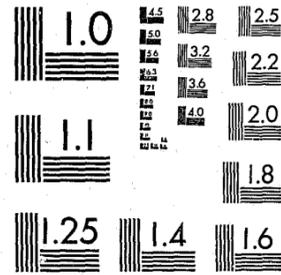


National Criminal Justice Reference Service



This microfiche was produced from documents received for inclusion in the NCJRS data base. Since NCJRS cannot exercise control over the physical condition of the documents submitted, the individual frame quality will vary. The resolution chart on this frame may be used to evaluate the document quality.



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Microfilming procedures used to create this fiche comply with the standards set forth in 41CFR 101-11.504.

Points of view or opinions stated in this document are those of the author(s) and do not represent the official position or policies of the U. S. Department of Justice.

National Institute of Justice
United States Department of Justice
Washington, D. C. 20531

DATE FILMED

5/28/81

X U.S.

Department of the Treasury
Bureau of Alcohol Tobacco and Firearms
Office of Criminal Enforcement
Program Development and Planning Division
Planning and Program Development Branch

SURVEY REPORT
ON
LAW ENFORCEMENT
PRODUCTIVITY MEASUREMENT

Prepared by
Carl A. Bond

April 15, 1980

NCJRS
OCT 27 1980
ACQUISITIONS

73030

TABLE OF CONTENTS

	<u>Page</u>
Abstract	i
I. Introduction to the Concept of Law Enforcement Productivity Measurement.	1
II. Survey of Current Law Enforcement Productivity Measurement Techniques	37
III. Analysis of Problems Found in Utilizing the Current Techniques	155
IV. Recommendations for Improving Law Enforcement Productivity Measurement	229
V. Conclusion	334
Bibliography	351

-i-

Abstract

The Bureau of Alcohol, Tobacco and Firearms (ATF), in evaluating program impact and effectiveness, arrived at the question, is there a practical way to improve productivity measurement for a criminal justice agency or program? This paper, a survey of contemporary criminal justice literature, attempts to answer that question in a meaningful way.

Several interesting evaluation and productivity measurement methods are reviewed. These include the application, to recidivism, of failure rate analysis, normally applied to machine breakdowns. Incapacitation and deterrence effect analysis methods are reported, as are works on violent behavior prediction, system modeling, patrol allocation modeling, inventory modeling, crime seriousness indexing, using the Delphi method of opinion convergence, and the "service package" concept.

What is an arrest? What is recidivism? How many crimes were committed today? These definitional and data problems are but a small example of the difficulties encountered in attempting evaluation and measurement in criminal justice, using contemporary methods.

But evaluating criminal justice agencies and programs, given the current definitional and data problems, can be improved. Implementation of the two following recommendations would greatly increase measurement capability.

1. Descriptive and quantitative modeling of the Criminal Justice System (CJS) found in the United States, as a total system concept defining interagency and environmental reactions and relationships.

For example, let's be able to answer the question, What will be the effect on police, courts, and corrections of increasing prosecutor manpower by 50 percent?

2. Weighting of crime seriousness to establish the relative value of criminal justice activities applied to committed or prevented crimes. Agency or program effectiveness would then be measured as a summation of efforts applied to activities relating to the relative values of the respective crime types.

The mere fact that attempts are made to measure CJS productivity implies that its goals, objectives, and activities have some value or utility. But how much value? The value is related to the seriousness of the committed or prevented crime.

Priorities reflect relative value, but the lack of an overall CJS structure or chain of command has precipitated priorities enumerated in light of each agency's self-serving goals, objectives, and activities. These may not be the optimum priorities for the system and society as a whole.

The research revealed significant interest in evaluation. Several interesting attempts have been made to improve measurement of criminal

justice and law enforcement productivity, primarily focusing on measurement methods. Traditional criminal justice evaluation, when it has been done at all, has usually been in terms of efficiency, but productivity measurement must also include measures of effectiveness. Effectiveness measures deal with external impact or the degree to which desired results are reached; efficiency measures deal with the degree of skill and economy displayed in producing these results.

Is it possible, in practical terms, to measure the productivity of a law enforcement agency or program? Not without innovation and effort.

Recognition is given to the various authors' works by use of end notes. The end notes consist of two parts - the first is the numerical bibliography citation; the second is the page number(s). For example, the citation 1:2 means that the material is taken from page 2 of the source listed as number 1 in the bibliography. The citation 1: refers to the entire work number 1.

The Report of the Advisory Group on Productivity in Law Enforcement

in 1973 stated, "the need for raising productivity has been identified as a national concern," a problem reiterated in President Carter's speech of July 15, 1979. The Advisory Group suggested national support in developing productivity improvement capabilities, and said, ^{54:v-1}

"Governments are challenged to provide more effective police services at a time when the growing desire for public safety is surpassed only by the increase in police costs."

In 1970, Ramsey Clark wrote, "Change is the dominant factor of our time... Today change is the main cause of crime and offers the best opportunity for its prevention. Change is created principally by two interrelated dynamics. Both are uncontrolled human conduct; population increase and ^{16:23} the application of science through technology."

One thing is clear after massive expenditures on crime control --- there is a need for careful reappraisal of the allocation of resources among and within the criminal justice system (CJS) 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 comparison among alternative programs... Unfortunately, in the case of almost every single aspect of the criminal justice system no such capability exists." "Decisions concerning the allocation of CJS resources are made mostly on the basis of past experience, intuition, ^{4:2-1} and wishful thinking."

Bottoms and Nilsson point out that, "no police agency in this country can look forward to meeting increased demands with proportionate increases in manpower" and that, "management sciences...can show... how to capture the manpower leverage offered by science and technology at acceptable cost." Optimal allocation of tax dollars among competing programs or agencies requires a knowledge of their performance. "Implicit in this optimal allocation is the determination of an optimal level of effort (expenditure) for each program."

According to Blumstein and Larson, "one central problem in improving law enforcement is the need to examine the total criminal justice system, comprising police, prosecution, courts, and corrections agencies, in an integrated way." CJS "has remained remarkably unchanged through the significant social, technological, and managerial changes of recent decades...partly from the insularity of these institutions and their relative freedom from external examination and influence" and also partly due to the "independence of the individual components of the system, each of which operates within prescribed rules to attain its own suboptimized objective."

Maltz quotes the "Iron Law of Political Dispersal," which states that, "in any democracy, there is a strong political pressure to expand every expenditure program to encompass a large number of geographic areas, and to spread the resources in the program across many of those areas, in order to build up a broad political base in support of the program."

Patrick V. Murphy has criticized the lack of research and analysis for planning and management in CJS, writing, "The normal situation is likely to be that chiefs do things because they always have or because they 'know' they are right without analysis." He points out the need for performance measurement, saying, "even if the law enforcement and criminal justice elements all used the most modern analytical management techniques to produce an ideal, perfectly coordinated system, the public and the political jurisdiction for which each element of the system works would have no basis for a reasonable expectation of what the system as a whole, or any of its elements, should be able to accomplish," and, "the police and the criminal justice system remain... unaccountable...and...unmeasured." The National Commission on Productivity

pointed out that police protection is costly and that traditionally, to improve protection against crime, more money has been spent on police, but by increasing police productivity, better services can be provided without proportionate cost increases. "The big problem, of course, is knowing whether the police are really doing the job." Many factors affect the crime rate, such as:

1. Police effectiveness
2. Proportion of low-income families
3. Mean age of population
4. Number unemployed
5. Population density
6. Effectiveness of courts and corrections

According to former New York Police Commissioner Patrick Murphy, "For various reasons the police have not been held accountable to any large degree for what they do," but "from now on public officials, reflecting public opinion, will become much more closely concerned with police expenditures, as other public agencies, will be required to institute new and more exact methods for determining what is done and how it can be done better without increasing costs. The pressures are mounting for public agencies to adopt the twin concepts of productivity measurement and productivity improvement which have been used for many years in private industry." Murphy says, "It is clearly the duty of elected officials to make a serious effort to determine if the money spent on policing is being used wisely and efficiently," and that, "any discussion of police accountability cannot ignore the fact that police departments are just one part of the larger criminal justice system that includes prosecutors, the courts, and the penal and correctional institutions." He quotes a study by New York's State Commission of Investigation which found that the so-called system of criminal justice in New York City was, "a jumble of ill-coordinated and inefficient agencies, each pushing its own budget and interests with no regard for any overall plan." Two common ways of dealing with management problems - reducing service or increasing expenditures - often decrease, not increase, productivity. Reducing services is reducing effectiveness. Increasing expenditures can signal decreasing efficiency. The manager who says that he needs more resources to be more productive may not really understand what productivity is. Because of the astronomical rise in the cost of public services, attempts "will be made to identify more

precisely than ever before what government agencies are trying to accomplish and what the true results of their activities are." ^{61:4-42}

According to Deutsch, "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." He terms the measurement process as a set of procedures and flows of information that describe the interaction between evaluation, the organization's environment, and the measurements and measurement strategies used. Deutsch points out that, "it is reasonable to expect government agencies to attempt to provide maximum services at minimum cost," and is troubled that, "very little consideration has been given to development of measurement strategies that do not require all possible data," since the theory of sampling is so well developed and applied. Deutsch also notes that performance measures often turn into quotas, and some evaluative programs have been so poorly designed that they ^{20:32-33} are considered as obstacles by the CJS.

The known relationships between resources of the CJS and its true outputs can best be described as tenuous, making a study and analysis program worthwhile to undertake, since elected and appointed administrators are entrusted by the public with monitoring the efficient operation of CJS agencies, and the public would benefit from the availability of regularly published and readily comprehensible information on the costs and ^{4:2-3--2-5} performance indicators of the CJS.

The management sciences, operations research, systems analysis, and computer sciences can provide valuable assistance to the police administrator who is subject to the constraints imposed by internal and external policies in planning problems, procedural problems, and problems of crime control. 12:22

Bottoms and Nilsson point out some benefits of applying the management sciences to the CJS: 12:22

1. Improved allocation control
2. Identification and evaluation of alternatives
3. Improved services
4. Improved effectiveness
5. Improved community relations arising from apparent increased professionalism

This leads to the concept of productivity, which according to Wolfle and Heaphy, "has something to do with getting work done," and "refers to the relationship between the resources used and the results produced," whereas, "effectiveness generally refers to achieving certain defined results or outcomes without regard to the cost of achieving them."

"Efficiency, on the other hand, refers to achieving any given result with the minimum expenditure of effort required to achieve that result."

Productivity, then, "is a combination of the effectiveness and efficiency concepts," asking "whether a desired result was achieved (the effectiveness question) and what resources were consumed to achieve it (the efficiency question)." 61:2-3

Increasing productivity in police services might be considered in four ways: 54:2-3

1. Improving police practices to the best level known, for better performance without proportionate cost increase.
2. Allocating resources to activities which give the highest return for each additional dollar spent.
3. Increasing the probability that a given objective will be met (e.g., assigning patrols at the time of highest crime incidence).
4. making the most of the talents of police personnel.

Put another way, the concepts of productivity and effectiveness are closely related. "Effectiveness is a measure of the extent to which a goal is achieved," whereas "productivity includes not just what was

accomplished but what resources were required to accomplish it," and "better productivity assessment...is an important step in the process of productivity improvement." Productivity improvement should be an ongoing, long-term process and an integral part of police management. 54:34

From still another source, Hatry, "productivity is generally defined as the amount of output obtained for a given amount of input," and "productivity improvement without productivity measurement is not possible." 61:86

Hatry also identifies a number of uses for productivity measurement: 61:87

1. Indicate the existence of particular problems.
2. Indicate the progress or lack of progress in improving productivity.
3. Identify areas in particular need of attention.
4. Serve as a basis for evaluating specific activities.

- 5. Provide agencies with the information necessary to set productivity targets.
- 6. Establish incentives for management and employees.
- 7. Improve specific aspects of productivity.
- 8. Account for government operations to the public.

The "first step to improved measurement is to understand how the various functions of police work relate to the broader mission of the department and the goals of...government." Although overall police performance may be judged by the general public on the basis of crime prevention or some perceived level of public security, the police are also responsible for noncrime-related and nonemergency services, such as:

- 1. Relations with other CJS members
- 2. Public order
- 3. Emergency response
- 4. Community relations
- 5. Nonemergency general services

When instituting a productivity measurement and improvement program, objectives should be realistic and quantifiable. "Early failures can seriously hinder continuation of a productivity program" and "a specific period of time should be allotted for achieving each objective." 61:138

20:34

The measurement process 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.
- 6. Improvement of the measure and/or measurement strategy.

The Advisory Group suggested a five-step process of productivity improvement: 54:4-5

- 1. Establishment of Objectives
- 2. Systematic Assessment of Progress
- 3. Search for Improved Operating Methods
- 4. Experimentation
- 5. Implementation

There are a number of indicators related to performance, efficiency, and productivity in the CJS, some of which are surrogate indicators.

Larson, et al., suggest that the correct approach is to select judiciously only a small number of indicators...making sure that each of them is representative of a large family of other possible indicators and then deal with only the small sample. Many studies have found indicators which apparently correlate with the crime rate, but no indicator has been shown to explain all of the variation in the crime rate and no set of indicators has been proven to be causally related to the crime rate. 4:2-23

Blumstein states, "The field of crime-control policy has traditionally been characterized much more by ideology and rhetoric than by reasoned inquiry and analysis," but recently we have begun to see, "some movement towards the use of analytical evidence." Blumstein also cites the concept of "general deterrence, that is, the punishment of some

individuals so an example would be provided to others"...a "symbolic effect." The principal concept underlying general deterrence is that of rational man, one who measures the expected risks and the expected benefits of a crime, then carefully chooses the alternative with the greater utility. But are all criminals rational? "Claims and counter claims are made as to the real extent and the possibility of controlling crime. The probability of a convicted criminal actually spending time in prison has been drastically reduced and this has profoundly altered the system." The rate of crime is somewhat dependent on the frequency and severity of sentencing due to these effects: 68:581

1. Deterrent
2. Rehabilitative
3. Incapacitative

According to Maltz, increasing criminal justice expenditures have precipitated the need for evaluation guidelines to:

1. Determine whether to continue a program.
2. Determine funding of programs.
3. Determine whether to expand a program.

Since no crime control program is effective against all types of crimes or criminals, it should be evaluated: 46:1

1. Before being expanded
2. Relative to other available programs

According to Glaser, two alternatives are available to society in dealing with the high social costs of crime:

1. Modify the social economic conditions that provide a strong incentive to criminal behavior.
2. Influence potential criminal behavior through the deterrent forces of crime control.

Glaser says that the "strength of the deterrence effect depends upon the probability of punishment," and that the frequency and duration of violations lead to detection by proactive, rather than reactive, police work. 32:1055

Greenberg delineated three effects of imprisonment:

1. Rehabilitative
2. Deterrent
3. Incapacitative
 - a. Selective
 - b. Collective

But is rehabilitation a reality? How do we measure crimes not committed because the criminal was deterred, incapacitated? The Rand self-reported crime study of California inmates revealed for their sample, "while offense rates decreased markedly over time, the probabilities of arrest, conviction, and incarceration all tended to increase," that the sample, "generally pursued crime opportunistically," and "prefer diversity to specialization." 18:116

Rand defines the functions of incarceration as: 18:118

1. Rehabilitation - intended to modify behavior from unlawful to lawful.

2. Deterrence - intended to alter the offender's perceived balance of the gains and costs of crime so that he desists (with general deterrence aimed at the population at large generating apprehension of risk).
3. Prevention - intended to forestall crime by making its target unattractive and difficult to reach.
4. Incapacitation - intended to remove criminal offenders from the community through incarceration.

There is a "trend developing toward the acceptance and use of the scientific method in acquiring knowledge about the CJS and its component parts," which, "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 the results." ^{4:4-1}

Larson, et al., cite two different types of evaluation in any CJS component:

1. Evaluation of any experimental programs
2. Evaluation of ongoing (routine) day-to-day operations

Thus they seem to ignore the long range planning and evaluation so ^{4:4-1} necessary to manage.

