of these objectives? At the present time there are three replies to these questions. The first reply, that of the ordinary social scientist, is that they are posed in such a general way as to be unanswerable if the answer is to be based on specific knowledge derived from quantified observations based on careful evaluation which is integrated into what one might call a policy calculus. Research on the criminal justice system is not yet adequate enough to tell us what is to be included in a description of the "system." It is insufficient to assure us as to what are the relevant processes within components that have been identified - relevant in this instance defined as having a demonstrable relationship to outcome data. And research to date is insufficient to tell us what the impact of existing components is on objectives because we lack data on relative influence on and interaction among the components, because we have but awkward measures of outcome, and because we cannot even be sure what the degree of consensus is on objectives. Further, even with consensus on objectives, we cannot be sure that the stated objectives represent the full range of consequential goals, for some of these must be inferred from an analyses of functions.* The arguments in support of these conclusions have been advanced in the preceding pages as we have considered evaluation itself.

* A further observation in support of the notion that objectives can be discovered through what sociologists call "latent function" analyses comes from Erikson who reviewed crime convictions in a closed and stable Puritan community. He found that the kinds of crimes charged varied dramatically over time (e.g. witchcraft, homicide, sex offenses, property crimes, etc.) but that the rate of convictions remained constant over decades. He argued that there was no reason to assume behavior change within the community and, further, that the allocation of resources to the criminal process being stable, that the system was itself not in disequilibrium. He argued that the community required a certain rate of convictions per annum and per capita, but that the conduct indicted was, within a range of disapproved acts, simply a matter of fashion. The conclusion is that the criminal justice process is itself essential to a Puritan society. Its unvarying rate suggests that who is charged for what is secondary to the need that arrest, conviction and punishment take place. Why there must be "deviants" is a matter of speculation.- Is it because the super-ego must be reinforced by public rituals? Because general deterrent depends on a given rate of public information as to consequences? Because punishment is itself a gratifying recreation? Because justice specialists, once the occupation is created, control sufficient power to guarantee jobs for their successors? Who can say?

The second more optimistic reply is that for a given sector of the system, or indeed for a linked series of components, one does have measures which assist in defining objectives; describing those features within the system that do and do not bear on these, do identify additional influences outside of the system which also bear on results, and finally, one can develop improved means for measuring objective attainment. The impact of planned intervention on any part of the system can be measured either in terms of short-term immediate effects on the apparatus itself, effects on the surrounding environment, or on the criterion (goal) measures themselves. A number of such studies have already been done. They share the characteristics which most contemporary social science research have; they are limited to particular locales, to specified operations, and to one or a few refined outcome measures. The results of such studies - which constitute much of the literature of criminology and some from the parent disciplines of sociology, psychology and the criminal law - give us our present "state of the art" picture.

Social science evaluation, while the theme of this section, dare not ignore the highly persuasive findings derived from other forms of observation and analyses. Consider the President's Crime Commission, American Bar Association commissions, the Knapp inquiry, IAACP department surveys, Federal jail inspection team reports, investigative reporting on police, courts or prisons, the work of legal scholars and the like. Such inquiries, of great scope, magnitude and importance, have contributed much more than social science to our knowledge of the criminal justice system. Because of their breadth, and because their virtue is that they escape the triviality which often characterizes social science research, it is likely that these non-scientific enterprises will continue to be the basis for the systems-wide appraisals which most state and national policy-makers require. It is likely that the contribution of the social sciences will be to provide refined measures which answer questions about specific impact in specific places and, more important, to offer new concepts about how to think about crime and justice in society.

The third reply to the major question about the attainments of the criminal justice system bears no "expert" answer but instead will reflect the conclusions of any observer based on his vantage point, information, goals and interests. From our standpoint two simple conclusions seem in order. One is that the system, overall, must "work" because it exists. "Work" here means that it is an accepted part of our society, massively supported by tax monies, employing hundreds of thousands of personnel and processing millions of citizens a year, and not challenged as to

its legitimacy by any noteworthy citizens. Indeed, even our revolutionaries would likely retain police, courts and corrections, only redirecting their efforts to oppose those defined as enemies and deviants in their ideologies. We suspect that the number of anarchists who advocate dismantling the entire apparatus is few indeed. Existence and support then prove efficacy, even if it is only the efficacy of tradition and virtue. The second conclusion is that the system obviously does not "work." Surveys of citizens, participating officers (police, courts, corrections and their associates in professions,) concerned authorities (legislators, mayors, etc.) and offenders,easily document dissatisfaction and the widespread conviction that the goals of crime prevention, detection, deterrence, corrections and justice are not being met. Indeed the fact of innovation - and possibly its rate - may be taken as a measure of dissatisfaction as well as reflecting the evolution of new goals. Whether the innovation be LEAA itself, DEA, the Police Foundation, or release, the Criminal Justice and Standards Commission, delinquency legislation, reform of the criminal code, expanded police departments, these social changes testify to public and professional consensus not so much about what to do, but that "something should be done."

The extent of dissatisfaction no doubt varies. Perhaps those in law-abiding homogenous communities will be only modestly displeased while those in high street crime areas may be upset indeed. Personal upset is, of course, not simply a function of actual experience with or risk of crime but is clearly dependent upon the aspirations one has for public safety and justice. Such aspirations are possibly linked as much to social views of the proper nature of society, the role of government, and the nature of man and to personality features such as punitiveness and insecurity as they are to knowledge of actual crime rates, crime control and justice costs in a comparative context. Indeed, one of the important potential contributions of research will be the determination of the sources of dissatisfaction with the justice system on the part of various groups. In any event, the present failures of the system as set forth by most observers are evident enough on the basis of the gross data before us. What is not known is the extent to which any intervention by government on the system, will have an impact on either effectiveness, that viewed as achieving the grand goals, or propriety and efficiency, as static or process measures.

Measuring

In the social sciences preoccupation grows with the sophistication of measurement. Research designs become more complex, apparatus becomes more expensive, statistical tests multiply, and the refinement of logic is virtuous in itself. There are many who hold that "if it can't be measured, it ain't there" although the truth of this proposition in relationship to a theory of knowledge is debatable. What we accept for the purposes of this paper is that a number of reasonable tools now exist which can assist in the evaluation of the criminal justice system, but that these tools are rather rarely applied. We suspect that the proportion of funds spent on sophisticated evaluation-related research on the criminal justice system compared to the expenditures of the system itself, or even more dramatically, the costs of crime overall, would be low indeed, probably much lower for example than cancer research outlays as a proportion of cancer treatment costs or the overall economic costs of that disease. We also accept, for the purposes of this paper, that preoccupation with numbers or the art of their generation, is easily come by and need not be relevant to much of anything including public policy. It is not simply that quantification in evaluation can be misleading or that their numerical equivalence cannot be assured (would the same index crime rates in Boston and San Diego reflect the same dark number rates in those cities; is equal value to be assigned to a homicide apprehension following a family fight and voluntary confession as to the arrest of a professional hit man?) but that one can over-refine the measures of events which are passing crude. To seek to measure the excellence of justice by calculating the rate of reversals on appeal for courts in town A versus B on criminal cases whilst overlooking an unsupervised plea bargaining rate of 90% in A and 0% in B illustrates that. It is, in the larger sense, a matter of fitting the measure to the need and that, like most of what we have discussed here, is evaluation strategy.