According to Lind, 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 subunit of an institution is performing with respect to some task or to some objective. 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. The process of evaluation poses both a threat and an opportunity to the individuals within the institutions being evaluated. An unfavorable evaluation may lead to disruptive changes, whereas a favorable evaluation may result in more resources. 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. ^{45:3}

The development of measures of performance should be part of an overall evaluation strategy that includes incentives for the adoption of sound evaluation practices. The evaluation procedures and performance measures that one develops must be consistent with the objectives of those who will use them. ^{45:3}

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. Evaluation, and measurement related to evaluation, may meet significant

resistance within the agencies responsible for implementing its use. Just as the principle of cost-effectiveness should guide decisions with regard to criminal justice programs, so it should guide our research strategy. 45:4

The very concept of evaluating the CJS 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. Implicit in evaluation is an objective or set of objectives and rules for determining whether or not one situation is better than another with regard to these objectives. Evaluation presupposes that, given any two situations we can determine which one is preferred given our basic objectives. Further it presupposes the ability to establish an ordinal ranking of alternative situations or that we can rank alternatives on an ordinal scale. 45:4-5

There are several ways that one can approach the problem of evaluation and each has different implications for the role of measurement. One is to define the objective or objectives that one wishes to achieve and then develop a procedure for assigning numbers to alternative situations so that a situation providing for a higher level of objective achievement is assigned a higher number than all other situations corresponding to lower levels of achievement. The objective function incorporates a value system as it pertains to how various objectives should be weighed. Given this approach, 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 which maximized the value of the objective function. This provides the evaluation and the problem of choices becomes that of finding the alternative that maximizes the objective function. 45:5

There are many objectives of the criminal law and of the system of criminal justice institutions that has developed to enforce it. Certain types of information are relevant to any evaluation. While there may be a number of goals or objectives of the CJS and while different men may subscribe to different goals, most people who think about the system behave as if crime control were one of the goals of our CJS. 45:6-7

For the CJS and its components, one can, through a process of questioning and analysis, identify higher level goals and develop procedures for measuring the achievement of these goals. Limitations in the use of 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. We cannot separate the effects of changes in the system from the effects of changes in other factors. The public's evaluation of the system is important just as is the level of crime and must be considered. 45:12-13

Two of the primary reasons for measuring performance and evaluating programs on the basis of performance are:

1. To enable us to better allocate our resources
2. To better manage the CJS

Decisions with regard to allocation and management require measurement and evaluation. The question of resource allocation within the justice producing system can be analyzed in terms of three separate but mutually interdependent decisions:^{45:13}

1. Decisions about how much of society's resources are to be devoted to the justice system.
2. Decisions about what proportion of the total expenditure will go to each of the production units.
3. Decisions within production units about the allocation of funds among specific production tasks.

The optimal expenditure for 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.

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, and the contribution to justice by any one agency is critically dependent on the operation of the other agencies within the system. 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 CJS as a whole. Each agency has to make decisions with regard to how it spends its funds on men and

material and to which task 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:^{45:13-15}

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?

The manager needs measures of performance that are task-oriented and measures of effectiveness related to basic objectives. To evaluate programs, he needs one 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. Even if we had an acceptable measure of the total performance of the CJS, 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 funding of our criminal justice institutions is fragmented between local, State, and Federal agencies, but it might apply to each level.^{45:16}

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. For the purpose of allocating resources among different branches of the criminal justice system, it is vitally important that we develop measures of workload and of performance for each of the component parts. We should also 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. We need more system modeling in order to effect a better balance between different parts of the CJS. It is important in analyzing allocation decisions ^{45:16-17} to keep the higher level objectives of the system in mind.

The effective use of performance measurement in evaluation is critical to justifying that a program of evaluation which supports the development of performance measures is cost-effective. 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 on measured performance. Crime rates and the level of crime by type are and will remain an important element ^{45:18} in criminal justice evaluation.

Lind describes research to promote the more effective use of evaluation of performance, in which researchers on Criminal justice evaluation are

provided with guidance on making criminal justice evaluation more effective, and how to implement better decision making based on performance measurement within various parts of the CJS. He says there is an apparent significant misallocation of resources among different criminal justice institutions. In some instances the system appears to have broken down because one part cannot handle the work generated by the other parts. We should evaluate the relative effectiveness of interrelated parts of the system, and, based on the appraisal, find ways of reallocating resources, or at least providing supplemental resources, to help those parts of the system that constitute a bottleneck. One way would be to address the technical problems of assessing the situation and the development of models to evaluate the impact of each part of the system as it relates to the workloads of other parts of the system. The institutional network that resulted in the existing funding decisions should be analyzed to see how it might be influenced and how a better allocation of our criminal justice resources might be effected. Crime data are important for evaluating criminal justice institutions and it is worth making a major investment to make crime ^{45:19-21} statistics and crime data a useful tool for evaluation.

Lind states that what Mr. Justice Cardozo said of law may be said of criminal justice, "Each man tends to see it through his own eyes." The CJS is obviously a multipurpose creature, if only because many people expect it to fulfill many purposes. A system that does not lower crime to a tolerable level is a failure. It may be meaningful to speak

of a typology such as "street crime," "white-collar crime," and "organized crime," or to place chief emphasis not solely on conduct, but also on the persons involved.
45:App. B,1-2

Next to economics, i.e., inflation and recession, many Americans today consider crime the number one domestic problem. Public alarm is largely founded in fact. Fully 60 percent of all major crimes against the person occur on the street or in other public places. Lind notes the increase in reported crime and says that depending on the offense, unreported crime is probably two to ten times higher than reported crime. A survey in the wake of Watergate showed that confidence in Federal enforcement has dropped. While street crime continues to occupy our attention, too little is said of white-collar crime, fraud, 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, stating that over an individual "life career" of 45 years, 70 corporations had an average of four criminal convictions each. Many liberals 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 socioeconomic class.
45:App. B,2-4

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. Liberals feel organized crime only "services" our moral failings. Political leaders also minimize its significance, yet a growing number of Americans see it as a threat. The economic price tag of organized crime was put in 1967 at twice that of all other crime combined and there is little reason today to revise that figure. Organized crime clearly affects street crime. Estimates place the percentage of theft related to the need to acquire funds for narcotics at 50 percent in our large cities.
45:App. B,5

Organized crime groups have not confined their activities to traditional criminal endeavors, but have increasingly undertaken to subvert legitimate businesses and unions. In many ways, organized crime is thus the most sinister kind of crime in America, dedicated not only to subverting American institutions, but our decency and integrity.
45:App.B,5

We inherited from England a medieval system of sheriffs, coroners and constables, devised for a rural society. Police work today is still largely looking, questioning and listening under the best of conditions and in the best departments, and includes a great deal of social service work.
45:App.B,6

Excluding automobiles, only about 10 percent of all stolen property is ever recovered. While statistics say something about clearance by arrest of crimes, is that the proper measure?
45:App.B,6-7

Lind says that 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." Although small in numbers, the impact of Federal agencies on criminal justice has been great. Additionally, the impact in recent years has dramatically taken the form of Federal aid to local and State law enforcement. How could we measure the impact of the Federal funds against national crime rates? ^{45:App.B,8}

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 is even more difficult because there are so few objective measures. Should arrests and convictions be our measure of effectiveness? How should we measure impact? Are all arrests the same? A danger of quantifying police work is that it may result in an unsophisticated sort of analysis. ^{45:App.B,8-9}

It is not possible to talk about criminal justice without talking about the courts. Criminal justice today is largely administrative, not judicial, the product of factors wholly unrelated to guilt or innocence or the protection of real liberty. If it is necessary to talk about courts, it is necessary, too, to talk about corrections. Our criminal justice system should be viewed as an integrated whole - even if it is not in practice. Corrections do not correct. Treatment is aimed at the offender, while many of the causes of his crime may be in his environment, which is left untouched. Probation or parole is often a joke. Our recidivism statistics, which are inadequate because they depend on catching an offender an additional time, indicate a measure

of our failure. The remarkable thing about our crime problem, is not that it is bad, but that it is not worse. The problems of crime involve more than criminal justice. Long-term solutions must be sought for underlying problems. Every part of the CJS remains undernourished. And the balance between parts is all out of shape. Virtually every aspect of the present system must be rethought. The rethinking must include questions touching on administration as well as theory. ^{45:App.B,10-13}

The immediate problem is how to balance our resources between the various aspects of the system so that we do not work at cross purposes. To do that, we must be able to measure what is being done. ^{45:App.B,13-14}

"There is a great deal of current literature available on performance measurements for the social services," but "many efforts totally ignore the difference between measures of effectiveness, measurement strategies, and measurement processes, and 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," which is not sufficient for CJS productivity measurement. In the CJS, "optimizing performance at the component level does not always lead to optimal system performance." ^{20:124-126}

"One of the factors that impedes coordination among the various agencies is a high degree of interagency rivalry," "characterized by refusals to exchange information and personal antipathy." ^{61:40} But in trying to optimize CJS performance there could be three types of evaluation measures: ^{26:26}

1. Effectiveness measures used to indicate the degree of success of a project or program in dealing with the target problems. These measures are end-oriented.
2. Efficiency measures used to indicate how well the project or program has been implemented (according to its plan). These measures are means-oriented.
3. Attitudinal measures helpful in interpreting the degree of project success.

Evaluation implies planning and is a seven-step process: ^{26:1-2}

1. Quantifying project/program objectives.
2. Establishing the relationship between objectives and organizational goals.
3. Identifying evaluation measures.
4. Determining data needs.
5. Developing methods of analysis.
6. Monitoring ongoing activities.
7. Performing analysis.

Maltz defines evaluation as "the process of determining the value or amount of success in achieving a predetermined objective. It includes at least the following steps: formulation of the objective, identification of the proper criteria to be used in measuring success, determination and explanation of the degree of success, and recommendation for further program activity." "The primary reason for performing an evaluation is to make the best possible decision." ^{46:4-5}

The problem of CJS evaluation has been addressed by asking how a particular CJS component, such as police, can use expenditure, employment, and resource allocation data to determine its efficiency and effectiveness. One method has been to compare itself with equivalent agencies. This approach ignores problems such as definitional differences in data among jurisdictions. ^{4:1-4}

The "trend toward research evaluation of programs continues ... in the past, persons have argued for this program or that one, mainly on emotional grounds rather than in terms of any conclusive evidence that the program accomplishes any significant alteration of behavior ... as more evidence develops from studies ... it will be possible to declare that program X has been shown to achieve a success rate ^{31:294-295} which could be compared with the success rate of program Y."

Measures can be classified as either being absolute or relative, with relative measures in unit-free ratio. Directness of the measure can vary from direct, which expressly evaluate performance, to indirect, which are separate from but related to the performance, to surrogate measures, which can be shown to be correlated to the performance being evaluated. Performance measures can be typified in at least six ways:

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.

"It is often useful to apply two or more measures together in order to increase the significance of the results obtained."^{20:24-27}

In 1971, Riccio studied the feasibility of building a model for the study of crime control administration, but found it difficult to isolate phenomena to be studied and said, "whether or not hard and fast laws of human behavior even exist is a matter of debate." He saw three types in the criminal population:

1. Rational criminal - weighs the costs and benefits of committing a crime - depends somewhat on economic conditions.
2. Crime or nothing - prefers crime to working - no alternative mode of life.
3. Narcotics addict who must steal to replenish his dope supply.

The three can have different crime rates which can be summed to arrive at a total crime rate.^{64:1-6}

External dynamics can influence data, such as:^{26:33}

1. Changes in policy.
2. Changes in administration.
3. Changes in economic conditions.
4. Developments in other urban programs.
5. Urban development.
6. Changes in the CJS or law.
7. Changes in the environment (e.g., the heroin price).

Two characteristics stand out in the present crime situation:^{3:190}

1. High apparent recidivism rate in such crimes as mugging, burglary, robbery.
2. High rate of unsolved crimes.

Avi-Itzhak and Shinnar cite four factors involved in the increased crime rate (1960-1970):

1. Input rate of new offenders exceeds disposal rate (the average offender age has become lower due to increased juvenile arrests, indicating increasing new offender population).
2. Police and prosecution effectiveness.
3. Sentencing and parole policy.
4. Behavior characteristics of offenders.

But in this changing environment, how do we isolate "police effectiveness"?

Police affect and are affected by other parts of the CJS. For example, effectiveness in preventing crime depends in part on the rehabilitative function of corrections. "A principal objective of the police is to prevent crime. Yet many police departments do not think positively and specifically about crime prevention." "The sum efforts of the police department theoretically are geared toward deterring crime; the very existence of the department serves notice on would-be criminals that society has the means to track down and apprehend offenders."^{54:10-37}

From Williams, "one criterion for judging the effectiveness of the CJS is its ability to reduce crime."^{53:263} This profundity ignores the fact that it requires a valid measurement technique to determine not only the reduction in crime, but even how much crime currently exists.

The PROMIS system, installed in the Washington, D.C., U.S. Attorney's Office, collects case data to aid the prosecutor in his day-to-day operations. "Presently, the system has a seriousness score for the defendant based on his past criminal conduct. The score is intended to identify persons with the most serious criminal backgrounds,"^{53:263} and includes the Sellin-Wolfgang score of each offense, which is a measure of offense seriousness, a good priority or utility indicator.

Two advocates of the incapacitation strategy were Attorney General Saxbe and then-President Ford, saying there is no question that incapacitation works to reduce crime, but if the number of crimes averted is small, the return may not be favorable for the \$10,000 annual imprisonment cost per convict. Shinnar argues, "we should restore the same incapacitation rate we had in 1960, and that we would reduce crime by as much as 50 percent through that incapacitation."^{53:4-5}

When measures of overall departmental effectiveness and ways to measure police work in all its variety are found, front-line officers will respond more positively and the enthusiasm of officers can be rekindled by making each responsible for a broader activity range, but "finding ways to measure the whole gamut of police activities will be much more difficult than determining, say, the number of television sets manufactured. Human interaction, which is what police work is all about, presents problems when it comes to new definition and measurement."^{61:43}

Although Hatry says that defining the product of police work is a major problem, he identifies four police functions:

1. Reducing, preventing, or deterring crime.
2. Helping to maintain a community feeling of security.
3. Apprehending persons responsible for crimes.
4. Carrying out noncrime-related functions, such as answering citizen inquiries

He also points out that many factors, demographic and economic, affect the crime rate, and also that multiple measures may be needed to adequately reveal productivity. Hatry states that "most major police products will represent the joint efforts of many employees, ...but for internal management purposes, periodic examination of the productivity of individuals, if undertaken properly, may be appropriate." In interpreting productivity measurement data, certain comparisons might be made:^{61:88-115}

1. Among time periods
2. Among police districts or groups
3. Among other governments
4. With estimates made by advance analysis

Two misconceptions are that police departments do not need productivity improvement programs, and that productivity measurement and improvement are limited to manufacturing industries.^{61:130} "There is no substitute for detailed and careful consideration of police statistics and operations - analyzed with experienced, professional judgment." "Organizations such as the International Association of Chiefs of Police, the Police Foundation, the Law Enforcement Assistance Administration, the American Society for Public Administration, and the FBI, are interested in the development of productivity concepts, measures, and practices for police work."^{60:3-11}

The decisionmaker needs to develop objectives or a statement of what is to be accomplished, then determine alternatives which offer some chance of attaining the objectives. To examine the reaction of the system to each alternative, a model, which is a simplified representation of the real world, can be utilized. Dean Gould has said, "the principles of scientific management used by industrial engineers can be applied with much success to the problems faced by public administrators."

The term, "police productivity," implies that it is possible to define at least some of the objectives of police work, that some measures of performance can be made, and that there are real tradeoffs between priorities which can be and ought to be made explicit as an aid to sensible decisions. Although the police officer is employed to enforce the law, prevent crime, and arrest criminals, he has many noncrime-related service type functions. Kiernan also noted the difficulty in police productivity evaluation due to the "many other agencies involved in police work." Riccio observes that the high cost of police services and the demand for more and better service causes managers, who are accountable for such expenditures, to need police performance information related to their goals, objectives, and activities, because "in order to improve productivity, the police manager must first be able to measure it." Productivity, concerned with the generation of valuable outputs via the application of inputs, encompasses:

1. Efficiency
2. Effectiveness
3. Quality

Murphy points out that, "only a fraction of the amount being spent to increase productivity in industry is being spent on increasing productivity in policing." He also observes that since police productivity depends greatly on other parts of CJS, the police cannot be expected to solve the crime problem alone, and that balanced spending among the parts is vital. "Some police departments ...are evaluated on the basis of public relations," causing "negative productivity issues." Hansen depicts the majority of law enforcement agencies as facing expanded workload with limited resources through application of advanced management methods and points out the most valuable resource - people, encouraging participation of all personnel in problem solving and system development.