It is beyond the scope of this paper - and indeed our competence - to set forth all possible research designs for evaluation work in the criminal justice system, nor can we list those which have been most useful for police-makers. Learning that would constitute a project in itself. It is also not possible to offer the consensus views of social scientists, criminologists in particular, as to what constitutes either the best evaluation measures, the best work done to date, or the portions of the criminal justice system which are most in need of

examination. Not only has that survey not been conducted, but we would recommend against it. Fashions rule in this profession as in any other, consensus is no proof of creative intelligence, and the factors governing popularity contests either for methodology, or investigators, are uncertain enough to suggest caution. In the light of these limitations we can only touch on major research approaches. Each approach requires at least some limitation of setting, at least some control over variables either through administrative or experimental manipulation and/or through statistical procedures, the use of standardized instruments or otherwise reliable and valid measurements, and the limitation of generalization to like settings, exclusive of major discoveries and provocative speculation.

The approaches which we expect to have the widest application will be as follows:

Surveys designed to sample from among populations to learn how traits of interest are distributed. From this,one constructs groups who share particular traits and, thereby, maps a social geography. One can learn with considerable accuracy about the distribution of personal experiences, attitudes, expectations, opinions, values, anticipated specific conduct, etc.

It is important to distinguish between the survey method itself as a way of reaching a representative sample and the particular instruments used to assess how a given person is to be described, as for example his attitudes, his personality, his criminal experience, his preference for legislation, his dissatisfaction with the police, etc. In each instance, one needs employ methods - be these interviews, questionnaires or tests which are pre-tested, known for their reliability, and also known to be valid.

Panels: A panel is a group of persons selected from out of a population of interest who are observed, interviewed, tested or otherwise used as a gauge over a period of time. To control for the effect of being a panelist, most studies require other matched or randomly selected people from the same population to be observed just before and after rather than continuously over time as occurs with the panelists. Panels can be particularly sensitive as informants registering their experience, opinions and the like during the course of efforts to introduce change in an agency or community. As with surveys,the particular tools used to measure changes, be these diaries, computerized experience recording apparatus, attitude scales or psychophysiological measures, must always be known for their reliability and validity prior to being used.

Observations on organizations or groups (including individual members.): A variety of methods present themselves including participant observation, sampling of members with questionnaires or for interviews, the use of previously quantified data on organizational activities and the development of new activity measures.

Career studies: These examine individuals or groups, thereof, over a course of time. They can be retrospective, prospective or cross sectional. They can utilize survey methods, participant observation, biographies, diaries, interviews with the subjects, interviews with those who know the subject, data from institutions in contact with the subject (work, health, arrest, school records, etc.)

Personality assessment: Trait (used here in the sense of any attribute observed) assessment among individuals (usually classified into groups) who are observed for their reactions to particular experiences, who are evaluated in terms of their prospects for engaging in specific future conduct, or who are assessed for their developmental changes over a long period of time, can all be important. In particular, insofar as a given personal characteristic can be shown to be the intervening variable between the actions of a system (e.g. jail) and later conduct, then the design of actions can proceed rationally. Consider the findings which show that remorse and anxiety peak for incarcerated offenders within the first few days or months of prison and then decline. If it could be shown that this peaking were an optimal time for release, then sentencing practices would be changed.

Biomedical, psychophysiological measures: When preventive intervention is expected to alter a trait which is otherwise associated with a high risk of an undesirable outcome, when intervention itself is feared to produce an undesirable outcome, or when intervention aims to alter a condition already judged as undesirable, biomedical measures are in order insofar as the condition is itself a physiological variable or if a biomedical measurement is correlated with the trait of interest. (A similar model of course applies to personality and attitudinal measures.) Examples occur when one uses bioassays to evaluate methadone maintenance program success, when one uses blood alcohol levels to test the efficacy of new drunk driving laws, when one evaluates the health of jailmates on intake and later, if one measures seizure rates and accident rates among licenses drivers taking or not taking an anti-seizure drug. If one were to give polygraph examinations (lie detector) to all police applicants okayed for hiring, where half had (randomly) been subject to new screening procedures and half had not, to determine if the new screening excluded felons or credit risks better than routine methods, that would be an evaluation using a psychophysiological measure as the criterion of program success.

Experimental situations: Whenever one wishes to test the outcome of intervention on a given behavior sample under highly controlled conditions, one will employ an experimental situation rather than the more uncertain natural or real life measures. Prior to release of high accident drivers sentenced to mandatory driver training, one might wish to test for minimum achievement by comparing their driving in simulator circumstances before and after training, requiring certain performance levels before restoration of a license. Inmates given psychotherapy and put in isolation for violent interracial strife in the prison yard might be exposed to experimental provocation observed in a one-way mirror observation room to determine if they have sufficient self-control not to

respond violently. Children instructed in a morals and law course could be tested, before and after, under situations where they have opportunities to cheat but are, unbeknownst to them, monitored. Such situations are not often used but can be tailored for narrow uses. Under different conditions such experiments test hypotheses about what interventions might work to alter views, conduct, personality or physiological responses or conditions.

Public records: Records of arrest, legislation, appropriations, and the like can be used retrospeetively to evaluated impact. Content analyses can identify changes in themes in legislation, political speeches, editorials, letters to the editor, front page news priority and the like. Illustrations occur in the evaluation of the consequences of the absence of a police force in Copenhagen following the police force's arrest by the Germans, in describing the effectiveness of drunkenness control legislation in England which set hours and times on pubs, and which restricted sales and increased the price of gin, or in measuring the impact of the British Indian campaign against thuggee (from an estimated 40,000 to zero Thug homicides a year in 20 years, praise be to the Raj.) Erikson's study of Puritans used public records and their content analysis to learn that charges changed but arrest rates stayed the same over many decades. Studies of the value of punishment and vengeance in writings by fascist, democratic and socialist leaders correlated these emphases with the legislative programs introduced (and Rokeach has demonstrated that a values inconsistency confrontation experience can change values in students for a six months period.) Currently public records of appropriations, crime reports, arrest, clearances, court dispositions, inmate populations, recidivism constitute the bulk of the measures of the crime problem, the operations of the justice system and inferences about impact. As earlier noted, public records are likely to be strongly biased by random and non-random errors. Their use for contemporary work should be approached with caution unless careful analysis has been conducted to show what influences records keeping and what phenomena are systematically excluded.