According to Forst, "one of the conspicuous features of the CJS in the United States is its fragmentation." He quotes the National Advisory Commission on Criminal Justice Standards and Goals as stating, "no element of the criminal justice system completely discharges its responsibilities simply by achieving its own objectives. It must also cooperate effectively with the system's other elements....Police agencies have a responsibility to participate fully in the system and cooperate actively with the courts, prosecutors, prison parole boards, and non-criminal elements." "Discretionary decisions within each organization of the criminal justice network are limited by legal mandate and by its internal structure and policies...agents of each organization exercise considerable unauthorized as well as authorized, discretion in making decisions." Discretion exists whenever an organization and its

agents make choices that are not generally open to re-examination by others.
32:679-685

According to Wilson and Boland, there is general agreement that the higher the clearance rate, the lower the total index crime rate; that differences in the risk of arrest are associated with differences in the crime rate; that patrol officers are a direct deterrent, that the arrest ratio is influenced by the level of resources and how the police use them, and that communities respond to increases in the crime rate by hiring more police. But are there valid measures here? If a police chief emphasizes clearance rate, it will be high. What does the offender perceive as his chance of of being caught? How do we measure deterrence? Do police use resources efficiently? And finally, how long will the public keep throwing money into a black box without an accounting of what's been accomplished with the money?
11:368-377

According to Riccio, the "value of a law in either natural or social sciences is embodied in its contribution to an effort to control a monitored activity" and he expresses a lack of confidence in the laws of the social sciences, which perhaps could be more accurately called collective past observations.
64:2
The social cost of crime control includes the direct operating cost of the CJS, which combined represents a major public sector expenditure. The allocation or assignment of these costs causes the following questions:

1. How do we allocate among agencies of the CJS?
2. How do we allocate to crime types?

3. What are the effects of interaction in changes of expenditure on any one part of the CJS?

Belkin, et al., advocate cost analysis based partially on, "costs associated with different crimes," and the derivation of cost-effectiveness measures comparing the relative effectiveness within the CJS.
7:12-14
But would a seriousness index apply better here?

There is some cost attached to the relative lack of coordination among CJS component organizations, because the "lack of a chain of command... ensures that conflicting objectives are designated." These costs could be expressed in economic or noneconomic terms, but how? "There is profound need for the development of the concepts of measurement strategies and measurement processes for CJS applications," and "there must be some effort made to develop models relating effectiveness achieved to overall organizational effectiveness," to remove existing conflicts and improve the overall performance of the CJS.
20:164-165
Focus on quantitative measures in evaluation suggests an ability to quantitatively predict the consequences of alternate programs prior to implementation, which implies the use of models of system behavior, such as Blumstein's JUSSIM which can be used on a CJS-wide level to assess the system consequences of changes in inputs such as resources applied, workload, or recidivism rate.
4:1-3

Each year "LEAA spends hundreds of millions of dollars on programs aimed at improving the components of the criminal justice system," but their staff members "often are not aware of other LEAA programs related to or even nearly equivalent to those in their own bailiwick," and "it is often difficult to obtain a final report of already completed projects."
4:1-1--1-2

Larson, et al., cite these deficiencies as showing a strong need for new mechanisms for appraising or evaluating LEAA programs, since evaluation implies a focus on measurable quantities, system inputs, measures of process, system outputs, and final outcome measures. According to Larson, et al., "LEAA should support the development of a formal CJS-formed evaluation methodology." Also, several quantitative measures, models, and methods exist that, if properly used, would help the CJS agencies to evaluate regarding specific operational problems, and handbooks could be written for this purpose. The fact that quantitative methods which would help police managers in operational decisions receive little use seems to indicate that the methods either have not been communicated or are not understood.
4:1-2--1-7

In its 1976 Research Highlights, LEAA said, "one of the major deficiencies in police administration is the lack of adequate performance measures. The need to develop sound productivity measures is becoming crucial with today's budget constraints. The American Justice Institute (AJI) is developing measurement systems more sophisticated than arrest rates and reported crime figures - both of which have proved inadequate for management purposes." The study was to have identified and ranked the objectives of police work. Then productivity measures were to be developed to correspond to these objectives. The new "performance measurement system" so developed does not offer substantial improvement, because AJI continues to rely on reported crime as a parameter, a number which few people have faith in. AJI even uses the "unreported crime," for which there are not even reliable estimates. Such measures certainly

lack utility and may even lack usefulness. The AJI does, however, wisely call for the use of effectiveness and productivity measurement.

26:24-25

To measure the contribution of projects to program goals:

1. "For programs within the area of prevention and postadjudication, the goals will be related to target groups of offenders or potential offenders for the purpose of decreasing the number of crimes they commit."
2. "For programs within the area of deterrence, detection and apprehension, goal achievement will be related to the number of crimes committed in target areas."

Maltz says it is possible to consider effectiveness from two perspectives:
46:2

1. External measures which relate to program success in countering crime.
2. Internal measures which relate to the manner in which the program achieved its results.

As an example of CJS evaluation, LEAA requires Project Grant Applications to include an evaluation component, and project objectives must be planned to meet goals. Consistency is needed in data collection as "data are the inputs to evaluation and analysis produces the output."
26:5-33

Data constraints are:

1. Existence of the data
2. Availability of the data.
3. Reliability of the data.
4. Cost of collecting the data.

Clark distinguishes several types of crime:

1. White-collar crime.
2. Organized crime.
3. Crime in the streets.
4. Crimes of passion.
5. Violations of regulations for public health, safety and convenience.
6. Revolutionary crime.
7. Corruption in public office.
8. Police crime (wrongful arrest).

and says "to think of controlling street crime while organized crime flourishes is to ignore their clear connections. Narcotics supplied by the professionals nourish thefts, burglaries and sometimes robberies and muggings." It may be possible that each crime type will require a different productivity measurement method. It can also be noted that "as expenditures for protection and deterrence increase, the losses to victims decrease," but for criminal activity to approach 0, then expenditures would have to approach infinity. Public costs for protection and deterrence are fairly well recorded, but private expenditures for protection by individuals and businesses are not well documented.

16:35-37

32:1057

Upon surveying the literature in the criminal justice field applying to productivity or effectiveness measurement, or evaluation of law enforcement effort, several things become clear:

1. There has been work in prediction as it relates to recidivism.
2. There has been work on deterrence and the incapacitative effect of incarceration.
3. There has been work done in modeling the CJS and in statistical techniques applications, along with uses of other techniques from operations research, business, and industrial engineering.
4. Very little work has been done to solidly delineate the dark areas of unreported crime to arrive at a defensible total crime figure.
5. Notwithstanding the Victimization Surveys and minute improvements in available data, too little emphasis has been given to supplying solid, reliable data for the academicians to try their new techniques on. To borrow a phrase usually heard in conjunction with computers, "garbage in - garbage out."

Prediction studies began with attempts to predict parole violation in the United States by Warner in 1923, and Hart in 1923. Attempts at weighting factors were made by Burgess in 1928 and by Vold in 1931. Two major stages of prediction study are:

1. Construction of an experience table.
2. A validation sample is prepared and the results obtained from the first sample are tested on the second.

If the results are acceptable, they are applied to future populations by references to a table of scores. Continual validation is desirable.

Conviction is the result of a sieving process involving detection, prosecution, proof of guilt, etc., and ... someone described as "not reconvicted" could in fact have reoffended without being reconvicted. A simple method of attempting a prediction of the likelihood of reconviction is the point score system. First used in 1928 by Burgess with the records of 3,000 prisoners in Illinois, a man is given one point for the presence of a characteristic "which, in a previous analysis of the sample, was found to be positively correlated with the dependant variable, reconviction." ^{66:25} The total points constitute the score for that individual. If the assumption of linear association with reconviction is correct, men with high scores will tend to be more often reconvicted.

A refinement of that method "that partially accounts for correlations among the variables is the linear discriminant function." ^{66:25} If the variables are normally distributed, an individual's score would be the addition $x_1 y_1 + x_2 y_2 + \dots$ over all the variables, where x_i is the value of the i th variable and y_i is the weight attached to that variable.

According to Williams, one method of trying to reduce the crime rate is to concentrate more resources on attempting to reduce repetitions in criminal behavior, since first offenders at the time of their first case in 1973 accounted for less than their share of the cases prosecuted during the year. ^{81:6-38} Identification of persons most likely to recidivate would allow better utilization of resources and perhaps a reduction

in the overall crime rate. Seven independent factors were hypothesized ^{81:11} by Williams as influencing recidivism:

1. Personal characteristics of the defendant
2. His previous criminal history
3. Actions taken by the police
4. Actions taken by the prosecutor
5. Actions taken by the defense counsel
6. Actions taken by the judge
7. Defendant's experience within the corrections system

Before conducting the analysis of recidivism using multiple regression, some simple frequency distributions were tabulated to understand how the dependent variables were distributed, and how they were related to some of the independent variables hypothesized to be important. The two questions which can be addressed by looking at recidivism ^{81:22} just within 1 year are:

1. What is the best prediction which can be made at the screening of a case as to the frequency and seriousness of a person's future recidivism, based on personal characteristics of the defendant, characteristics of the defendant's criminal history and characteristics of his/her current case?
2. What actions taken by the CJS have an effect on the probability of recidivism?

The question of the best possible prediction of recidivism at the ^{81:23} time of screening was addressed using three measures of recidivism:

1. The number of cases brought by the police in 1973, with the characteristics of the defendant at the time of the first case as the independent variables.
2. Whether or not the defendant recidivated within 6 months, based on his characteristics at the beginning of the 6-month period.
3. Seriousness of the second case within 6 months for those who were rearrested, based on his characteristics at the beginning of the 6-month period.

All of the personal characteristics of the defendant had an effect on the number of arrests in 1973 at the 2 percent confidence level or less, except for two variables -- whether the defendant's last job was held for 6 months and whether the defendant had a physical disability or bad health. Of the six variables describing a person's criminal history, three were significant at the 5 percent level and three were not.

The number of previous arrests for crimes against persons, if the first arrest was for auto theft, and whether or not the defendant has used an alias were not significant. Looking at the variables which describe the case for which a defendant was arrested, significant relationships with recidivism were found for the type of case, its seriousness score, the number of codefendants, and the relationship between the victim and the defendant. The seriousness of the crime had a negative effect on the probability of recidivism, although it may be that the explanation behind the negative relationship for seriousness is that persons who commit serious crimes are likely to be incarcerated and unable to recidivate.

A separate analysis was made to determine whether the variables available at screening could also be used to predict rearrest and seriousness of the second case if the defendant was rearrested. The conclusion is uncomplicated: men who commit serious felonies are likely to commit another serious crime if they do recidivate.

Assignment of a case to the Major Violators Unit in the Washington, D.C. Prosecutor's Office had a positive effect on recidivism for each of the four measures of the frequency of recidivism: rearrest, re prosecution, reconviction, and rearrest within 6 months. One possible explanation is that since the Major Violators Unit is supposed to be targeting on persons who have committed a misdemeanor but have serious criminal histories, such persons would be a group likely to be recidivists.

There was not a single action taken during case processing that had a significant effect on the seriousness of the second case if the defendant was rearrested within 6 months. According to Williams, one of the primary findings of this analysis is that personal characteristics of the defendant and characteristics of his criminal history were more important determinants of recidivism than any of the actions taken by the CJS during case processing. Another was past criminal history seems to be a good predictor of future criminal activity.

Avi-itzhak and Shinnar say that the expected incarceration per crime committed is derived by dividing the total number of crimes committed in a given period into the average number of people in prison. One of the characteristics of the present crime situation is the high rate

of recidivism in certain types of crimes, such as mugging, burglary, and robbery. Another dominant characteristic of the present crime situation is the high rate of unsolved crimes.

If one let the random variable D describe the number of convictions during a complete criminal career, the probability of the offender never being convicted is denoted by P₀.
3:191

$$P_0 = P(D=0) = \sum_{x=0}^{\infty} \left(\frac{\lambda_0}{\lambda_0 + \eta} \right)^x (1 - q_0)^x \frac{\eta}{\lambda_0 + \eta} = \frac{\eta}{\eta + \lambda_0 q_0}$$

The probability of the offender never being convicted again given that he has survived the ith sentence is denoted by P_i.

$$P_i = \sum_{x=0}^{\infty} \left(\frac{\lambda_i}{\eta + \lambda_i} \right) (1 - q_i)^x \frac{\eta}{\lambda_i + \eta} = \frac{\eta}{\eta + \lambda_i q_i}$$

The distribution of D can easily be expressed in terms of P_i and q_i, i = 0, 1, 2, ...

$$P(D > 0) = 1 - P_0$$

and

$$P(D > m) = (1 - P_0) \prod_{i=1}^m (1 - P_i) \Theta_i, \quad m = 1, 2, 3, \dots$$

The expected number of convictions in a criminal career, E(D), is given by

$$E(D) = \sum_{m=0}^{\infty} P(D > m) = (1 - P_0) \left(1 + \sum_{m=1}^{\infty} \prod_{i=1}^m (1 - P_i) \Theta_i \right)$$

The expected number of convictions, given that there was at least one conviction is denoted by ϕ :

$$\phi = E(D | D \geq 1) = \frac{E(D)}{1 - P_0} = 1 + \sum_{m=1}^{\infty} \prod_{i=1}^m (1 - P_i) \Theta_i$$

Greenberg cites a California study, in 1958, initially under the guidance of Leslie Wilkins from England. The Research Division of the California Department of Corrections entered the field of parole outcome prediction
34:544-575

from base expectancies. The base expectancy scale assigns a score to each inmate according to possession or absence of certain historical characteristics. It predicts from past observation the percentage of inmates for each particular BE score who will have favorable outcomes; the higher the score, the greater the possibility of favorable parole

outcome. Favorable outcome was defined as no return to any prison from parole, no jail sentence of 90 or more days, or not PAL (parolee-at-large) over 6 months. The scale scores range from 0 - 76, accumulated for whichever of the following characteristics are applicable, with high scores
34:544

favoring parole:

- 12 - arrest-free period of 5 or more consecutive years
- 9 - no history of any opiate use
- 8 - not more than two jail
- 7 - not committed for burglary, forgery or checks
- 6 - no family criminal record
- 6 - no alcohol involvement
- 5 - not first arrested for auto theft
- 5 - six or more consecutive months for one employer

- 5 - no aliases
- 5 - first imprisonment under this serial number
- 4 - not more than two prior arrests
- 4 - favorable living arrangement

The base expectancy score method continues to be a valid measurement and predictive device for male felon parolees. Although the BE was created to predict favorable outcome within 2 years, it has some validity for predicting returns to prison, in that the percent of returns generally increases as the BE score level decreases. The only way the parole board could retain all recidivists in prison would be to release no one. As the number of detained persons declines, the number of confined non-recidivists declines as well, but so does the percentage of recidivists.

According to Greenberg, the most elaborate attempt to develop methods for predicting violent recidivism was that of Wenk and Robison, who studied the violent recidivist offenses of California Youth Authority wards. These youths have a higher rate of overall recidivism and a higher rate of return to violence than adult parolees, making them a logical target for a policy of selective incapacitation.

Using a different approach, a psychologist developed a multiple regression prediction using 18 variables, and concluded:

"...it appears to be feasible to develop, in this sample, at least, an index of violence proneness that would correctly identify over 50 percent of those individuals violating parole by violent offenses at the cost of misclassifying no more than 10 percent of those not returned for violent offenses."

34:548

A recidivism rate is not to be confused with the percentage of parolees who are eventually reconvicted or returned to prison. The rate involves not only the number of violators or violations, but the period of time over which the violations occur as well. Thus, the finding that the recidivism rate was no more than 20 percent in the first year and less thereafter, would not contradict a statement that more than half of all parolees are eventually returned to prison. Persons returned to prison are in a legal sense far from homogeneous. Some have been found to have committed felonies, others misdemeanors, and the majority have not been found to have violated the law at all.

Parole, instituted in the decades following 1870, reflects views of crime causation held in the late nineteenth and early twentieth centuries.

Crime was compared to a disease and penal administrators to physicians who could cure criminals of the personal pathology that led to their initial involvement with crime. The parole system was seen as an administrative device that would simultaneously permit the retention in prison of those whose disease had not been cured and the speedy return to prison of those who were beginning to relapse. Another mechanism involves the effect of the high recidivism rates generated by the parole system in reinforcing public stereotypes of released prisoners as especially dangerous. This contributes to the difficulties faced by released prisoners in such areas as employment and social life. For the 1970 male parolees returned to prison, the bulk of the returns were for property crimes not involving confrontation with a person, or violations of the drug laws.

Because data are most complete for index crimes, Greenberg attempts to estimate the rate at which released prisoners commit index crimes, and the corresponding magnitude of the incapacitative effect of imprisonment, assuming that the recidivist crimes would not be committed were the parolees not released. Since most of the current interest in incapacitation concerns offenses against the person, it must be pointed out that most of the index offenses are those involving theft where no confrontation with a victim occurs. Criminal activity need not be of uniform intensity or character throughout the crime career, and parole recidivism data suggest that recidivism rates decline with time, or involvement in criminal behavior declines with age. ^{34:559-561}

Unlike subsequent editions, the 1965 edition of Uniform Crime Reports distinguishes index from nonindex arrests in summarizing information about crime careers. Index crimes can be committed by persons in any of the following categories: ^{34:563}

1. Those who have already acquired an arrest record on a nonminor charge.
2. Virgins who are committing an index offense for the first time and who are arrested for it.
3. Virgins who are not arrested for the index crime they commit.

The number of persons with nonminor arrest records and whose crime careers have not ended is VT, while r is the number of crimes each commits per year. Denoting by V the annual number of virginal arrests for an index crime, we have the following inequality: ^{34:563}

$$C > V_I + rVT$$

where C represents the annual number of index crimes committed in the United States.

An independent, though somewhat crude estimation of the amount of recidivist crime can be obtained using parole recidivism data, to determine from information about returns to prison with new commitments or allegations something about the magnitude of involvement in new crimes, whether or not detected. The official clearance rate allows us to link the number of reported crimes with the number of arrests. Parole recidivism data tells us the number of returns to prison for index offenses. ^{34:568}

Greenberg quotes former U.S. Attorney General Ramsey Clark:

"Much of our crime is caused by the inhumanity of our prisons and by our failure to rehabilitate those we send to them.... better than one-half of all the people who leave prisons return convicted of a subsequent crime..."

This is an extremely misleading depiction of our crime problems. More than half of those who leave a prison do not return with a new criminal conviction. The rate of return to prison is indeed high, but most returns are not the result of new convictions. ^{34:575}

Statistical comparison of the recidivistic behavior of different groups is used to compare relative effectiveness of rehabilitation programs,

usually defining recidivism as the number of releasees returned because of a new violation. One alternative method, known as Failure Rate Analysis, is not dependent on random samples and does not entail discarding data, but still prevents bias due to different total exposure times. It is commonly used to analyze equipment failures and is drawn from reliability theory. Modeled as a process in which individuals from a population fail in time, the probability function then is:

$$P(t) = P r \text{ [an individual will fail within the time period } (0, t)]$$

Harris and Stollmack describe the F-test and comparison of release programs to compare failure rates. Working with this method, they found challenge to the commonly accepted hypotheses that the propensity to return to criminal behavior decreases with the time after release from prison.

Harris and Moitra describe recidivism as, "tendency to lapse into a previous behavior mode," and cite these measures of failures:

1. Arrest
2. Escapes
3. Convictions
4. Incarcerations

They define failure rate as the number of failures observed during period of concern as a fraction of total time in which the failure events could have occurred (the total of all times the individuals are exposed to the "hazardous" environment). The value of failure rate (H) "provides an estimate for the (time) rate at which individuals could be expected to fail out of a given program over particular enrollment risk period."

$$H_i = \frac{\text{number failed in observation period } i \text{ (say, one day)}}{\text{total man-days of exposure in day } i}$$

$$= \frac{\text{number failed on day } i}{\text{number started on day } i}$$

which estimates the conditional probabilities of failure in each period i, given success up to that point

(In order to smooth out the function)

$$h(t) = \text{Prob [failure on day } t, \text{ given success until } t]$$

then the Probability distribution function (CDF) for the random length of participation

$$F(t) = \text{Prob [program time } \leq t]$$

$$= 1 - \exp \left[- \int_0^t h(u) du \right]$$

then the density is the derivative $f(t) = dF(t)/dt$

Their work led to lifetime models which have the distribution functions:

$$F(t) = 1 - e^{-\lambda t} \quad (t \geq 0; \lambda > 0)$$

and

$$F(t) = 1 - e^{-\lambda t^\alpha} \quad (t \geq 0; \lambda > 0; \alpha > 0)$$

"The failure-time density function, for those individuals who will fail, $f(t)$ is assumed to be exponential by both Stollmack and Harris, and Maltz and McCleary as:

$$f(t) = \lambda e^{-\lambda t}$$

and the definition of the failure rate or hazard function $h(t)$ is:

$$h(t) = \frac{f(t)}{1-F(t)}$$

Harris and Moitra say that their method of predicting recidivism basically attempts to apply the concept of measuring "rates of failure relative to time, as is commonly done in device reliability modeling and mortality analysis in biostatistics," which rely on the form:

$$\text{failure rate} = r = \frac{\text{Number of failures observed during period of concern}}{\text{Total time in which the failure events could have occurred}}$$

Failure can be arrests, escape, conviction, incarceration, etc., and "a recidivist is thus one who succumbs to one of these defined hazards within an appropriate time period." Webster defines recidivist as an habitual criminal.

Harris and Moitra feel the introduction of failure-rate analysis to criminal justice provides a better statistical tool for measuring the performance of different rehabilitation techniques or programs. Their objective is to introduce concepts and procedures that make more efficient use of available data and standardize measurement so that more valid and timely comparisons are feasible. The use of the failure rate in analysis begins with the computation from the raw failure/success data of the quantities

$$r_i = \frac{\# \text{ failures in observation period } i \text{ (e.g., one day)}}{\text{Total man-days of exposure accrued in day } i}$$

Harris and Moitra, in order then to provide a more rational function form for the potential program times of subject individuals, smooth out the observed (and usually erratic) set of $\{r_i\}$, and call the smoothed form

$$r(t) = \% \text{ failures on day } t \\ = \text{Prob}[\text{failure on day } t, \text{ given success up to } t],$$

from which one can derive the probability distribution function for the (random) length of participation. The most common forms of smooth failure rates that satisfactorily describe empirical patterns of correctional failure over time are either a constant failure-rate model or one where the failure rate has a simple dependence on time. These lead in turn to lifetime models which are called the exponential and Weibull, respectively. A statistical test of that data to determine whether they are indeed well fit by one of the simple smooth models is the F-test, which has been shown to perform quite well against a wide range of reasonable alternative hypotheses.

The application of the technique generally involves two stages. First observed or empirical failure rates are computed and the statistical distributions associated with their failure times established. By itself, this step could provide some significant information on the incidence of future recidivism. The second stage in this kind of effort is the comparison (also using the F-test) of failure rates of alternative programs in order to evaluate program effectiveness.

This is not the first effort to model time-phased criminal justice/law enforcement problems as stochastic models, but earlier efforts were strictly probabilistic in nature and did not concern themselves with such statistical problems as estimation and hypothesis testing. Statistical methods used in criminal justice program evaluation have for the most part been crude and highly inappropriate. Higher costs and decreasing revenues available to all sectors of the criminal justice system require

that funds be funneled into areas that impact most on the control of crime. Harris and Moitra find that one cannot incorporate recidivism into population projections as simple percentages, but instead must use a time-based measurement.
35:85

Harris and Moitra applied developments in the failure-rate methods, the modeling of program outcomes as failure patterns or rates over time, in a number of specific cases including to the data from Kantrowitz's California parole study. General adoption of this kind of methodology would lead to the development of standardized techniques of evaluation. It would also lead to more effective data-base management, studying recidivism, applying and extending statistical techniques commonly used in reliability engineering and mortality modeling where the statistical phenomena are essentially the same. These phenomena are failure or nonfailure within some period of exposure to a "hazardous" environment, which is relatively insensitive to the varying gradations of outcome definitions including offense scaling, but nevertheless is adaptable to the most universal set of problems possible.
36:194-195

According to Van Alstyne and Gottfredson, one trend in the development of prediction in criminal justice has been toward increased statistical sophistication of methods including log-linear analysis.
77:172-174

The development of statistical prediction techniques has been a central methodological and substantive area of concern to criminologists for over 50 years, and statistical prediction techniques have had applications for parolee and probationer risk assessment (e.g., Glaser 1955, 1964; Gottfredson

and Ballard, 1965; Ohlin, 1951; Mannheim and Wilkins, 1955). Prediction techniques have also played an important role in structuring discretion in the criminal justice system, such as at the parole release decision (e.g., Gottfredson et al., 1975) and at the sentencing decision (Wilkins et al., 1976), by helping to make explicit the policy underlying these decisions. Statistical prediction methods have been useful in evaluation research as statistical controls, when experimental designs that incorporate random assignment to experimental and control groups were not feasible (Gottfredson, Gottfredson, and Garofalo, 1977). Development has proceeded from simple measures of discrimination (Burgess 1928; Glueck and Glueck, 1930), to multiple linear regression (Mannheim and Wilkins, 1955), and to various numerical taxonomic techniques (Wilkins and Macnaughton-Smith, 1964; Fildes and Gottfredson, 1972). Each statistical development in this historical progression has been justified as being theoretically more appropriate for the constraints of the data. Concern over the levels of measurement of most predictor candidates and their potential joint effects led to the suggestion that the use of predictive attribute analysis (Wilkins and Macnaughton-Smith, 1964) may provide greater predictive efficiency. Violations of the homoscedasticity of variances requirement concerning multiple linear regression led to the suggestion that some forms of numerical taxonomy might prove useful (Gottfredson, Ballard, and Lane, 1963). Solomon (1976) has introduced another statistical tool, log-linear analysis, that may be used as an aid to criminological prediction, adapted by Solomon for the construction of parole prediction tables. The log-linear technique provides a way to examine all of the

possible complexities of a multiple contingency table and to isolate the most important associations and interactions, furnishing two useful tools (Davis, 1975):
77:174

1. It provides estimates of the effects of variables acting alone or in conjunction with others, focusing the analysis on those effects demonstrating a significant contribution to the variation in cell frequencies.
2. It provides a means of indirectly testing hypothesized relationships among the variables by setting up competing models and comparing the expected cell counts of these models with the actual observed cells.

Stollmack has done some work attempting to predict prison population. According to Stollmack, the most common method for predicting incarcerated populations appears to be extrapolation of linear trends determined by linear regression techniques, relying totally on past data of the number incarcerated. He goes on to say, there are two basic types of descriptive mathematical models:
73:142-161

1. Deterministic
2. Probabilistic

"The former assumes that all parameters of the process are known and invariant while the latter assumes that the process is affected by a multitude of factors which can be predicted only in a probabilistic sense."
73:142

Deterministic models of population growth assume that the rate of increase (or decrease) in population at time t, N(t), is a function of the size of the population at that time

$$\frac{d}{dt} N(t) = f[N(t)].$$

where $\frac{d}{dt} N(t)$ is the time-derivative of N(t). When it is assumed that the function, f, is a simple linear function with a proportionality constant, B,

$$\frac{d}{dt} N(t) = BN(t)$$

The deterministic model for the number incarcerated is developed from two very simple axioms:
73:143

1. The rate of change in the population at any point in time is equal to the input rate minus the release rate.
2. The release rate at any point in time is equal to the turnover rate at that point in time multiplied by the actual population at the same point of time.

These two statements can be combined in a single mathematical expression as follows:

$$\frac{d}{dt} N(t) = \lambda_t - \mu_t N(t)$$

where:

$N(t)$ - population at time t

$\frac{d}{dt} N(t)$ - rate of change in the population at time t.

λ_t input rate at time t.

μ_t turnover rate at time t.

The prediction problem boils down to one of estimating the parameters λ and μ .
73:143-144

For the Probabilistic Model, the population growth phenomenon is classified according to probability theory as a stochastic process; i.e., a time-varying probabilistic process. The models applying to population prediction are referred to as continuous-time, discrete state stochastic models which are treated extensively as part of the analysis of queuing systems, reliability theory, and general counting processes. The predictions made using stochastic equations are expected values; i.e., they represent our best estimate of what the "average" population would be in a given month. The term "average" implies an ability to view the process over several identical periods of time. The expected value, then, for a given month after the initial period should correspond to the average for that month over all such trials with the process.

The probabilistic model is developed in two stages and the results of the two stages are added together to obtain the overall prediction model. In the first stage the model used is referred to in the literature as a "simple death process model." The second stage is handled as an infinite-server queuing model providing a "server" for every new arrival with service rate μ . A recidivism rate, $1-p$, can be incorporated into the model writing:
73:147

$$\frac{d}{dt} N(t) = r + (1-p)\mu N(t) - \mu N(t)$$

Thus, the relationship between the probability of not recidivating and the length of stay can be seen as the critical factor.

A closed system model incorporates separate turnover rates: r for parolees returning as violators, and q for those succeeding and being released upon expiration of sentences. The turnover rate is the inverse of the average length of stay unless the distribution of this variable is markedly nonexponential. If length-of-stay data are not available, and if the input rate has not recently undergone a significant decrease or increase, the ratio of the number released during any period of time to the average population during that time can be used as an estimate of μ . Basically the input rate is the product of several other rates such as the arrest rate, the rate at which arrest cases are indicted, the conviction rate, and the incarceration rate.
73:150
The parole projection given in this model is entirely a derivative of the projection for the number of incarcerated felons.

Recidivism can be anything up to ten times more frequent in offenders who have received suspended sentences than in those who have actually had to serve their terms of imprisonment, according to Fontaine's study of 1 week's (11/22-28/65) criminal court cases in France (1/300 of the total annual criminal court turnover in France at the time). He found:
29:138-143

1. As a function of age, once-only offenders are distributed normally. In France in 1965, the function was defined by an average of 30.5 years with a margin of 14 years on each side.
2. The population of habitual offenders manifests itself according to a period of half-life of 7 years, commencing at 14.5 years of age.

- 3. The tendency towards recidivism of habitual offenders decreases by 2 percent per year of age, commencing at 14.5 years.
- 4. Recidivists commit 1 1/5 times more known offenses than first offenders.

If n crimes committed during a given year in small area with population M, if each crime has 1 victim and each member of the population is equally likely to be a victim, the probability that a given member of the population has not been affected is, according to Avi-Itzhak 3:185-213 and Shinnar:

$$\left(1 - \frac{1}{M}\right)^m$$

For large values of M, assume $m \rightarrow \alpha M$ where α = average crime rate in the area then:

$$\left(1 - \frac{1}{M}\right)^m \rightarrow e^{-\alpha}$$

The probability that a person living in an area with a crime rate α will not be affected in a given year is $e^{-\alpha}$.

If the average lifespan is 70 years, then the probability of not being affected during a lifetime is $e^{-70\alpha}$ and the probability of being a victim of at least 1 crime in a lifetime is approximately $1 - e^{-70\alpha}$.

If one assumes the individual offender at the start of his criminal career commits offenses at a Poisson rate λ_0 ; at this stage the probability of a crime being cleared and the criminal being prosecuted and convicted is q_0 .

If the first conviction results in a period s_1 , during which the offender is neutralized, there is probability ϕ_1 , that the offender will emerge still active at the end of s_1 ; then the probability of his criminal career terminating during s_1 is $1 - \phi_1$.

If the random variable s_i describes the length of incarceration resulting from the i^{th} conviction (s_i may be = 0); ϕ_i is the probability that the offender survives (emerges still active) after the i^{th} incarceration, λ_i is his Poisson rate of offenses after surviving the i^{th} sentence and q_i is the probability of each such offense leading to a conviction. Then if we assume s_1, s_2 are statistically independent and that the length of the criminal career is exponentially distributed with mean $1/n$, the survival probabilities are:

$$\phi_i = \int_0^{\infty} e^{-s} dF_{s_i}(s)$$

3:196

"Repeat probabilities tend to increase with severity of prior records."

If one lets x describe the total number of offenses committed by an individual offender during his life, then the expectation of X, $E(X)$ X size criminal population = crime level. $E(X)$ is practically independent of the exponential assumption for distribution for length of criminal career (exponential is roughly same as constant career length). The expected number of offenses during a criminal career with exponential career length distribution is:

$$E(x)_e = \frac{1-p}{q} \phi_e = \frac{\lambda}{1-\lambda q} \phi_e$$

with constant career length:

$$E(X)_c = \frac{E(\phi_c)}{q} = \frac{1-p_0}{q} \phi_c = \frac{1-e^{-\lambda q}}{q} \phi_c$$

According to Wilkens, "Instead of providing a general statement of the probability for an individual to be reconvicted it would be possible to make estimates in conditional terms, such as:"^{80:}

Pr/a = if an offender is released and goes to his old address, his chance of reconviction is x%;

Pr/m = if an offender is released and goes into the armed forces, his chance of reconviction is y%

As an example of recidivism, a Georgia study cites: ^{13:181} Returns to prison within 6mo. 1yr. 2yr. 3yr. 4yr.
2.7% 7.7% 14.6% 18.6% 21.4%

NISHIMURA, in predicting recidivism in Japan, determined a number of independent variables effecting recidivism. His variables were:

- | | |
|-------------------------|-----------------------------|
| 1. age | 6. runaway |
| 2. type offense | 7. stealing |
| 3. motive of act | 8. concealed weapon |
| 4. unrecorded offenses | 9. living condition |
| 5. frequency of truancy | 10. family history in crime |

He established a score and cut off score. Validation study showed ^{52:118-120} 76.2 percent in conformity with his inventory of variables.