Methodological studies: Some work concentrates on methods and statistics themselves. Sellin and Wolfgang's new method for rating crime severity, based on a psychophysical analog requiring subjects to discriminate among crimes according to seriousness is a most careful example. So, too, are studies which reveal marked differences in relative rates of prevalence (e.g. alcoholism) which different statistical tests yield. Were one to determine intercorrelations among discrete measures in one community, as for example police morale, victim satisfaction with police response, offender appraisals of police competence, prosecutor appraisals of police evidence preparation, per capita taxation for police support, police salaries, police corruption rated in response to controlled opportunities (i.e. experimental entrapment) etc. one would be embarked on a methodological enterprise. The aim might be simply to see which measures correlated one with another, to see what common factors emerged; it might be to see which ones in combination best predicted another or, by (path) analyses, one might hope to infer which were antecedent (causal) to others.

Organizational records: Most agencies keep records about pertinent events such as personnel turnover, absenteeism and sickness, auto accidents and repair costs, complaints and citations, down time for expensive equipment, loss to inventories attributable to waste, pilferage, accident, performance measures for personnel such as arrests or citations among police,

cases adjudicated among judges, ratings by supervisors, performance on examinations and the like. Insofar as the intervention imposed seeks to change some aspect of any agency, it may well be that these and other organizational records can be used in, before and after, evaluation. As with any such effort, the goodness of the records must itself be established.

It can also occur that the records of organizations not directly acted upon by an intervention effort can be used to measure change. Examples are the liquor sales records before and after a temperance campaign, the hand gun sales rate before and after publicity about street crime reductions, venereal disease rates reported to public health agencies after a police campaign to close down brothels, Custom's seizures of heroin before and after a special interdiction effort abroad, or the number of minority applications for justice-related jobs before compared to after the inauguration of a police community relations program in a minority neighborhood. One might be able, for example, to examine the rate of use and growth of private security advisors and personnel (guards, burglar alarm installations, etc.) in two communities, one with a program of mandatory sentences for all second offender shoplifters and burglars, the other without (necessarily controlling for store size, inventory values, security service prices, etc.)

Covert and/or deceptive measures: These are used when one does not wish a subject or group to be aware that observations are being made; they can utilize humans or technical apparatus for the primary observations; they can be done, as many experiments are, simply by giving a false explanation to subjects; they can utilize willing informants who make covert observations on a group - heroin dealers or policemen for instance. One can plant observers posing as police officers in a department or lawyers posing as new employees in a prosecutor's office. They can employ hidden gadgets, as for example, tape recording juries with the judges' permission (that happened only once!) using lie detector apparatus built into a chair about which the subject is unaware, taking fingerprints from "anonymous" forms used in a prison survey, filming from concealment the stripping of a car planted as "abandoned" on a street. The ethical problems in such approaches are evident and "human subjects" requirements now prohibit many of these deceptive privacy-invading activities. When an agency is in charge of and conducting its own evaluations these methods are still in use.

Inventions: Most that is history, political science or sociology is attention paid to new developments, be these revolutions, Populist movements, the growth of multinationals or, in the administration of justice, "new looks" such as work-furlough programs, open-door jails, juvenile or neighborhood peer trials and sentences, new criminal codes, diversion programs, and the like. These latter are all innovative responses which presume either dissatisfaction with existing methods or the growth of new ideas. There are also new developments on the crime side, be these rising rates of marijuana or cocaine use among the young, transnational fraud, computer thefts, skyjacking, urban guerilla terrorism, and the like. These are ordinarily explained in terms of new criminal opportunities, new legal definitions of what is criminal, changed population characteristics,

changed individual experiences (i.e. stress, family breakdown, etc.) and the like. On either side, the innovations in response to crime itself or to problems in the justice system, or trends in crimes and criminal legislation constitute social signals which can also be evaluation measures. What does the growth of pre-trial diversion in town A mean when that is not occurring next door in town B? Analyses of the events in those two towns may show that town A was feeding far more defendants to trial than the courts could digest and, further, that concerned community groups were willing to take responsibility for sponsoring half-way houses, therapy groups or what-have-you. Now take town C which also has a court system quite unable to process cases but which has a district attorney who simply tells the police to reduce their referrals to him. Assume the police respond by reducing their arrest rate of armed robbers from 10 to 2 a month while reported armed robbery rates are the same. Assume now that community leaders demand change, a study group comes on the scene, and, in lieu of tax money to hire more judges, they recommend diversion on the model of town A. There will be several outcome measures including the rate of "reception centers" for pretrial diversion, a change in police arrest rates, and the reported armed robbery rate. A wise evaluator will also monitor dark number reports of armed robbery, will have trusted observers monitor police robbery investigations, will keep talking to the DA and the judges, and will watch what goes on in the agencies supervising diverted cases. But the innovations will be one measure of change. (The same kind of work can be done at the other end, as for example, the many "total" neighborhood delinquency control programs which sought to reduce juvenile predispositions to crime (Cambridge Somerville, Shaws Chicago are a project, many others). Here one starts with the innovation - let us say a recreation center, home visits and counseling - and measures changes in juvenile crime (for example, by follow-up interviews and police records monitoring.)

The foregoing suggest the range of methods used (but not always useful) in social science research to evaluate either conditions at one point in time, processes ongoing, the impact of interventions designed to produce change, or tests of hypotheses about the relationship of one to another variable of interest in an evaluation, crime and criminal justice activities. It is seen that the breadth of such enterprises can vary dramatically, from a national survey of citizen victimization in dark number crimes or of felony histories among ordinary people to studies of one group or organization, to individual assessments in experimental change-inducing circumstances.

Proposals for Action

This section serves to highlight what has gone before. We believe the following activities will be of value in crime and criminal justice evaluation of impacts on and of the system. The emphasis here is on more general studies rather than highly localized detailed efforts.

A. LEAA should review the expenditure of funds to determine what proportion of the funds spent in support of the criminal justice system are spent on sophisticated research and evaluation. These ratios should be compared to the research expenditures as a function of the support to and cost of other human and social problems, as for example, national defense, health or education. The findings may suggest new allocation priorities.

LEAA should review its own support for evaluation research to determine to what extent that has been guided by a rational strategy for evaluation management. Have policies been consistent, has there been peer review of the work done, have the requirements for knowledge been derived from explorations with the outstanding managers in the criminal justice system? Furthermore, has that management anticipated resistance, monitored its own work to learn when evaluation has not led to needed innovation, and identified the conditions under which evaluation does lead to change?

B. LEAA should review its expenditures and sample representatively from among agencies receiving support. After determining the before-aid level of functioning, especially with respect to the adherence by these agencies to minimal standards for operation, the agencies should then be assessed for their progress to adherence to minimal standards for operation. Whenever possible, their achievement of longer range criminal justice standards should be noted. The task them is to identify the conditions - a profile - which characterize successful versus unsuccessful aid programs. Future activities would then focus on learning how to induce change in agencies characterized by these pessimistic profiles. Support without experimentation should be given only to agencies with profiles indicating that minimal standard achievement can be facilitated by support.

C. With respect to the objectives of the criminal justice system, a study should be undertaken which seeks to array the objectives which exist for the various components in the system. This will include formal goals and those inferred from observations of what is happening in reality. One would expect to identify major groups in the population who agree on objectives, and to learn what characteristics of persons and groups are correlated with support for an objective. One would expect to create a social map showing the knowledge and expectations various groups have for each of the components of the system, how important their expectations are, what dissatisfactions exist. In the study of these expectations, satisfactions and dissatisfactions, one would want to learn the extent to which given positions

are derived from immediate experience with crime or the criminal justice system, and which positions are part of larger personality, social, political perspectives. Any such mapping expedition would anticipate that individuals will hold multiple goals for the system, that these need not be consistent or of equal importance, and that they may or may not be validated by real knowledge of crime and justice. Provision must be made for distinguishing between operating standards to be met, long range expectations, and immediate service and gratifications derived from one or another aspect of the justice system. One would look separately at the police, judges, offenders, victims, attorneys, various citizen strata, etc.

D. With respect to conventional output measures of the system, used to infer performance, as for example, index crimes, arrest and clearance rates, trial delays, recidivism and the like, studies are required in a variety of locales which would identify the major influences on and error sources in these statistics.

E. With respect to justice system records, a crucial need is the analysis of records systems themselves to identify major error sources and the reasons for these. Research would then focus on demonstrating how corrections can be made in major agency records keeping, including provision for the routine monitoring for error.

F. Work should be done to develop methods for finding and elicting reports from victims, both those experiencing dark number crimes and reporting victims so that local justice agencies, perhaps in association with universities or contract research groups, can routinely gather information from victims which will reflect on crime frequency, assessments of the police, court and corrections response and recommendations for both improved justice performance and increased citizen self-protection.

G. Work should be done to develop methods for sampling and eliciting reports from a variety of offenders, including those in various stages of justice processing and those still at large. These methods should be capable of routine use at the community level to assess the dark number, police-court-corrections adequacy, and to contribute to recommendations for improved crime protection and efficiency in the justice system.

H. With regard to the description and definition of organizations, events, people and processes which interrelate with the governmental institutions ordinarily

defined as the criminal justice system and to influence outcome measures, an expanded mapping is in order. This mapping will include internal studies of the process through which people enter the system, that is, the intake side invoking the police and definitions of crime by primary and secondary groups such as families, school, work associates, neighbors, etc. The mapping would include studies of the events and outcomes associated with sideways referrals out of the system through diversion, whether this occurs at the beat or station house level, in the DA's office conference between attorneys and the judge or at trial. As more and more reception centers for diversion are generated, it becomes increasingly important to understand what they are and do, how it is they are used, and what influence they have both on the processes of justice and on suspects or offenders criminal careers. In the same way, increased knowledge of the "gatekeeper" institutions on intake side will not only tell us more about objectives for and functions of the criminal justice system, and the discrepancies among these for various groups, but should point to possible ways for shoring up the social controls within these groups so that fewer referrals (complaints) are initiated which require the response of the justice system.

I. As external influences and operations are identified which prove to be important for the functions of the justice system, it may become apparent that traditionally used performance or outcome measures of the system (crime rates, etc.) are strongly affected by these, e.g. intake and diversion features. Research on these should be undertaken to learn how traditional and to-be-devised performance measures are affected by intervention within these often informal or non-public institutions. Insofar as these interventions (in family, gangs, school, work, etc.) are part of planned experimentation by components of the criminal justice system, as for example, LEAA as a support agency, it may be found that there are new optimal areas for producing impact on and by the justice system. Thus, one foresees that by expanding the description of what is relevant to the justice system and conducting research within these previously excluded sectors, research which aims to show how product measures of the system itself (e.g. crime, arrest, trial, recidivism rates) can be altered, new priorities for and definitions of the justice system may emerge.

J. Past evaluations have produced important findings with regard to the effectiveness of system components. These (for example, that prison rehabilitation does not occur and that whatever the prison experience, recidivism rates are about the same) suggest that experimentation must take place within traditional components

of the system which provides for the evaluation of major innovations. Such evaluations must provide for probing in context to indicate unanticipated side effects, should require studies of the new processes themselves so one knows what internal events are relevant to the outcome measures generated, and should utilize multiple measures of outcome so as to determine varied effects and their intercorrelation.

K. Methodological studies are in order. These go beyond the need for studies of record system error or bias to consider the yield of various statistical tests, as for example, in calculating prevalence rates for given crimes. One example is found in the statistical research designed to produce better estimates not only of current heroin use, but to predict the rate of future uptake of use. Another example occurs in the comparison of six statistics for relative alcoholism prevalence under two conditions. At present there are few if any efforts to show intercorrelations among various operating and outcome measures of the sort which would allow identification of common factors, covariance, maximal predictive power when weights and interaction are taken into account (as in multiple discriminant analysis) and so on. Methodological studies bearing on justice system product measures should not be restricted to comparisons of findings from statistics derived from the same data or assumptions. One needs compare different research methods for their yield. This is being done now by LEAA in its routine comparison of dark number with reported crimes. One might, for example, routinely sample from among a national panel of police detectives to compare their reports on the diligence of their investigations to the yield of the court system defined as the ratio of sentences to incarceration over all other dispositions by each type of offense. When studies of the sources of error on existing outcome measures show consistent bias that is beyond administrative control (viz a recent narcotics arrest study in NYC showed low enforcement diligence in high black crime areas because of officer fears while high narcotics arrest rates were obtained in white junkie and middle class grass-user neighborhoods,) one will want to introduce alternative product measures (e.g. in the NYC case one could compare narcotics use rates produced by a panel of heroin dealers, the Braden autopsy formula, interviews requested use estimates among acquaintances from among all arrested persons (regardless of charge), and a criterion study conducted in high and low police fear neighborhoods. It is only through studies on error and bias source, the comparison of various data-gathering methods, and examination of the yield of various statistical

methods applied to such data that one can expect to increase not only the accuracy of outcome measures, but knowledge of the relationship of system operations on outcomes. One can also test refined methods, (as for example the Sellin and Wolfgang crime seriousness index) against their cruder counterparts (the FBI index) to assess improvements which, if great, can then be subject to demonstration programs designed to make refined measures usable in normal reporting. One can, in addition, expect to expand the regular measures of system performance, for example, adding to the victim, offender and participant routine feedback inquiries, periodic "career progress" interviews with a national panel of offenders followed over many years. Repeater victims might also be seen periodically as panelists over the years. Methodolical work which aims to train observers and develop for them reliable measures for rating agency functions are also much needed. When the analysis is to be of day-to-day functions from which inferences are to be made about objectives and from which one hopes to identify direct influences on outcome measures, the importance of having such reliable quantifiable measures dare not be understated. As methodological queries, all such enterprises ask, what can we learn from this device that we don't learn from others? Is what we learn relevant to assessing performance? Can it be learned efficiently enough to make learning it worthwhile, given our priorities for system achievement and for learning? These questions are the essential ones in methodological studies in evaluation. They can be answered only through a program of research, a type considerably more centered, sophisticated, and more consistently well supported than has been in evidence in LEAA, or elsewhere, to date. A commitment to evaluation is a commitment to genuine research.