Both murder and felonious assault are mostly committed by career criminals ^{68:599} and could be prevented by incapacitation, according to Shinnar and Shinnar.

According to Forst, the view that incarceration deters crime is accepted by many as an article of faith; to others it is equally obvious that it does not. Gordon Tullock reviewed a number of studies that addressed the deterrent effect of punishment and concluded: . . . "we have to opt ^{30:1-21} either for the deterrence method or for a higher crime rate."

A key study upon which Tullock and others have based such a prescription for reducing crime was conducted by Isaac Ehrlich. A major empirical finding of Ehrlich's article is that public expenditure on law enforcement activity has paid "in the sense that its marginal revenue in terms of ^{30:2} a reduced social loss from crime exceeded its marginal cost." Using 1960 data, Ehrlich estimated that "a 1 percent increase in expenditure on direct law enforcement would result in about a 3 percent decrease in all felony offenses," taking explicit account of circular causation.

Using 1970 data, Forst used a simultaneous equation estimation technique in order to separate confounding factors and, unlike Ehrlich, found the crime rate to be virtually insensitive to cross-state variation in either the probability or length of incarceration. The index crime rate is the principal dependent variable of analysis. The crime rate reflects the number of offenses against the average person in a community no less ^{30:3-9} than it does the number supplied by the average person.

Forst's model attempts to enlarge upon Ehrlich's description of offenders ³⁰ and to reflect characteristics of potential victims. This ensures that

the variables we are most interested in analyzing are purged of their dependence upon factors that have been omitted from previous investigations.

Ehrlich's estimate of unit crime elasticities come directly out of a regression of the crime rate upon five variables:
30:14-15

1. Probability of apprehension and imprisonment
2. Average time served by offenders in state prisons
3. Median family income
4. Percentage of nonwhite residents
5. Income inequality

Forst used 13 explanatory variables. Two variables that are particularly potent in reducing the regression coefficient of the probability of imprisonment are the population migration and density variables. Forst concluded that the relationship that Ehrlich finds between the probability of imprisonment and the crime rate is primarily spurious.
30:16

Fleisher in 1966 attempted to measure the effect of economic factors on delinquency and found "the overall effect of income on delinquency to be negative" and that "unemployment appears to be a cause of delinquent behavior." Weicher in 1970 added to Fleisher's explanatory variables, and concluded: "Traditional sociological 'taste' variables, such as 'anomie' and 'the absence of a strong father figure in the home' have significant effects on delinquency; economic variables appear to exert no effect." Forst also concludes that Ehrlich's analysis concluding that imprisonment substantially deters crime appears under scrutiny not to be as convincing as he and others have concluded.

But, evidence that arrests do, indeed, deter crime has been presented by Carr-Hill and Stern.
30:21

Research on the functions of imprisonment has begun to provide quantitative, empirical knowledge of its rehabilitative and deterrent effects, but less is known about the incapacitative effect of imprisonment. Greenberg distinguishes between selective incapacitation, the prevention of crime through physical restraint of persons selected for confinement on the basis of a prediction that they will engage in forbidden behavior, and collective incapacitation, crime reduction accomplished through physical restraint no matter what the goal of confinement happens to be. He says continued imprisonment of only the violent offenders would have prevented a very limited amount of homicide, at the cost of imprisoning several thousand offenders whose initial offense involved violence, but who did not, to the best of our knowledge, become involved in a new homicide or manslaughter after release.
34:542-543

According to McGuire, "incarceration removes criminally productive individuals from contact with free society, interrupting their criminal careers. For each individual confined, the incapacitation benefit is the value of the crimes avoided." For the agency responsible for their being there, it is the summation of these benefits and is related to:

1. Number confined
2. Likelihood to commit crime
3. Loss per crime

The "incapacitation benefits in any period are inversely related to the magnitude of the displacement effect, a measure of the longrun supply elasticity of criminal activities."^{48:13}

According to Deutsch and Malmborg, one question which the Blumstein - Nagin model did not address involved the determination of the relative impacts of incapacitation and general deterrence for a given imprisonment policy, which has potentially important implications for the effectiveness of a policy.^{23:139-147} Deterrence tries to reduce crime by posing a threat of punishment, while incapacitation reduces crime by isolating the criminal from the rest of society through imprisonment. Deterrence operates to reduce criminality, by posing a threat of punishment for any crimes they might commit, while imprisonment can reduce crime through incapacitation. They model the magnitude of the incapacitative effect as directly related to λ , the rate at which offenders commit crimes while free in period t . The number of crimes prevented by the incapacitative effect during period k can be determined by accumulating the portion of the incapacitative effect in those previous periods, which was operative in period k .^{23:142}

Deutsch and Malmborg found the incapacitative effect of the current sanction level was significant, yet clearly a subordinate effect to general deterrence, and suggest incapacitation entails about 20 percent of the effect of sanctions under current policy. Blumstein and Nagin had speculated in their theses that about 30 percent of that savings was due to incapacitation and about 70 percent due to deterrence.^{23:144-147}

According to works edited by Brounstein and Kamrass, society keeps a constant proportion of its population in prison, regardless of the crime rate.^{53:8-10} Then, if crime rates go up, the CJS raises the threshold for going to prison, rather than simply sending more people to prison.

The estimate of crimes averted by imprisoning a criminal with crime rate λ for S years is simply λS . We have good evidence on arrest rate, but we have very poor evidence on crime rate. And any estimate of the effect of incapacitation inherently requires information on individuals' crime rates. An incapacitation policy relies on an implicit assumption that an individual's future criminality can be predicted and is sufficiently high to warrant imprisonment. The stable imprisonment rate we have in the United States for the past 45 years -- 110 per 100,000, with a coefficient of variation of only 8 percent -- does indeed represent an important constraint on imprisonment. The fragmentary deterrence evidence seems to suggest that the "certainty" of punishment deters more than the "severity."^{53:10}

Von Hirsch noted that "predictions of dangerousness have historically been used to justify confining mentally ill persons." Canada has a "Preventive Detention" law for persons with multiple convictions. Also, under the Maryland Defective Delinquent Law, persons convicted for the first time of any of a wide variety of offenses, may be indefinitely confined based on two sets of criteria for decision:

1. Quasi-psychiatric.
2. Demonstrated behavior indicating danger to society.

"Since 1933, Illinois has made regular use of statistical prediction techniques in its parole system" based on an actuarial prediction table of parole outcomes. Ohio, California, and Colorado have also developed formal prediction tables for parole decision use. Von Hirsch also notes a need for reasonably precise legal standards of dangerousness. The "prediction method used must be subjected to careful and continuous validation."
78:717-725

Deutsch and Richards noted that several analytical and simulation models have been used for evaluating the CJS and that modeling the performance of the CJS is the mainstay of many ongoing studies of crime. The CJS models currently being used for evaluative purposes are either analytical or similar in character. The analytical models have been used chiefly to predict the recidivism rate, whereas digital simulation models have been used variously for forecasting resource requirements, reducing court delays, and predicting CJS operating costs and recidivism rates. The Generalized Network Simulator (GNS) is seen as a vehicle by which such efforts may achieve their modeling objectives.

The analytical model form first appeared in the 1967 Presidential Commission's Task Force Report: Science and Technology. Christensen developed several simple but illuminating models. One model forecasted the number of first offenders who are arrested per year, while other models approximated the number of convictions that could be expected during any recidivist's criminal career. The analytical models that have appeared in the Criminal Justice literature are aggregate in nature.

Belkin, Blumstein and Glass developed a feedback model of the CJS which contained only two components of the CJS: combined police and judicial component, and corrections component. Their objective was to model the entire criminal career.
22:1

The first model to possess recognizable policy variables was developed by Avi-Itzhak and Shinnar, and later refined by Shinnar and Shinnar. They modeled the criminal career of an offender and incorporated the incapacitation effect of the CJS into the model formulation.
22:2

Two policy variables were included:

1. Length of incarceration.
2. Effectiveness of the police and the prosecution.

By assuming an incapacitative effect,

$$E(x) = \frac{\lambda T}{1 + \lambda qJS},$$

where q and S are the policy variables representing the joint probability that an offender is both arrested and convicted and the actual time served in prison, and J is the conditional probability that an offender is incarcerated following conviction. Another policy model was formulated by Blumstein and Nagin, which examines the deterrent and incapacitative effects of incapacitation on the crime rate.
22:2

Unlike their analytical counterparts, the simulation models have emphasized the operations of the CJS as opposed to the characteristics of the offender population. They deal directly with the issues of CJS policy-making. Whereas the performance measures of the analytical models have been the crime rate, the performance measures of the simulation models are varied.
22:3

The first serious attempt to model the operations of the CJS was by Navarro, Taylor and Cohen. Their model, called COURTSIM, makes use of the General Purpose System Simulation language to trace, on a day-to-day basis, the paths along which offenders progress through the Washington, D.C. judicial system. Unlike the COURTSIM model, the JUSSIM model does not deal with individual offenders; consequently, queuing phenomena cannot be examined. The JUSSIM model is driven by a forecasting function of the total arrest rate, and has the ability to capture the essential characteristics of the CJS and estimate the cost of alternate system loads.

Blumstein and Larson also introduced an extension to the JUSSIM concept called JUSSIM II, a feedback model wherein offenders are tracked from the point of their first arrest to the point where they finally leave the CJS for the last time. JUSSIM II includes measures of criminal recidivism. JUSSIM II determines the most serious crime for which an offender is charged by invoking the Markovian assumption; the current offense depends solely upon the type of the immediately preceding crime. Wolfgang, Figlio and Sellin tested this assumption with their male birth cohort and found this model to be an acceptable representation of crime-switching behavior.

Pittman was able to estimate future system loads and the crime mix given the number of first offenders who are arrested and convicted. The expected number of times the offender is rearrested, the average sentence length, the expected criminal profile, and the expected

career cost of an offender were all computed analytically under steady-state conditions.

In 1972, a queuing model of the entire CJS, called DOTSIM, was designed, incorporating a model of offender recidivism similar to that demonstrated by Blumstein and Larson. DOTSIM, like COURTSIM, follows each simulated offender through the CJS, and has the capability of delaying the processing of offenders whenever the demand for a particular resource exceeds its supply.

Simulation models have provided a great deal more flexibility than their analytical counterparts, but future work in modeling the CJS could combine the attributes of both model forms.

In 1973, the Advisory Group on Productivity in Law Enforcement noted that the most common data used for judging overall police performance are crime rates, such as compiled in the Uniform Crime Reports published annually by the FBI. Some hope was offered for getting more accurate crime data through victimization surveys, confidential and detailed surveys of scientifically selected samples representative of the population as a whole, to detect the true number of crime victims. These may provide new measures for crime control and crime prevention programs, and may show why crimes were not reported, as well as the victims' attitudes toward the police and police service.

The Advisory Group chose three objectives of police patrol for consideration:

1. Deterrence of crime.

- 2. Apprehension of criminal offenders.
- 3. Satisfaction of public demands for noncrime services.

These three objectives are closely related. For example, better noncrime services enhance the image and public support of the police department, thereby strengthening crime deterrence and apprehension efforts. To meet these objectives, the police force carries out a variety of activities, any one of which may contribute simultaneously to one, two, or all three of the objectives. The activities include observation, response to calls for service, enforcement of the law, investigation, maintaining order, and various administrative and postarrest activities. Since any one activity may contribute to all three objectives, and since the objectives themselves are interrelated, the measurement and analysis of the police force can be a complex undertaking. A payoff comes in using existing manpower to the greatest advantage.

54:13-14

Many police departments keep statistics needed to compute productivity measures adaptable for widespread use, and the range of performance for a variety of measures suggests a potential for productivity improvement in most departments. A simple measure used to help determine the ability of management to make manpower available for patrol is:

54:15

$$\frac{\text{Patrolmen Assigned to Street Patrol Work}}{\text{Total Patrolmen}}$$

A measure used to indicate the extent to which patrol time in the field is being committed to patrol activities is:

54:16

$$\frac{\text{Man-Hours of Patrol Time Spent on Activities Contributing to Patrol Objectives}}{\text{Total Patrol Man-Hours}}$$

A principal objective of most police departments is to deter crime, but no persuasive relationship between overall patrol activities and crime deterrence has been established. In the absence of a direct measure of deterrence, three types of substitutes might be used:

- 1. Existing reported crime indices used with discretion.
- 2. Victimization surveys.
- 3. Quantitative measurement of activities which professional judgment suggests contribute to deterrence.

54:23

A measure for apprehension productivity is:

$$\frac{\text{Arrests Resulting From Patrol Surviving the First Judicial Screening}}{\text{Total Patrol Man-Years}}$$

54:24

An apprehension productivity measure is:

$$\frac{\text{Felony Arrests Resulting From Patrol Activities Surviving First Judicial Screening}}{\text{Total Patrol Man-Years}}$$

In providing noncrime services, a force's productivity may be determined by the following measure:

54:27

$$\frac{\text{Number of Noncrime Calls for Service Satisfactorily Responded To}}{\text{Man-Hours Devoted to Noncrime Service Calls}}$$

Three kinds of action might be considered for getting a larger proportion of the patrol force in a position where they can contribute:

54:30-31

- 1. Use of nonsworn personnel.
- 2. Combining jobs.
- 3. Transfer of services.

The amount of time which a patrolman assigned to street work can actually devote to important patrol activities can be increased in at least three ways:
54:31

1. By reducing his responsibility for nonpatrol activities.
2. By better use of time-saving equipment.
3. By simplifying necessary but time-consuming administrative chores.

An obvious but often overlooked element in effective patrol is to have people on call when and where they are most needed. In team policing, officers are assigned to the teams on a permanent basis, they are permitted and even encouraged to develop flexible work schedules which enable them to make necessary followup investigations of crimes. Possibly the most innovative potential for improving patrol effectiveness is to make patrol work anticipatory rather than reactive.
54:34

An econometric study of the factors contributing to crimes against property and the factors determining the effectiveness of law enforcement activity directed against these crimes was carried out in 1969. Among its findings are:

1. Deterioration of labor market opportunities for youths, particularly nonwhites, was one of the principal factors responsible for rising per capita offense rates for economic crimes.
2. Increasing school enrollment rates for youths have had an ameliorating effect on the rise in crime rates for some types of crimes.
3. The decline in police effectiveness measured by the ratio of offenses cleared by arrest to known offenses has encouraged criminality and induced higher rates of growth in per capita offense rates.

The report also found that "approximately 98 percent of the rising trend of economic crime is explained by the worsening of economic conditions."
54:44

An indicator of the quality of personnel programs is the total departmental turnover compared to its manning strength, calculated for both sworn and nonsworn personnel:

$$\frac{\text{Total Turnover During the Year}}{\text{Total Number of Department Personnel}}$$

The result must be qualified by the type of person who leaves, determined by performance appraisals. Unnecessarily high turnover can be traced, among other factors to poor management, improper selection and assignment criteria, and few opportunities for growth. Very low turnover is equally undesirable, since it is symptomatic of organizational stagnation and lack of growth. Four major programs contribute to maintaining the quality of personnel at the highest levels:
54:52-53

1. Recruiting
2. Selection and assignment
3. Training
4. Organization development

One measure of recruitment effectiveness is:
54:53

$$\frac{\text{Number of Man-Years Lost Due to Unfilled Vacancies}}{\text{Total Authorized Man-Years}}$$

Another measure to indicate the effectiveness of the recruitment program in attracting people who not only meet entrance requirements but also perform satisfactorily on the job is:
54:53

Total Number of New Hires Who Perform Satisfactorily After "X" Months
Total Number of New Hires

Organization development is less well understood and not as widely accepted as a police department activity, but is in every way as important as other activities, and while it is difficult to measure and evaluate, its importance demands that the effort be made. 54:58

The ingredients of a "productive" organization include open communication between levels, an incentive system that rewards inter-departmental collaboration rather than competition, the confrontation of differences, participatory decision making, and an organization structure that allows for flexibility. Considering how much time is spent communicating, the impact on productivity of even a small improvement can be enormous, with better coordination, saved time in explaining and repeating, fewer mistakes, and a better feeling about the department and the public because of improved performance. 54:59

Chapman developed a simultaneous model of crime causation, police output and demand for police and therefore incorporated this interdependence, finding police labor positively related to police output and property crimes more important than other types of felonies in increasing the demand for police. 14:48 Crime has been said to depend upon such varied things as the number of associations between criminals (Sutherland, 1939), the environment, the chance of being caught (Becker, 1968; Ehrlich, 1973) or the economic system of the country (Bonger, 1969). Chapman noted there are many different trends of thought that

can be utilized to explain crime, and saw two major currents of thought, an ecological school and an economic school as really variations on the same theme. 14:49

The ecological basis of crime can be related to an economic approach. An economist would postulate that, given his environment, a criminal would act in a rational manner to maximize utility. The ecologist would say that the environment is causing the criminal to act, whereas the economist would say that the criminal is acting, taking his environment into account. Chapman says Bentham was one of the first political economists to seriously analyze crime in this manner. He developed two basic criminological concepts: (1) human action can be reduced to one formula of motivation: pursue pleasure and avoid pain; (2) to prevent people from engaging in uncontrollable orgies of criminal behavior as they follow this precept, checks and sanctions may be established by legislation. 14:50

Becker's model was all-inclusive. It not only developed an equation for the supply of crimes but also a system that included losses to society from crime and from enforcement. It also developed an optimal enforcement rule. The basic Becker equation for the supply of crimes relates the number of offenses committed by any individual to his probability of conviction, his punishment if convicted, and a portmanteau variable that reflects other possible influences.