The foregoing proposals for action each comprises an evaluation research program area. The studies in each program can be done using conventional methods to yield quantifiable data and, more importantly, knowledge about the justice system's performance.

APPENDIX B

MEMORANDUM ON CRIMINAL JUSTICE AND MEASUREMENT

TO: Robert C. Lind

FROM: G. Robert Blakey GB

DATE: April 17, 1975

You asked for a memorandum on criminal justice and measurement. What follows are first thoughts. They do not repeat what we talked about at your house over the weekend fully or in detail. Hopefully, what follows are further details and ideas.

I. Purpose

What Mr. Justice Cardozo said of law may be said of criminal justice. Each man tends to see it through "his own eyes." Our earliest traditions saw the criminal law simply as a substitute for revenge, an instrument to promote the peace. Religious thought, particularly following the Reformation, tended to identify law and morality. Crime became sin, and punishment found its rationale in expiation. The Enlightenment soon followed the Reformation. Criminal proscriptions then became a means of crime prevention through deterence. Modern man, on the other hand, has turned away from any view bottomed on moral or individual responsibility. Crime is now seen as the product of environment or the aberrations of personality. Reformation is the goal. Its explanation lies not in concepts like original sin or conscious choice, but in poverty, passion, discrimination or disease.

Debate over the purpose of criminal law thus is a debate whose terms have changed over time and which is today still unresolved. It is safe to say, however, at least a couple of things about "purpose." No one save ideologues now contends that there is "a" purpose to the criminal justice system; it is obviously a multi-purpose creature if only because many people expect it to fulfill many purposes. It seems beyond argument, too, that the success of the system is significantly to be measured in terms of men's minds. Has justice been done is a question in fact, but also a question of society's perception of that fact. Does the operation of the system promote respect for the social morality it represents? Deterence, too, obviously is a question of perception -- explicitly so. Although the success of the system in crime control would seem to be factual, on a closer examination it, too, is an issue in men's minds. A system that does not lower crime to a tolerable level is a failure. Consequently, I suppose it might be possible to

ascertain what people want from the system and then ask them to evaluate it. Satisfaction with its operation may well be a measurable quantity -- at least in theory.

I wonder, however, if the costs might not exceed the usefulness of your data. How frequent and representative would your surveys have to be? Once a year? National, statewide, citywide, etc.? For criminal justice planning purposes, what if people expect contradictory things? Suppose you are doing the best that can be done, and it's not good enough? Do you then see your task in "P.R." terms?

What troubles me most here is the abstract or general nature of these so-called objectives, i.e., what people might try to measure. I am inclined to think you ought to set your focus on narrow gauge objectives. Looking at "justice" or "crime control" through deterence or rehabilitation or incapacitation is just too abstract to be useful.

II. Crime

Traditionally, we have classified crimes in terms of individual actions -- murder, rape, robbery and the like. When the task is the measurement of a social problem and society's response to it, this system is not only inadequate, but misleading. Crime must be analyzed in terms of more than legal definitions. For present purposes, it is, therefore, more meaningful to speak of a typology such as "street crime," "white-collar crime," and "organized crime," that is, to put chief emphasis not solely on conduct, but also on the people involved.

Next to economics, i.e., inflation and recession, many Americans today consider crime the number one domestic problem. Rightly or wrongly, they fear crime, and the crime they fear is the crime that affects their personal safety, especially on our public streets. While our primarily rural population of yesterday could view crime as characteristic of the remote and immoral city, our primarily urban and suburban population of today views crime as more directly threatening.

Rather than walk in their neighborhoods at night, many of our people stay behind locked doors. Poor people spend money they cannot affort on a taxi, because they are afraid to use public transportation. Moreover, the fear of many is a fear of the stranger, often the stranger of another race. Fear of crime makes our people want to move from the neighborhood in which they live. It is a fear which seems to be ubiquitous, and it is a fear which is impoverishing the lives of many Americans. P.T.A. meetings and church services are not held at night. Library facilities and other cultural opportunities are underused. We have become an intimidated society. As I noted above, I suppose you might measure this fear, though I question how you would relate it to the system of criminal justice. In

addition, people are reluctant "to get involved" in aiding a victim of crime. Thus crimes go unreported to the police, for people feel--perhaps rightly--that the police cannot do anything. People are not only afraid, but hopeless. This too presents a problem for evaluation, should it be based on "reported crime," for if people lost fear and then gained hope -- two goods -- the number of "crimes" might go up.

Public alarm, however, is largely founded in fact. Fully 60 percent of all major crimes against the person--rapes, robberies, and assaults--occur on the street or in other public places. There are honest disputes about the figures, but our best indicators tell us that the various forms of street crime are increasing, with some exceptions, notably, robbery. Overall, since 1967, reported crime has increased about 55%, while population has increased 5%. Violent crimes are up about 67%. Property crimes are up about 53%. Depending on the offense, unreported crime is probably two to ten times higher than reported crime. Certainly, there is reason for concern about street crime. But how do you measure not people's fear, but what it is based on--crime--when the "dark figure" and its relation to reported crime is largely unknown?

Yet all public fears are not justified. Apart from robbery, most crimes against the person--murders, rapes, and assaults--are committed by family members, friends, or other persons previously known to their victims. So it is not a stranger you must fear. The motivation behind these offenses is largely beyond reach of a threat from the criminal justice system; it is not part of the "crime problem," if by problem you mean something that is "solvable" as opposed to "punishable."

It is true that race is a factor in street crime. The ratio of Blacks among perpetrators is about four times that of whites. However, if other factors relating to socio-economic conditions are correlated, there seems to be no interrelation between race and crime. Moreover, street crime is not significantly interracial. Nonwhites are also disproportionally victimized by street crimes--with the exception of larceny. A Black male in Chicago, for example, runs a victimization risk six times that of his white counterpart; a Black female, eight times.

The statistics, such as they are, tell us something else. While crime is on the rise in suburban areas, it is still largely a problem of the city ghetto. Both the criminal and his victim tend to be of a broken home, poor, uneducated, unemployed, and a member of a minority group. Street crime is, therefore, plainly related to socioeconomic factors. If so, the criminal justice system cannot do much about these factors, so it cannot do much about this kind of crime. Measuring these kinds of crimes, therefore, is not measuring the success of the system. And it is on these offenses that our UCR statistics are largely based.

Fortunately, the statistics also tell us something more. For the first time in seventeen years, crime decreased in 1973, although it has begun to rise again recently. The decrease was only 2%, but that decline broke a seemingly irreversible trend. The large core cities having a population in excess of 250,000, moreover, recorded an 8% decrease. Some people have suggested that this drop was not real and that it was more than just a coincidence that mayoral elections took place in these cities. This too is an issue. Maybe the measuring stick is politically biased. What can science do about that? People's perception of their safety, however, brightened. In 1970, 62% of our people felt crime increasing in their neighborhoods; today only 48% believe their neighborhoods are becoming less safe.