In general, for public service functions, output has been defined to be a function of inputs, input quality, and service delivery conditions.

For police, this involves problems in defining output and isolating factors, aside from inputs, which can be legitimately expected to influence the public delivery of that output. Not only must the relevant quantity demanded be determined, but it must also be decided what the independent variables are which affect that demand. For police this demand could be for the inputs of the production functions; e.g., patrolmen, and quite possibly, one of the determinants of demand would be crime, as explained by crime function. 14:50-51

It is generally assumed that the higher the payoff from going into crime, the greater the chances of an individual participating in the illegal activity. The environment colors the individual's perception of the payoff. Two basic elements underlie a supply-of-crimes equation:

1. The idea that individuals must have distinct preferences that dictate their rational behavior in certain situations.
2. In general, an individual will try to maximize the utility he gains from earning criminal and noncriminal income.

Generally, it can be said:

$$N_1 = f_1(E_i) \quad (1)$$

with: N_1 = the amount of crime committed because of the situational element

E_i = environmental situation i

f_1 = functional relationship for this equation (1).

The individual must do three things to maximize his utility:

1. Equate the ratio of marginal utilities of expected legal and illegal income to the ratio of the wages of illegal and legal income.
2. Equate the ratio of the marginal utilities of leisure income and illegal income to the illegal wage received.
3. Equate the ratio of the marginal utilities of leisure income and illegal income to the legal wage received.

The rate of substitution between the amount of illegal income earned through crime, and the amount of legal income earned is also equal to the ratio of the two wages; and so it can be said: 14:51-52

$$N_2 = f_2(W_c/W_y) \quad (2)$$

N_2 = amount of crime that a person commits because of this element

W_y = expected legal wage

W_c = expected criminal wage

f_2 = functional relationship for this equation

There is a great deal of interrelationship between the "environmental" variables and the "economic rationality" variables. Included within the crime supply equation must be the arrest rate variable that the potential criminal considers. Since the true criminal is unconcerned with false arrests, the total arrest rate is the relevant variable, and 14:53

$$N_4 = f_4(E_i, W_c/W_y, A/N) \quad (4)$$

with: A/N = the arrest rate

f_4 = functional relationship for this equation

Although police do many things, the attention of the public is often oriented toward how many arrests police make.

The arrest rate,

$$A/N = g(C, S) \quad (5)$$

with: C = set of physical police inputs

S = set of service conditions

g = functional relationship for this equation

The quantity of police labor that is hired in a particular city should depend upon its price--the wage rate--and other societal variables that impact on the citizens' perception of how necessary police are.

$$C_1 = h(N_1, V, W) \quad (6)$$

with: C₁ = per capita sworn policemen

V = set of socioeconomic variables that influence the demand for police

W = wage (price) of police

h = functional relationship for equation (6)

Equations 4, 5, and 6 constitute a general three-equation simultaneous model of crime and police response. This model thus has three dependent variables: crime rates, arrest rates, and per capita police. But these are interrelated; in fact, the arrest rate has an influence on crime, police impact the arrest rate, and crime rates can influence how many police are hired. As the relevant independent environmental variables, service conditions, and socioeconomic variables are specified,

a more accurate description of the relationships between police and crime is possible. The entire system can also be partially illustrated graphically demonstrating the interrelationships between the wage rate, the number of per capita police, the arrest rate, and the number of per capita crimes.
14:53

Chapman also cites two crime equations for two types of crimes:

1. A property crime equation
2. A violent crime equation

For the property crime equation, the economic variables are considered to be the relative wage variable, the chances of being arrested, and likelihood of being employed. For the violent crime equation, the economic variables are income, employment prospects, other criminal activity of the potentially violent criminal, and the chance of arrest. Environmental-type variables for the property crime equations include the extent of discrimination that the potential criminal faces, the social class of the individual, and the environmental factors that influence the potential success of the crime. Since little is known about the environmental factors that influence violent crimes, the degree of discrimination was utilized as the proxy variable for the environment. The economic rationality variable for the property crime equation, named WAGE, is defined to be the average per capita illegal wage rate for criminal activity in a specific city divided by the average per capita legal wage for legal activity of that city. The per capita illegal wage is, of course, unknown. Yet, it can be estimated

by the total losses that occur because of felony crime divided by the city population, that is:

$$\frac{\sum X_i L_i}{\text{POPULATION}} = \frac{W}{C}$$

X_i = number of reported property crime i

L_i = gross loss per property crime i

If the average illegal gains were equal but legal income differed in separate cities, the model predicts that the city with the higher median income would experience less property crime. The economic variables that influence violent crime are not expected to be as directly influential as the wage ratio is expected to be on the property crime rate. 14:55-56

A person involved in the commission of a property crime may be forced to commit a violent crime in order to successfully complete the property crime. To the extent that property crimes are economically motivated, then violent crimes, as joint products of economic crimes, are economically motivated. 14:56

The police production function relates a simple measure of one type of police output to physical inputs and service conditions. The service condition set of elements contains two basic influences:

1. Noncrime demands made on the police within the city.
2. Degree of cooperation that the police receive from the city populace.

These might be measured by the percentages of workers who use public transportation to get to work in the city and the percentage nonwhite in the city, where the first should be positively related to arrest rates and the second, assuming poor police/minority relations, should be negatively related. The crime rate should have two effects on the demand for police: 14:57

1. Actual need for protection by businesses and residents in the community.
2. The second effect is more psychological and grows as the fear of being victimized grows.

There are several interesting points that are apparent in the property crime equation. The wage rate variable is quite significant, however, its elasticity is relatively low. The most important variable in terms of elasticity is the employment variable which indicates that for a 1 percent increase in the percent labor force employed full-time, there is a 1.8 percent drop in the crime rate. The violent crime equation is quite strong with almost all of the variables being highly significant and all of the signs as expected. Property crimes are quite important as an explanation for violent crimes, for every 1 percent increase in the property crime rate there is a 1.1 percent increase in the violent crime rate. Arrest rates may significantly retard property crimes, and per capita policemen can significantly increase arrest rates. Property crimes have a much stronger effect on the demand for police than violent crimes, thus the citizen

might be interpreted to live more in fear of property crime than
14:59-61
violent crime.

In general, Chapman's study has indicated that a simultaneous approach
to examination of the relationships between crime and police can be
fruitful, that both economic and environmental variables play an important
role in the explanation of crime, that the chances of being arrested
can significantly retard property crime, while having little effect
on violent crime, and that property crimes have a large effect on
14:61-62
the demand for police.

Alt and Deutsch point out that broadly speaking, Massachusetts' Gun
Control law would a priori be expected to make an impact on two major
components of the CJS: the police and the courts. In order to test
this hypothesis they used the multiplicative empirical-stochastic model
of order $(p,d,q) \times (P,D,Q)_s$ in modeling the monthly crime occurrences
1:549-551
as tabulated in the Uniform Crime Reports.

Pfeifer and Deutsch presented a three-stage iterative procedure for
building space-time models, falling into the general class of STARIMA
models, characterized by autoregressive and moving average terms lagged
24:i-30
in both time and space. This model class collapses into the ARIMA
model class in the absence of spatial correlation. A flexible class
of empirical models, the multiplicative autoregressive moving average
model family, together with the model building procedure commonly
referred to as the Box-Jenkins method have proven very useful, but only
24:1
to single series data.

An alternative to univariate time series modeling is multivariate
time series modeling, which attempts to simultaneously describe and
forecast a set of N observable time series. A further refinement
of a general multivariate time series model can occur if the system
to be modeled exhibits systematic dependence between the observations
at each region and the observations at neighboring regions. This
phenomenon is labeled "spatial correlation." Models that explicitly
attempt to explain these dependencies across space are referred to
as space-time models. The Space-Time Autoregressive Moving Average
Model (STARMA) is characterized by linear dependence lagged in both
space and time, where first order neighbors are those "closest" to
the site of interest, and second order neighbors should be "farther"
24:l-4
away than first order neighbors, but "closer" than third order neighbors.

The most pressing question encountered when attempting to utilize these
forms of space-time models is, which of the model forms is most appropriate
for the data at hand? In a manner completely analogous to that of
univariate time series, STARMA processes are each characterized by
24:7-13
a distinct space-time partial and auto-correlation function.

According to Deutsch, there has been an increasing usage of statistical
methodology to analyze law enforcement problems. In modeling the
monthly crime occurrences as tabulated in the Uniform Crime Reports,
multiplicative empirical-stochastic models of order $(p,d,q) \times (P,D,Q)_s$,
as proposed by Box and Jenkins, have been employed, forming a starting
point for a quantitative evaluation mechanism. The need for the

adaptation of the control chart concept to the law enforcement scenario is easily visualized as an ongoing means of evaluation. It is desirable for the policy maker to receive information regarding program effectiveness as quickly as possible after the commencement of such a program. 25:5-6

For a temporal sequence of crime occurrences (Z) for a given index crime, the general form of the multiplicative model of order (p,d,q) x (P,D,Q) is given by 25:6

$$\phi_p(B) \phi_p(B^s) \nabla^d \nabla_s^D z_t = \theta_q(B) \theta_q(B^s) a_t$$

where $\phi_p(B)$ and $\phi_p(B^s)$ are the nonseasonal and seasonal autoregressive operators, $\theta_q(B)$ and $\theta_q(B^s)$ are the nonseasonal and seasonal moving average operators, ∇^d and ∇_s^D are nonstationary and seasonal differencing operators and S is the seasonal lag. When there is no seasonal component (P=0, D=0, and Q=0) the multiplicative model reduces to the ARIMA model of order (p,d,q) which is given by 25:6

$$\phi_p(B) \nabla^d z_t = \theta_q(B) a_t$$

Deutsch says the procedure is capable of detecting even small shifts with a high degree of accuracy, and concludes multiplicative autoregressive moving average models with an imbedded shift parameter, to capture potential changes in future crime occurrence, can be utilized after being transformed to a linear model representation. 25:17-22

He goes on to say that several computer programs have been developed such as:

1. AGGRE, a program which aggregates a 115 by 115 matrix of flows randomly into a minimum size matrix without losing any off diagonal flow.
2. AGGFLO, a program which must be supplied externally the parameter IBULL, with the actual data between the months 1 through IBULL obtained for the aggregated map.
3. LINEAR, a program which requires external input.

Deutsch also studied the effectiveness of the new law in Massachusetts as a deterrent to carrying guns and the commission of gun-related crimes, the offenses of homicide, assault with a gun, and armed robbery for a change in their occurrence levels. However, he failed to recognize important variables such as the Gun Control Act of 1968, passed during his data base period.

Intelligent forecasting of personnel needs is important to the public administrator in any political climate. In today's atmosphere of tax and budget cutting, it is vital and informal methods of estimating and justifying future personnel requirements are no longer convincing. In order to defend the legitimate needs of their agencies, public administrators must be familiar with empirical methods of personnel forecasting, methods that can be objectively defended as logical and sound, and methods that can be used with the framework of a zero base budgeting system. Among the approaches that might meet these needs are the population ratios method, the standards method, the needs approach, the economic determinates method, and the program specific method, according to Waldron and Altemose. 2:385-389

If the size of the population served determines the number of criminal justice employees needed, computing personnel requirements is a simple process. Multiplication of the projected population to be served times the ratio of personnel to population results in the number of needed personnel. The population method can be refined by breaking down the population by demographic characteristics such as age or sex and applying different ratios to these different parts of the total population. It can be further refined by analyzing employee requirements by job classification. Most public officials are familiar with this numbers game, since the population ratio method is the most common empirical method in use today. The reason that the approach is often more useful in political gamesmanship than in honest forecasting is that national or state averages frequently mean little or nothing to an individual police department. Since the goals, organizational structures, and duties of criminal justice agencies vary so greatly, the "average" number of employees may be far too little for one organization and far too many for another. The population ratio method also ignores the fact that the population served is only one factor in an agency's workload, and it may well not be the most important factor. ^{2:385-386}

The standards method corrects both of the major limitations of the population ratio approach. Agency workload is calculated not merely by raw population estimates, but by multiplying the population by a standard -- the number of services to be offered per person in the population.

$$\text{Employees needed} = \frac{\text{population} \times \text{standard of service}}{\text{employee productivity}}$$

Besides being more accurate than the population ratio method, the standards method has further advantages to the administrator besieged with employee associations or unions on one hand and angry taxpayers on the other. The method focuses on the two factors most important in determining the costs of government services provided and productivity. The many standards available for use in the various occupational specialties used in criminal justice agencies make this method of determining needs potentially objective, and therefore defensible. ^{2:386} However, arbitrary selection of standards invalidates the method.

The needs assessment method is similar to the standards method in that it compares workload to productivity. The key difference is that the requirements for service are empirically determined. Instead of using a standard for the number of services to be provided per client, we determine in our agency exactly how many services per client are needed. Its limitation is that the collection of the information necessary may be difficult and expensive. However, it is by far the most defensible method discussed so far, since the figures used are not based on national averages or on somebody's standards, but on what is really happening in the agency in question. ^{2:386-387}

The economic determinates approach regards the budget and therefore the number of employees, as fixed. The population served, the services provided, and/or productivity per employee must be altered to fit the budgetary constraints. The economic determinates method

can be refined to fit whatever level of budgetary sophistication is desired. If the necessary information is available, the needs assessment formula can be used instead of the standards formula. Rather than altering only the level of services or the population served, or productivity, two or all three of these variables can be adjusted to fit the budget. If it is possible to significantly alter salary levels (e.g., by substituting paraprofessionals for professionals) the salary level can be used as a fourth variable. The advantages of the economic determinates method to the administrator who must work within a predetermined budgetary figure are obvious. This specific objective information on the effects of proposed budget cuts has been instrumental in defeating unwise tax limitation amendments at the local level. 2:387

The program specific method addresses organizations which may have unusual staffing needs that are not addressed by the preceding formulas. Police and correctional agencies must maintain certain staffing patterns around the clock, despite the fact that these staffing patterns may be far from cost-efficient. Problems such as these require a detailed analysis of the program in question. The advantage of the program specific method is its ability to address special needs. Its major disadvantage lies in the fact that special needs are usually based on subjective judgment. Although the administrative judgment may be highly accurate in its perception of needs, justifying the need objectively and defending it often proves difficult. 2:387-388

According to Chaiken, et al., the State planning agencies, using their Federal and State funds, have added to direct funding to researchers from such Federal agencies as the National Science Foundation or the Department of Justice to aid development and implementation of overall CJS models. 15:20-45 The best known overall CJS model is JUSSIM, designed by Belkin, Blumstein, and Glass in the Urban Systems Institute at the School of Urban and Public Affairs, Carnegie-Mellon University. Detailed flow, cost, and workload estimates were made for the State, and the model was run, using the distribution of reported crimes as input. Output measures from the model were judged as reasonably good predictors of real-life observations. This work included recidivism feedback, which was to become one of the important features of the second-generation model called JUSSIM II. JUSSIM II, an interactive feedback model for criminal justice planning, takes account of the feedback effect of recidivists on the types and numbers of crimes. JUSSIM provides the user with estimates of the first-order effects on the workload and costs at each of the system processing stages under each of a number of proposed changes. The model forces the user to quantify his intuition about the interactions between one part of the CJS and another. 15:20-21

The JUSSIM model is an interactive computer program that operates on a data file representing the user's criminal justice system. The CJS must be modeled by the user as a linear steady-state production process where crimes and associated offenders are the basic unit of flow, and the processing stations are the different stages through

which the arrested offender passes. The CJS model is constructed by the user and is often graphically displayed as a flow chart. Individuals, both recidivists and new offenders in society, perpetrate crimes. Some crimes are detected, some not; some crimes are reported, some not. Reported crimes are processed by the police, arrests are made, and a fraction of arrestees are charged with a crime. These arrests become cases to be processed by the courts, and those convicted are assigned to the corrections subsystem. Parole and eventual release return convicted individuals to society. The emphasis of the model is on the units of flow, usually offenders, criminal acts, prisoners, etc. At each stage the units of flow consume resources, such as the time of police officers, and the model calculates the rate of consumption of the resources. The output of each stage goes to alternative stages in proportions called branching ratios. Real-life flows are more complicated than those that can be modeled. JUSSIM is not a case-by-case simulation in which each offender is followed through the system, but rather considers offenders in aggregate groups whose behavior can be described by the branching ratios.