Paradoxically, this decline in crime has not been accompanied by an increase in public confidence in law enforcement. In 1970, local law enforcement was rated "good to excellent" by 64%; today, the figure has fallen to 58%. Confidence in federal enforcement--perhaps in the wake of Watergate--has dropped from 60% to 42%. This, too, presents an issue. Suppose people's perceptions are unduly influenced by the sensational? Should we let that influence our measures?

Nor has the criminal justice system itself faired better. In 1967, only 26% were willing to say that the system really "does discourage crime much." Today, the figure has dropped to 18%, hardly a vote of confidence in the system as a whole.

While street crime continues to occupy our attention, too little is said of white collar crime--fraud, embezzlement, tax evasion, price-rigging, double dealing in securities and the like. If we want to measure "real problems," these will be missed if we measure perceptions, since these "crimes" are not perceived as "crimes." In 1949, Sutherland published his seminal study, White Collar Crime, an analysis of the crimes committed by 70 of the largest manufacturing, mining, and mercantile corporations in this country. Over an individual "life career" of 45 years, the 70 corporations had an average of four criminal convictions each. Today, however, no one systematically collects statistics on these crimes; attention is focused elsewhere. People do not find white collar crimes in the headlines each day, so it is not a matter of interest or concern to them. How do we get at this issue? This is an aspect of the dark figure not generally noted.

Many liberals also do not seem to speak of white collar crime because it does not fit neatly into their ideology. How can it be "crime" if it is not the product of ignorance, poverty, discrimination, or disease? Many conservatives do not speak of it either; they are embarrassed for they might have to attack members of their own socio-economic class. Its impact on our nation, however, can be great. The final damage figures cannot yet be calculated, but the recent Equity Funding Corporation indictments, for example, include $430 million in bogus insurance policies. Economically, petty theft merely transfers wealth, usually from the rich to the poor. Yet price-fixing and the

collusive allocation of markets undermine what is left of our free enterprise system, most often transferring wealth from the poor to the rich and unfairly eliminating small businesses. When an individual in the upper middle class evades taxes, less revenue is available for needed social services, and persons whose wealth is more visible must make up the difference. There is a cause for concern here, too, but plainly more knowledge is needed before a more accurate evaluation (or specific action) could be productive. But I don't know how to suggest that you measure the victimization of people who do not know they have been "had" except in a generalized sort of way.

While people are concerned about street crime and white collar crime is largely ignored, our attitude toward organized crime is strangely ambivalent. A majority of our people probably do not believe that a group like the Mafia even exists. Others, particularly liberals, feel organized crime only "services" our moral failings. Touched by its corruption, many political leaders also minimize its significance. Yet a growing number of people see it as a threat. The hard core of organized crime today consists of 24 highly structured groups operating in our largest cities across the nation, but concentrated mainly in the Midwest and Northeast. Their internal organization is patterned after the Mafia groups of Sicily. They are, however, more than mere criminal cartels. They are also para-governments within our society. They are active in professional gambling, the importation and distribution of narcotics, and loan sharking, each an offense parasitic, corruptive, and predatory in character, and whose chief impact is on the urban poor. The economic price tag of organized crime, moreover, was put in 1967 at twice that of all other crime combined, and there is little reason today to revise that figure. In addition, organized crime clearly affects street crime. Estimates, for example, place the percentage of theft related to the need to acquire funds for narcotics at 50 percent in our large cities.

Significantly, too, the organized crime groups have not confined their activities to traditional criminal endeavors, but have increasingly undertaken to subvert legitimate businesses and unions. The viability of large spheres of our economic life is threatened. More importantly, these criminals have everywhere established corrupt alliances with the processes of our democratic society: the police, the prosecutors, the courts, and the legislatures. Freedom from legal accountability is secured, often under a rhetoric of liberty. In many ways organized crime is thus the most sinister kind of crime in America. In a real sense, it is dedicated not only to subverting American institutions, but our decency and integrity. The cause here for concern is real, but how would you put a quantified label on the price tag? Burglary, yes; bribery? The crime commission's figure was largely based on an estimate of the cash flow of gambling. It really did not get at the "cost" of organized crime.

III. Criminal Justice System

To understand the administration of criminal justice in our society today, we must first appreciate the problems of the administration of justice in a largely stable, homogeneous, pioneer, agricultural community of the first half of the 19th century. It is then necessary to understand the problems of the administration of justice in our mobile, heterogeneous urban industrial community and the resulting difficulties involved in meeting those problems with legal doctrines and social institutions first inherited from England and then adapted to an American society of the last century.

We inherited from England a medieval system of sheriffs, coroners and constables, devised for a rural society and fashioned out of the history of the struggle between the Crown and Parliament for political and religious freedom in the 17th century. A professional police force was then unknown. Not until 1844, in fact, was a unified night and day police force established in this country, first in New York City. Its primary function was street patrol. Today, there are more than 450,000 people working for approximately 40,000 separate police agencies that spend more than 3.9 billion dollars a year. Nevertheless, nearly one-half of the personnel of every city department remains primarily involved in street duty. But most of their time is spent in doing social service work. Little of it is true crime-fighting. The rest back up the patrolman in detective work and staff positions. Police work, in short, has changed comparatively little since 1844.

Popular fiction notwithstanding, scientific crime detection is still a limited tool in police work. The radio and the automobile have had a greater impact on day-to-day law enforcement than the best the modern crime laboratory has had to offer. Police work today is still largely looking, questioning and listening--under the best of conditions and in the best of departments.

The performance of the local police, moreover, is hardly impressive. Our criminal justice mortality tables need to be perfected, but they do give us rough ideas--order of magnitude measurements--of what is going on. Putting aside the question of community relations--too many times non-existent in the ghetto--the clearance rate by arrest of reported crimes against the person in 1972, for example, was fortunately high--82% for murder, 57% for rape, and 66% for aggravated assault. Yet recall that here most often the perpetrator was known to the victim--and it has not been in this area that the greatest growth of crime has occurred. Under present law enforcement practices, the 1972 statistics demonstrate, in short, that crime does pay. Only 30% of the robberies, 19% of the burglaries, 17% of the auto thefts and 20% of the larcenies were even cleared by arrest. Excluding automobiles, only about 10 percent of all stolen property is ever recovered. The year 1973 was, unfortunately,

a typical year. Indeed, the trend is moving in the wrong direction. Since 1967, the overall clearance rate has dropped 8%. Note, too, that these figures are clearance by arrest only. Something will be said of convictions later. While these statistics say something about clearance by arrest of crimes, is this the proper measure? Suppose we looked at criminal careers eventually interrupted by contact with the system? I would suppose that there are few people who make a career of crime who do not eventually get caught, but I cannot document this. Yet if I am right, the system may be operating tolerably well to catch criminals even if it does not solve crimes. Which is the proper measure?

The truth is that the nation's local police officers, despite some important advances, remain poorly paid, undertrained, and overworked--and arguably ineffective at stopping crime. Professionalization is only just beginning. All too often, politics, if not corruption, taints their work. The experts agree that too few departments are well organized. There is a need for area-wide planning and state-wide coordination, but local law enforcement remains largely fragmented, complicated and frequently overlapping. America remains in too many places too often a nation of obsolete, small police forces. Too few states have attempted to set up state-wide standards. Police personnel has remained stable, while population and crime have increased. How much could be done just with more adequate personnel has been often demonstrated. The needs of local law enforcement, in short, are massive.

As the state's chief executive, the governor of most states must see to it that the laws of the state are enforced. The governor, too, is usually the head of the National Guard, which backs up police agencies in civil disturbances. Traditionally--and today in some 47 states--the state attorney general, however, is the state's "chief law enforcement officer." In this capacity, he generally has supervisory power over law enforcement in the state. State concern with law enforcement, moreover, is widespread. No less than 28 states have programs for police training. Another 31 have identification and laboratory facilities open to the local police. On the other hand, statewide police and highway patrol forces, now found in a majority of states and in all of the populous states, save California, have been a comparatively new development. Principally because local county sheriffs or constables--institutions, with a few notable exceptions, now anachronistic--were unwilling or unable to enforce the law, state police agencies were established, first in Pennsylvania in 1905. Today, they are the fastest growing of the police agencies; they employ more than 54,000 people, and they have been characterized by a high degree of professionalization, largely free of corruption, and not dominated by politics. Usually, it is their job to train local officers, maintain state-wide laboratories, keep state-wide intelligence and other files, and otherwise back up local forces. In many ways, state law enforcement is the hope of the future.

Unlike the states, the federal government has no common law jurisdiction in the area of criminal justice. Like Topsy, the federal police agencies have "just growed." In 1789, the Revenue Cutler Service was started. Since then, innumerable different agencies have been established. Enforcing 2800 federal statutes, they employ 36,000 full-time men and directly spend some $500 million a year on crime. Located in various departments of the federal government, the chief agencies are, among others, the Secret Service, the Bureau of Narcotics, the Internal Revenue Service, and the Federal Bureau of Investigation. Although small in numbers--New York City has almost as many police officers as the whole federal establishment--the impact of the federal agencies on criminal justice (federal, state and local) has been great.

The impact in recent years, moreover, has dramatically taken the form of federal aid to local and state law enforcement. In 1967, the President's Crime Commission recommended a "national strategy" to fight crime, which called for planning by state and local crime prevention and control agencies to be followed by substantial federal contributions to states and cities for improved law enforcement. Congress responded to the Crime Commission's call with the passage of the federal aid program of the Safe Streets Act of 1968, subsequently extended for three years in 1970 and extended again for two years in August 1973. To date, this program has put more than 4.2 billion dollars into our criminal justice system. Candor requires the comment, however, that as large as this figure is, federal funds in 1972 accounted for only 6% of the total criminal justice system expenditures. Obviously, the major role--financial and otherwise--is still state and local. How then could you measure the impact of the federal funds against national crime rates?

While the attention of state and local agencies has been primarily directed at street crime, a major share of the burden of responding to white collar and organized crime has fallen to the federal government. Evaluation of the federal effort, however, is difficult, because there are so few objective measures. But what information we do have is encouraging. Eighteen Federal Strike Forces are in operation in major cities all over the United States. Progress is being made against the hard core of organized crime. In April of 1973, the Attorney General was able to testify that one-half of the leadership of the core groups of organized crime is now under indictment. The growth of the federal effort has indeed been dramatic. In 1968, 38 indictments involving 38 defendants were returned and 23 convictions were obtained, while in 1972, these figures rose to 76, 121 and 60, respectively. With the enactment of the Organized Crime Control Act of 1970, the federal agencies acquired the legal tools they needed. The failure to do more has been attributable, basically, to a failure to secure legally admissible evidence. Organized crime has shown a willingness to threaten, bribe, or murder those who would testify against its members.

The need for such modern evidence gathering techniques as court-ordered electronic surveillance, moreover, is clearly being established. Indeed, Title III of the 1968 Safe Streets Act has yielded, including state and federal data, 6,956 arrests and 2,495 convictions to date. Other trials are pending. This is surely an impressive figure, which should put to rest any suggestion that electronic surveillance is "ineffective" in combating organized crime. While the outlook is good, much still remains to be done.

But should arrests and convictions be our measure of effectiveness? Has gambling bee controled, narcotics been eliminated, or fencing ceased? This program has brought the lawless to justice; it has not substantially reduced crime. How should we measure its impact? Has it turned young men away from crime by showing that big men get caught? How do you assess criminal careers that never develop? You can always say, "without us it would have been worse!" Can you ever measure what might have been, but was not?

In addition, arrests or "collars" raise another point. Are all arrests the same? Is the arrest of the street robber the same as the arrest of a Mafia leader? The danger of quantifying police work is that it may result in this unsophisticated sort of analysis.

Another point needs to be noted here. The police component of our criminal justice system is not one, but many. How do we attribute to one the "success" it has achieved. It could also be related to what others have done. The crook does not care who catches him, and as the power of state agencies has grown and the jurisdiction of federal agencies expanded, it may be realistically said that each can act in the other's traditional spheres of activity. Measurement here of effectiveness may well be beyond our present means to factor out the contributing efforts of various aspects.

It is not possible to talk about criminal justice without talking about the courts, both state and federal. Popular fiction makes much of the drama of the criminal trial. It is supposed to be, of course, a contest between the forces of good and evil from which the truth emerges--established in a process characterized by a high regard for individual rights. Attention also is focused on the work of our appellate courts, particularly the Supreme Court. There we are told rights are vindicated and justice done. The facts bear little relation to popular fiction. Data from New York City shows how resources are being overwhelmed by volume now. Manhattan has about 120,000 felonies. They result in about 30,000 felony arrests. The arrests result in about 3,000 felony indictments. The indictments result in about 300 felony trials. How can you test anything in systems that have broken down?

Well over half of all criminal defendants are poor--and poorly represented at trial. Seldom is a defendant judged by a jury of his peers--his guilt is usually established as a result of a bargaining process with the prosecutor in which the merits of the case is not always the chief consideration. One need only note the plea recently taken on in a federal court in Baltimore. Popular fiction notwithstanding, criminal justice today is thus largely administrative not judicial. When cases do go to trial, moreover, the conviction rate for the offense as charged is seldom different from the exoneration rate, a disturbing result from any point of view. And some 20 to 25% of these convictions are reversed on appeal.

In large measure, this result is the product of factors wholly unrelated to guilt or innocence or the protection of real liberty. Our courts today operate with rules of evidence and criminal procedure that were fashioned in an age dominated by the death penalty and fearful of the suppression of religious or political dissent. Law enforcement was consciously debilitated. Insistence upon the common law rights of an Englishman of the latter half of the 18th Century--pressed to the limits of their logic in our formative years by a Puritan and pioneer distrust of all government itself--has now produced a complicated, expensive, and time-consuming process, which has largely broken down. The adjudication of guilt or innocence, in short, is at best a matter of chance. A system designed for the leisurely pace of a rural society is now operating as a mass production scythe of the poor and a means of avoiding legal responsibility for the rich. Volume alone--now more than a staggering 7.5 million cases a year--makes a mockery of justice. However one may feel about so-called "civil liberties"—some are real, while others are specious—it is true that the effectiveness of the system in vindicating these values is at the expense of crime control—at least in the short run. How do you balance off the competing values in both the long and short run in a measured way?