JUSSIM calculates the downstream flows, the total costs, resource requirements, and workloads in a disaggregated form to provide the user with cost, resources, or workload for each stage, crime type, or subsystem. It was designed to operate in this interactive mode so as to make it accessible to the user who does not know computer programming, and to bring his judgment into the analysis process.

CONTINUED

1 OF 4

The process of preparing test case data for JUSSIM helps the user recognize the importance of indirect consequences of policy changes.

One important contribution of models of the overall CJS is to the development of statistics and to the collection, aggregation, and comparison of data that are consistent and compatible across the entire CJS. The interactive mode, allowing users to work from remote consoles, trains the user to think of the CJS in terms of system ramifications and gives the model the characteristics of a "management game." JUSSIM, or models of similar type, can be viewed as a catalyst for establishing the data collection and decisionmaking organizations necessary for improvement in any CJS. While JUSSIM is intended to be run on time-sharing computers in an interactive mode, it could easily be operated in a batch-processing mode for running a large number of test cases.

Verification refers to an examination of the internal workings of the model to make sure the model does what the model builder intended. The JUSSIM model is quite simple mathematically and has been verified. Validation, which means examining and testing models to see if their predictive and descriptive capabilities are accurate, has not been conducted on the JUSSIM model.

Each of the proposals that can be tested by using JUSSIM has an impact on offender flow processing, and the following types of proposals might be tested:

1. Effects of drug offender diversion programs.
2. Costs and savings of changes in the bail release program.

3. Impact of a police crackdown on burglary.
4. Impact of an increase in psychological counseling during incarceration.

JUSSIM's primary value is as a catalyst for developing a data collecting and policy recommending organization for the entire CJS, and for focusing attention on the implications of changes in one part of the system on other parts. Implementation requires the development of a description of the overall CJS in terms of flow and stages.

The CANJUS project was undertaken by the Ministry of the Solicitor General of Canada with the objective of developing a comprehensive simulation model of the Canadian justice system. The decision was made to employ the existing JUSSIM model. The name CANJUS refers to the project.

PHILJIM is an adaptation of JUSSIM designed to fit the needs of the Philadelphia Planning Council, viewed as a management tool to aid in deciding where to direct available funds for improving the CJS. PHILJIM is a linear model with a somewhat larger number of user options than JUSSIM. Because it predicts one year into the future at a time, it is not a steady-state model in the same sense as JUSSIM. PHILJIM has an option that accumulates backlog cases when resources such as the courts cannot handle the input load.

JUSSIM II was designed to include recidivism, the major source of feedback among CJS components. Recidivism has a time delay effect on the CJS so that the impact of any new program to change the CJS may not be felt

for several years hence. Recidivists may switch crimes and are reintroduced into the CJS crime-committing stage at later times, representing the various time intervals between the commission of crimes. 15:33-34

A Dynamic Offender Tracking Simulation (DOTSIM) model was developed, but is still viewed as a prototype model developed not for on-line implementation but for experimental use to demonstrate the usefulness of such models. Upon implementation it would provide a means for discovering and testing alternative planning policies.

Planning policies that are addressable with DOTSIM include those of the JUSSIM model and its descendants as well as questions relating to queuing delays and the random nature of the processing of offenders.

DOTSIM objectives are to:

1. Reflect the actual procedural step-by-step processing of offenders through a CJS.
2. Represent the correct utilization of the CJS resources at each procedural step.
3. Determine the time required for each step.
4. Determine queuing delays that result from unavailability of resources.
5. Account for information transfer delays.
6. Assign priorities to the processing of any crime type.
7. Use historical or desired policies.
8. Assign fully burdened direct and indirect costs based on utilization at each step.
9. Handle recidivism and any type of offender feedback.
10. Differentiate recidivists and virgin arrests.

Use of the model enables planners to predict resource workload and cost, as well as the extent of delays occurring in the operation of the CJS. To use DOTSIM, a system flow chart representing graphically the sequencing and interaction of offender flows and a historical data base on CJS operations must be constructed, and key parameters from it provided as input for the program. The model requires the distribution of the lengths of time spent on each processing step in the form of minimum, maximum, and most likely times. 15:36-38

The CJS Training Model was developed under a Georgia State Crime Commission grant in the summer of 1972. The designers of the model intended that it be a training aid similar to management games used in business schools or to war games used in the military. The theory behind the use of models of this type is that a user's understanding and decisionmaking skills in the criminal justice area will be improved by using the training model. The model is a deterministic simulation model. 15:40-41

Overall CJS models have already had some impact on the synthesis and analysis of planning policy in the CJS. One of the major benefits has been the indirect training of CJS planners that takes place while using models. The experience gained from overall CJS models has been beneficial primarily from the learning that has taken place on the part of both users and model builders. 15:44-45

Deutsch shows a total crime model expressed in equation form as;

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} - 0.45a_{t-12} + a_t, \text{ indicating the}$$

current level of crime incidence is dependent on the last value,

the value 1 year ago, the value 13 months ago as well as a part of the year ago residual and a current shock. From the difference equation 21:7-8 form of the model, minimum mean square error forecasts are obtained.

Robison, et al., note that the prediction of violence in offender populations has long been a dream of correctional decisionmakers, but simple classification procedures and multivariate approaches failed to yield an operationally practical prediction instrument that would warrant implementation in actual preventive or correctional practice. 65:393

According to Wenk, Robison, and Smith, a large proportion of the public is alarmed about criminal violence. They usually stay out of "dangerous areas" and sometimes purchase weapons. Both of these responses are likely to elevate the overall level of public danger:

1. Opportunities for victimizing those who remain may increase.
2. Chance of accidental injury, posed by the presence of a weapon in the home, may surpass the likelihood of deliberate injury by an intruder.

Citizens do not seem as concerned about the probability of personal injury as they are about the possibility of injury from a specific source, the stranger. 65:394

There is little doubt that the known offender in general and the known violent offender in particular are more likely than members of the public at large to commit an assaultive act. 65:394

In 1965 the California Department of Corrections Research Division developed violence prediction scale which employed, as predictor items, commitment

offense, number of prior commitments, opiate use, age, length of imprisonment, and institution of release. The result was the identification of a class of offender, 14 percent of whom could be expected to violate parole by a violent or potentially violent discovered act. This likelihood was nearly three times as great as that for parolees in general. If a perfect corrective intervention method were developed and applied to the members of the violence-prone class, all acts of violence by this class could be prevented. However, since this class is such a small part of the parolee population, only 8 percent of total violence on parole would be prevented by its isolation and special treatment, leaving 92 percent of violent parolee acts occurring as usual. 65:395

In the California Department of Corrections Parole and Community Services Division, all parolees released to supervision are classified into one of six categories according to past aggressive behavior. The usefulness of this classification procedure can be examined by analyzing the effectiveness of decisions based on a correctional setting. Special precautions taken for identifying and handling the violent offender seem unwarranted. 65:395

The class of offenders with the highest level of violent recidivism is composed of subjects who had been referred to psychiatrists. This group does not have a higher-than-usual rate of general recidivism, but nearly one in five of the violations that do occur are violent. 65:399

Considering the rarity of the phenomenon (only one in forty exhibited subsequent violence), it is difficult to imagine that, even with the most refined techniques, one could do much better than, say, to double

the best rates obtained previously. A useful violence index might be constructed if different predictive equations were developed for each ethnic group and if multiple analyses of variance were applied. 65:400

Works on operations research as applied to law enforcement, edited by Brounstein and Kamrass, indicate that techniques such as regression analysis or discriminant-function analysis are used to identify the predictive characteristics of a successful probationer. These techniques provide specific weights allowing the decisionmaker to appraise the relative importance of the various characteristics. 53:275

The first model developed for adult probationers was done in 1932 by Monachesi, followed by other models for adult probationers, including those of Sheldon and Eleanor Glueck, D. Glaser and R.F. Hangren, D.V. Babst, D.E. Frease, A.W. McEachern, and others. These studies identified specific characteristics that appeared to be related to success or failure on probation. Some of these characteristics were: marital status, property possessed by probationer, previous criminal record, crime for which convicted, size and type of community in which convicted, length of maximum probation sentence, unemployment at termination of probation, age at conviction, size of family, usual occupation, socio-economic level, probationer's predominant values, type of leisure associates, stability of residence prior to probation period, economic dependence or ability to obtain self-sufficiency. 53:276

Several techniques have been used to develop prediction models. One widely used technique employs the Burgess method of weighting the variables found to be predictive; each item is given a weight of unity, but it does not take into account the possible interrelation among the factors. The Glueck method provides for weighting each variable although, as in the case of the Burgess method, it ignores intercorrelation that might exist between the variables. The Glueck method assigns each variable a given weight based on the maximum percent difference between any subclass of the variables and the recidivism rate of the entire sample. 53:276

To validate a model, a score for each probationer must be computed.

Probationers having high scores are expected to be successful. Once

the model scores are computed, statistical methods are used to determine if there is a relationship between a person's model score and his actual outcome. If the statistical test shows that a nonchance relationship exists, the model is considered valid. Statistics such as the t-test can be used to validate the models. For the t-test, the average score for the successes is compared to the average score for the failures. 53:278

California Base Expectancy Forms 61A and 61B were derived using multiple regression analysis on 875 parolees, and validated on sample of 900 parolees. A third model was developed using association analysis. Another model, labeled Salient Factors, was developed for the U.S. Board of Parole using a sample of 225 male Federal parolees. Chi-square tests were used to determine which characteristics were predictive of poor risks on parole. These characteristics were then weighted using the

Burgess method. Other models are the Oregon State Probation Model, a judgmentally modified version of a California model; and the Newark Probation Model, a judgmentally modified version of a California model. 53:278-279

58:80

According to Carlson and Palmer:

1. Linear regression can be used to measure nonlinear relationships.
2. Linear regression can deal with subsets within a sample.
3. Linear regression can be used to measure relationships where the explanatory variables are not independent of each other.
4. Qualitative variables must be handled differently from quantitative variables.
5. Unless pooling or weighting techniques are used, the results of linear regression with a binary dependent variable are inefficient and imprecise.
6. Maximum likelihood regression methods allow a researcher to use Tobit and multivariate logistic models which have many applications in criminal justice research.
7. Validation methods when improperly used can be misleading.

The maximum likelihood principle is based on the intuitively appealing idea of choosing those parameters from which the actually observed sample in most likely to have come. Any particular sample is more likely to have come from one population than from others. Inherent in each likelihood function is a distribution of the error term. Therefore, maximum likelihood

regression analysis does not require a normal distribution of the error term, neither does it require that the relationships between the parameters and the error term be linear. However it does require that the researcher know and specify the distribution of the error term. 58:74

There are two types of models that can be estimated using maximum likelihood methods which are of great potential value in criminal justice. The first is a Tobit model, used when the dependent variable is constrained at a maximum or minimum value. This model can be used in criminal justice research; e.g., consider the sentencing of judges. Assume that the dependent variable is the length of the prison sentence.

If the judge is allowed to place people on probation, there will be many observations at zero. A Tobit model can simultaneously estimate the probability a person will be sentenced to prison and the expected sentence given that the person is imprisoned and not placed on parole.

In the case of recidivism, a researcher may be concerned with the probability that a parolee will not complete his term without being returned and, given that he is returned, how long it takes. The model can be used to estimate the probability a person will return to crime and the number of crimes he will commit if he does. The Tobit model cannot, however, employ a binary dependent variable. The proper maximum likelihood method to use with a binary dependent is called a multivariate logistic model of the following form:

$$Y = 1 / (1 + \text{EXP} - (a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots))$$

The multivariate logistic model assumes that the independent variables interact; i.e., the effect of one explanatory variable depends upon

the size of the other explanatory variables. The logistic model is a special case of a more general model which can have any number of mutually exclusive outcomes whose probabilities sum to one. A multivariate logistic model gives estimated probabilities of each of the three possible outcomes. It does not require any ordering or scaling of the outcomes. In the case of two possible outcomes; i.e., a binary dependent, the distribution of the error term used by the multivariate logistic model is a discrete distribution. 58:74-75

One of the most common uses (or abuses) of linear regression analysis in criminal justice is estimating the probability a person will not return to prison after his release. 58:75

If we define principal variables of analysis as:

- P: Probability of a court outcome desirable to the police
- D: Delay in apprehension; number of days from offense to arrest
- W: Probability that witness will not cause the case to drop out of the court
- N: Number of witnesses cited at the time of screening by the prosecutor
- E: Whether tangible evidence, such as weapons or stolen property, was recovered by the police
- X: Vector of variables that describe the primary arresting officer
- C: Vector of control variables

We then can write:

$$P = P(D, W, N, E, X, C)$$

The likelihood of a court outcome desirable to the police is determined by the length of delay in apprehending the suspect, the quality of testimonial evidence, the existence of tangible evidence, certain characteristics of the arresting officer, and other factors. 53:60-62
to be determined empirically.

A model based on the individual police officer has these performance measures: 53:62-63

1. Number of cases accepted by the prosecutor.
2. Number of cases in which the defendant was convicted.
3. Proportion of arrests made by this officer that were accepted for prosecution.
4. Proportion of arrests that ended in conviction.
5. Total seriousness score for the crime or the defendants, or both, of all cases accepted by the prosecutor.
6. Total seriousness score of all cases that terminated in conviction.

Blumstein and Larson address recidivism using a simplified Markov model, wherein each offender, after committing a crime, is apprehended with a probability P_A and, if apprehended, incarcerated with a probability P_I . Assuming that the actions of the CJS have an effect upon the offender's future behavior, the probabilities of committing at least one more crime are P_{R1} , P_{R2} , and P_{R3} , depending on whether the offender was:

1. Not apprehended
2. Apprehended but not incarcerated
3. Apprehended and incarcerated

10:214-215

They define three different types of recidivism:

1. P(C/C) = Probability that an offender commits at least one more crime/he has just committed a crime.
2. P(A/A) = Probability that an offender is arrested at least once more/he has just been arrested.
3. P(I/I) = Probability that an offender is incarcerated at least once more/he has just been incarcerated.

$$\text{Then } P(C/C) = \frac{P_{R1} (1-P)_A + P_{A R2} P (1-P)_I + P_{A I R3} P P}{1 - (1-P)_A P}$$

$$P(A/A) = \frac{P_{R2 A} P (1-P)_I + P_{R3 A I} P P}{1 - (1-P)_A P}$$

$$P(I/I) = \frac{P_{R3 A I} P P}{1 - (1-P)_A P - P (1-P)_A P}$$

Using this same model we can compute the average number of career crimes, arrests, and incarcerations.

10:216

- n_{C/C} = Mean number of crimes committed in criminal career/at least one crime is committed
- n_{A/C} = Mean number of arrests in criminal career/at least one crime is committed
- n_{I/C} = Mean number of incarcerations in criminal career/at least one crime is committed

The more complex feedback model, which includes effects of aging and crime-type switching, demonstrates that a rehabilitation program

that causes only a small but measurable reduction in recidivism probability could well have a substantial effect in reducing the total number of crimes committed. A complete description of the recidivism mechanism requires not only the values of the crime repetition probabilities but also an estimate of the time until recidivism occurs.

10:218

The average number of crimes committed during the course of a criminal career is 1/(1-P). If the average time between crimes is T years, then the average time between the first and last crime (or the average length of a criminal career) is [1/(1-P)-1] T years. The average number of crimes committed per year during the course of the individual's criminal career is 1/T crimes per year. This is the "contribution" of one committing individual to the crime rate during a year.