We experienced in the last several decades, too, a revolution in criminal procedure, led by the Supreme Court, which has been both good and bad. The elimination of the violence of the third degree--now largely an accomplished fact--was a reform crying for implementation. In 1936, for example, in Brown v. Mississippi, the Supreme Court overturned a conviction because it was based on a confession obtained by physical torture. No other result was possible. But as the High Court moved on and on to more and more attenuated questions of "fairness," the single-minded pursuit by some jurists of individual rights--defined by an 18th Century ideology, but applied to a 20th Century society--threatened finally to alter the nature of the criminal trial from a test of the defendant's innocence or guilt to an inquiry into the propriety of the policemen's conduct--with possible adverse consequences to us all. For example, in 1957, in Mallory v. United States, a rape conviction was overturned merely to discipline the police because they questioned the defendant at the wrong time, even though his confession was fully voluntary and trust worthy. The defendant, after his release, raped another woman. Was that a success for the system?

The revolution in criminal procedure that reached its peak in the 1960's, however, now seems to have come to an end. The liberal block no longer dominates the court. President Nixon's unprecedented four appointments, voting together roughly 70% of the time, have managed to return balance to the work of court in the criminal justice area. In the last year of the Warren Court, the Court heard 26 appeals in criminal cases. The prosecution won only 8 or 31%; 16 or 63% were reversed. In contrast, during the first term of the Burger Court (1969-70), the Court heard and decided 29 appeals in criminal cases. This time, the prosecution won 18 or 62%. Percentagewise, this was a 100% turn around. The "search and reverse" policy of the Warren Court is apparently gone, and this new attitude toward law enforcement has been widely perceived and applauded. Indeed in 1973, the court received a "good to excellent" rating from 37% of our people--the first rise in approval since 1969, a year that marked the nadir of a long period of decline in expressed public confidence in the Court. But should we evaluate the work of the court—a component of the system—by popular surveys? Suppose constitutional values are long run minority values?

While we all have sympathy with the goals of the Warren Court--accurate fact finding, the recognition of human dignity, the preservation of privacy--it may be rightly questioned whether the judicial process is the proper instrument for the moderization of criminal justice. The tool of reversal, the only one generally available, is much too blunt. Among other things, it maybe argued, better training, higher pay, in short, fair minded professionalism would do more for these goals than judicial exhortations. The improvements necessary cannot be

implemented by the judiciary. They must come from legislative and executive action. It is in this direction that the hope of the future lies, and it is to these sorts of remedies that I hope we will now turn to carry forward needed reforms. This, too, poses a problem. We are not measuring a static phenomenon. How do you measure a moving line?

If it is necessary to talk about courts, it is necessary, too, to talk about corrections--probation, institutionalization, and parole. Our criminal justice system, in short, should be viewed as an integrated whole--even if it is not in practice. On any given day--nationwide--our correctional institutions are responsible for a staggering 1.5 million offenders. Each year they spend some $1.5 billion. Here, too, the central role of the states is evident. All states have prison and parole systems. There are 45 states that operate or underwrite adult court and probation systems. Yet corrections does not correct. Life in many institutions is at best barren, at worst unspeakably brutal and degrading. Most institutions are confronted with serious problems of rackets, violence and homosexuality. Treatment is aimed at the offender, while many of the causes of his crime maybe in his environment, which is left untouched. Probation or parole is often a joke. Caseloads of the probation or parole officer are out of proportion. Although over half of the adult offenders are "supervised," frequently this supervision is limited to a 10 or 15 minute interview once or twice a month with an officer carrying over 100 other cases. Trained officers are an exception. Parole decisions are often related to little more than institutional capacity--the necessity to release or parole as many as are imprisoned. Personnel resources and adequate facilities are everywhere in short supply.

We speak of rehabilitation, yet 80 percent of those in corrections are involved in custody and maintenance. Twenty-five of our major state institutions are over 100 years old. Our recidivism statistics, which are inadequate because they depend on catching an offender an additional time, indicate a measure of our failure. For example, in a selected F.B.I. study conducted in 1972, 65% of those who were arrested had been arrested before two or more times. On the average, these individuals had been arrested four times, and their criminal careers spanned four years and eleven months. These repeat offenders tended to repeat for burglary, robbery, narcotics, and fraudulent checks. Frequently, there was a progression towards violence. This data, although incomplete, sadly document the existence of a persistent or hard core offender, who contributes substantially to the overall crime problem.

No short survey can do more than touch on the high points. There are, however, certain conclusions that can

be drawn from a survey of criminal justice today. The remarkable thing about our crime problem, however it is seen, is not that it is bad, but that it is not worse. The problems are everywhere massive. For too long now we have dissipated our energies in ideological debates over issues often of symbolic value only--consensual homosexuality, capital punishment, police review boards, coddling of criminals, and the like. The problems of crime, of course, involve more than criminal justice. Long term solutions must be sought to underlying problems. But symptoms must be treated, too. Every part of the criminal justice system remains undernourished. And the balance between parts is all out of shape. Cops get too little, but compared to corrections officers, they are rich. Balance is almost as important as absolute figures. There is insufficient manpower, and what there is, is not well paid, trained or organized. Our legal theories were developed in another age to deal with other problems. We have tended, moreover, to forget that the liberty of the King's subjects presupposes the establishment of the King's peace. Virtually every aspect of the present system must be rethought--and the rethinking must include questions touching on administration as well as theory.

IV. Conclusion

All of this may seem unduly pessimistic. Our failure to provide adequate resources to the system may mean that it is achieving none of our expressed goals, to say nothing of our real goals. The system is in a state of collapse, and people are beginning to learn the truth which means that things will get worse. Indeed, this may be a dilemma for those who would measure it. Suppose it operates on men's minds? If it is perceived as effective—rightly or wrongly—it will be effective. By telling the world how bad things are, we will change people's perspectives and make things worse!

My basic suggestion, as I outlined to you last weekend, it to forget about effectiveness or impact except in the most limited fashion. We should examine the efficiency of the system itself. First, the data is at least possible to obtain without overcommitting resources. Measure how well the system processes cases—offenders, witnesses, etc. This is a classic problem of imput and output. It says nothing about impact. Develop standards for relative resources—police vs. courts—courts vs. corrections etc. Develop federal vs. state analysis.

These seem to me to be possible goals. Society is not going to give us what we "need" anytime soon. Other things will always take priority — schools, welfare, national defense. The immediate problem is how to balance

our resources between the various aspects of the system so that what we do does not work at cross purposes.

This would offer a concrete enough goal. More is too ambitious now. Consequently, it is pie in the sky. Efficiency, yes; effectiveness, no.

U.S. DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
WASHINGTON, D.C. 20531

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF JUSTICE
JUS-436

SPECIAL FOURTH-CLASS RATE
BOOK

END