10:218

Program crime is a vector simulation language for the administration of justice system which was motivated by the need to model the adult felon administration of justice system. Larson specified the preliminary structure of the simulation language, including the five arithmetic sub-routine calls, the three vector combination subroutine calls, and the general method of the associated bookkeeping procedures. McBride modified most of the earlier routines and added sensitivity analysis and feedback capabilities. Program Crime is a computerized implementation of the overall generic CJS model. The program is a set of subroutines that provides the user with a block diagramming language; i.e., there is a one-to-one correspondence between a block on a flow chart and a computer instruction. This capability makes it possible to revise the model or even

change it completely with a minimum of effort and without interfering with the rest of the program. Given a reported crime rate for a jurisdiction in a particular year, a set of policy-related probabilities, and a set of certain aggregated fixed and variable costs based on data for several previous years, the program will generate costs and flows for that year at each stage of the CJS. The effect of changing various policies, flows, or costs can be measured by using the sensitivity analysis routines. In this way, the critical points in the system can be quickly isolated. Incremental flows and other quantities can be computed for each additional person inserted into the system at a particular stage and charged with a particular crime.

10:223-224

The cost and flow breakdown and the sensitivity analysis program both operate on an "open-loop" structure; i.e., the input is the total reported crime rate and the offenders who "drop out" of the system who are subsequently rearrested are not specifically taken into account. The closed-loop feedback model, on the other hand, is based on offenders (those arrested), not on crimes. The input is "new offenders" and, using probabilities of rearrest, a portion of these re-enter the system, contributing to the number of total offenders. A sensitivity analysis can be made on this closed-loop model to find, for example, those factors that would be most helpful in reducing recidivism. "Career costs" can also be found by linking the closed-loop and open-loop models. One new offender of a particular age and initial crime type is injected into the feedback model. The resultant total number of arrests is then used to compute an input to the open-loop

model, which then computes the various costs resulting from the single new offender.

10:224

10:224-225

There are five basic assumptions inherent in the overall system model:

1. Offender flows and costs at various points in the system can be allocated by crime type.
2. Probabilities can be assigned to describe the likelihood of various outcomes at the decision points in the system.
3. Changing a policy at a particular decision point does not affect the policies at other decision points.
4. The model is "steady state" with respect to policies and costs.
5. The system is linear; i.e., all costs and flows are linearly related to the input crime rate.

The sensitivity analysis routines compute two types of "derivatives" that measure the response of the system to incremental changes in offender flows at various points in the system. By using these derivatives, one can find both the expected cost per offender and the percentage increase in the number of offenders inserted at any stage in the CJS.

10:225

10:227.

The data requirements include three types of information:

1. Input crime rate for the jurisdiction under consideration.
2. Cost, workload, and manpower data associated with the various subsystems.
3. Probabilities of various outcomes at a particular decision point in the system with all data given by crime type.

The closed-loop feedback model was devised in order to study the effects of policy changes in the criminal justice system on recidivism. The feedback model includes age as well as crime type as an independent variable. 10:237-238

Shinnar and Shinnar developed a model to predict the incapacitative function of prison sentences. Consider a criminal population of N criminals. N is considered to be a constant in time, and an equal number of criminals enter and leave the system. The average length of criminal career is T, and individual career lengths are exponentially distributed. Further assume that during his career a criminal commits crimes in accordance with a Poisson process with rate λ. This rate is uniform for all identified criminals and applies only during the time at which the criminal is free. Each time he commits a crime he may be arrested and convicted with a probability q, or he may be continue undetected with a probability 1-q. If convicted he may be committed to prison with a probability J. The probability of going to prison, having committed a crime, is therefore qJ. For each commitment he stays in prison for a time S. S is not the sentence length, but the total time he stays in prison for each conviction. For each crime he is on the average detained for an average time qJS. Assume the number of criminals is unaffected by the crime policy. The number of crimes committed by each individual during his criminal career is x. 68:586-587
If the CJS does not intervene, the expected value of x is E(x) = λT.

If during his life the criminal is convicted of a crime and sent to jail then E(x) is reduced. The time a recidivist criminal spends out

of prison is on the average 1/λqJ; inversely proportional to the frequency with which he commits crime and the probability of being sentenced to prison having committed a crime. The expected fraction of the time he is free to commit crimes is therefore

$$\frac{\frac{1}{\lambda q J}}{\frac{1}{\lambda q J} + s} = \frac{\text{Average time between commitments}}{\text{Average time between commitments} + \text{Average time in prison}}$$

If there were no prisons (or the criminal is never caught) the total number of crimes committed by a criminal has an expected value of

$$E(x) = \lambda T = (\text{length of career}) \times (\text{number of crimes per year})$$

If he is incapacitated during a fraction of his career, E(x) is reduced.

The ratio of the number of crimes committed by him under a given policy, 70:587-588
to his expected number for zero incapacitation:

$$\frac{E(x) \text{ at a given } qJS}{E(x) \text{ of } qJS \text{ equal to zero}} = \frac{\lambda T / (1 + \lambda q JS)}{\lambda T} = \frac{1}{1 + \lambda q JS}$$

We can also express the effectiveness of the policy as the number of crimes that are prevented due to the CJS, given by

$$\text{Effective reduction} = 1 - \frac{1}{1 + \lambda q JS}$$

If data for q and J are unavailable we can also use q_A, the probability of being arrested for a crime, and J_A, the corresponding probability of going to prison, having been arrested, since q_AJ_A is equal to qJ S. 68:588-589

For any such policy to be effective at all, J_A must be close to unity.

We really don't know λ, but we can measure qJS rather accurately for each type or class of crime. It is simply the number of criminals

confined in all prisons for a specific crime divided by the number of crimes per year. We can also get estimates for q, from which we estimate λ .
68:589

A somewhat unexpected result of the model shows how sensitive incapacitation effects are to changes in CJS performance. If λqJS is small, increasing qJS will initially have significant effects on prison population. Incapacitation can have major effects on the crime rate, according to Shinnar and Shinnar.
68:590

They attribute the majority of unsolved crimes to criminals who are convicted at least once. The fraction of criminals in the model which are never convicted is $[1 - e^{-qXT}]$. This is crucial, since 70 percent of all safety crimes in the United States are never solved, and in New York City this fraction is higher. If most crimes are committed by criminals who are never caught, then no incapacitative policy will work until there are means to catch them at least once. The first and most important assumption in the model was that most crimes are committed by recidivists, since most crimes that are solved by either arrest or conviction are committed by recidivists.
68:592

Extensive studies showed that above 90 percent of the arrests investigated were based on solid evidence and depending on the nature of the crime, 65-80 percent of arrests lead to convictions related to the arrest. The parameter qJS, the average time actually spent in prison for each crime committed, according to Shinnar and Shinnar, is the true index

of the effectiveness of the criminal justice system. The effectiveness of the criminal justice system is given by
68:603

$$1 - \frac{1}{1 + \lambda qJS}$$

A total crime increase can be partly due to an increased number of criminals, especially juveniles. In the United States, the number of persons in the crime prone age (15-30) increased by 40 percent between 1960-1970 but the total number of prisoners remained constant.
68:604-605

At present, the chance of a mugger being arrested for a given mugging is about 12 percent, and his chances of imprisonment after being arrested about 10 percent. His total chance of going to prison is only 0.012 or 1.2 percent. A qJS of 0.5 means that for each crime the criminal spends, on the average, half a year in jail. We can increase the value of qJS in different ways, either we increase the length of stay or increase the probability of a criminal getting convicted and sent to prison.
68:605-606

Shinnar and Shinnar conclude that one of the main effects of prison is simply temporary incapacitation and that any factor that decreases the chance of a criminal to get convicted has a direct effect on increasing crime rate in an almost proportional way.
68:607

Bottoms and Nilsson cited the Chicago Police Department's development of its motorized beat assignment policy on the basis of O.W. Wilson's weighted workload scheme, to have each motorized beat unit evenly divide its time between the response function and the preventive patrol function. They advocate queuing theory to estimate the

number of units required in the response force at any time of day, computer simulation to test the effect of alternative mission assignment policies on availability, and concepts of search theory, originally developed for finding submarines, and computergraphics to arrive at assignments of preventive patrol units. 12:24-26

In the problem of how to measure the effectiveness of preventive patrol activities, there is a analogy between the problem of estimating search effectiveness in antisubmarine warfare and the effectiveness of preventive patrol. They propose as a measure:

$$Q = \frac{\frac{C}{N}}{\frac{T}{A}}$$

where:

C = Number of on-view arrests by preventive patrol in given area

A = Area patrolled

T = Total time spent by patrol units in Area A

N = Amount of reported crime during T in A

to represent the arrest rate per unit area search time. 12:26

An article in the Australia and New Zealand Journal of Criminology cited a weighting method now being used by the Classification Board at Wi Tako, although the knowledge that prediction will be wrong in approximately one-third of the cases reviewed makes the Board Members less dependent on the method than they might have been had the analysis been more successful. The variables and weights used in the discriminant function analysis were: 55:26-30

1. Type of offense
2. Drinking at time of offense
3. Age at conviction
4. Previous court appearances
5. Nationality
6. Marital status
7. Jobs in previous 12 months

According to Avi-Itzhak and Shinnar the expected number of reported offenses during a criminal career, $E(X)$, is given by

$$E(X) = \frac{E(D)}{c} = \frac{1-P_c}{q} \phi = \frac{1-e^{-\lambda q}}{q} \phi_c$$

In general q , the probability of conviction after an offense, is the measure of effectiveness. However, $1-P = \frac{\lambda q}{1 + \lambda q}$ and in a system where λ is uncontrollable, the term $1-P$ stands for effectiveness ($1-P$ is increasing with q). 3:215-216

To illustrate system sensitivity of a population variable such as the steady-state number of inmates in a state correctional institution, for the crime of robbery, $W(2)$, the associated derivative $[\frac{\partial W(2)}{\partial N(2)}]$, for the California model is computed to be 0.23. This quantity has two possible interpretations, according to Blumstein and Larson:

1. For every additional robbery reported, on the average, an additional 0.23 man-years are spent in prison by an individual found guilty of robbery.

2. For every additional robbery reported, the steady-state prison population is increased by 0.23 inmates.

The first interpretation is given in terms of time spent in prison, 9:196 the second in terms of the population in prison.

"A variation of the concept of incremental flows (or costs) is the concept of elasticity. Here, the derivative of the incremental flow is modified to a percentage derivative... As an example, the elasticity of the number of detective man-hours for auto thefts with respect to the number of auto theft arrests is computed to be:

$$\frac{N_d(G)}{W_d(G)} \cdot \frac{\partial W_d(G)}{\partial N_d(G)} = \frac{\partial W_d(G)/W_d(G)}{\partial N_d(G)/N_d(G)} = 0.62 \text{ (for the California model)}$$

This elasticity can be interpreted to mean that for a small increase of X percent in the number of auto theft arrests per year...the 9:203 detective workload for auto thefts would increase by 0.62 X%."

To illustrate system sensitivity regarding adults found guilty of robbery, N (2), the associated incremental flow is $[\frac{\partial N(2)}{\partial N(2)}]$ the first derivative of the number of guilty robbery defendants with respect to the number of reported robberies. In the California model this is calculated to be 0.08. Two alternatives interpretations could 9:195 be given to this number:

1. For each additional robbery reported there would be, on the average, an additional 0.08 adult defendant found guilty of robbery...
2. In a randomly selected reported robbery, the probability that the robber would be found guilty of that robbery is 0.08.

Nagel and Neef wrote, "operations research is the study of the application of mathematical techniques to the choosing among various alternatives that decision or decisions that will maximize some quantitatively 50:7 measured goal." Three common methods are cited:

1. Linear programming
2. Inventory modeling
3. Decision theory

Linear programming or linear optimizing can be defined as a geometric or algebraic procedure whereby one finds the optimum allocation of something between two or more alternatives in light of certain goals and given constraints or conditions.

Inventory modeling can be defined as a geometric or algebraic procedure whereby one finds the optimum quantity or optimum inventory in a situation where doing either too much or too little will result in excessive costs or unduly low benefits.

Decision theory can be defined as a geometric or algebraic procedure whereby one chooses among alternatives in order to maximize given goals 50:7 in light of probabilistic or uncertain events.

Nagel and Neef illustrate the inventory lot size method as a system 50:48 of three equations:

1. A rising cost equation of the form $TCP = a(P)^b$, where P = the degree of effort expended pursuing a policy; TCP = the total cost of pursuing the policy; a = the amount of TCP incurred

if one unit of effort is expended in pursuing the policy $b_1 = a_1$ positive number to which P is raised to show the degree of increasing TCP costs from additional units of P.

2. A falling cost equation (with regard to the same policy problem) of the form $TCQ = a_2 (P)^{b_2}$, where TCQ = the total cost of pursuing the opposite of the policy; a_2 = amount of TCQ incurred if one unit of effort is expended in pursuing the policy; and b_2 = a negative number to which P is raised to show the degree of falling TCQ costs from additional units of P.
3. A total cost equation that represents the sum of the left sides of the first two equations and has the form $TC = TCP + TCQ$, where TC = total costs of pursuing a policy which incurs relatively high costs if too little or too much of the policy is pursued. The object of the model is to find the value of P where TC is a minimum.

The costs of not incarcerating violators who violate are:

1. Crime costs
2. Later arrest costs

Whereas, the costs of incarcerating a violator who would not commit additional crime are:
50:53-54

1. Jail costs
2. Court costs
3. Lost GNP
4. Bitterness costs

Sellin and Wolfgang developed a seriousness index to rank the gravity of offenses. They indicated their "major" purposes for scaling offenses were:

1. To select from multidimensional features of delinquency a single dimension, taking into account the relative gravity or seriousness of delinquent acts.
2. To produce an empirical, objectively ascertained set of components of delinquency that would be examined by socially significant groups whose evaluations could be used as a basis for scoring.
3. To arrive at a system of weights for delinquency events for use in the construction of an index.

They cited Fechner's Law: "The psychological measure of a physical event is equal to the logarithm of the physical measure (multiplied by a constant of proportionality)...for intensive physical variables, the psychological measure is simply the logarithm of the physical variable."
67:236-238

In the seriousness measurement, the freedom in the range of possible responses available by the magnitude estimation technique provides intrinsically more information about the raters' judgments than the severely limited categories. They also noted that,
67:273-342

1. While a rating of seriousness does not directly yield information on reportability, there is some connection.
2. A measure may be considered valid if there is a high correlation between the predictions derived from it and what actually occurs.
3. By interpreting the seriousness index as a ratio scale, with zero indicating the absence of an offense, we are employing a strong

method of analysis which becomes a powerful tool for examining empirical relations.

As an example of seriousness index utility, PROMIS (Prosecutor's Management Information System) in Washington, D.C., utilized a measure of the urgency of a case for prosecution, of the following form:

$$U = p_1 w_1 SW + p_2 w_2 BE,$$

where

U = Judged urgency of the case for prosecution,

p = Subjective probability of winning the case,

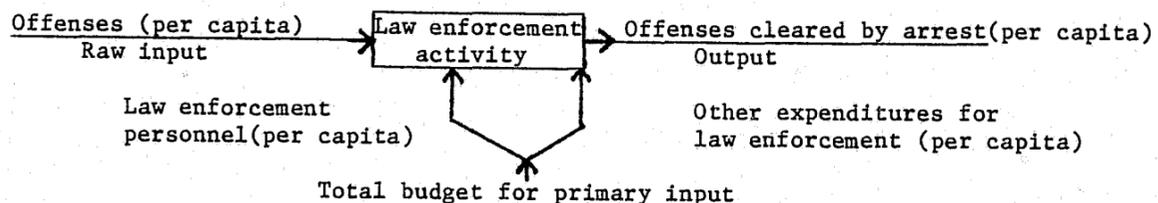
SW = Seriousness of the offense on the Sellin-Wolfgang Scale,

BE = Base expectancy, a measure of the likelihood of the offender's

recidivism, based on the work of Gottfredson, et al.,

w_1, w_2 = weights appropriate to seriousness and base expectancy, respectively.

Glaser sees law enforcement as a production process:



Then the production function may be represented by:

$$C_i ; q_i(t, OF_i, EMP_i, OTE_i), (i = 1, 4 \text{ crimes})$$

where C represents crime cleared by arrest

OF = Offenses reported to police

EMP = Law enforcement employees

OTE = other (nonpayroll) law enforcement expenditures, in real terms (adjusted for inflation)

As is typical of almost all production functions, we expect diminishing returns. 32:1070-1071

According to Glaser: 32:1076

1. Increased expenditures for law enforcement tend to increase clearance ratios.
2. Increased clearance ratios may be regarded as increasing the probability of arrest and/or incarceration for the offenses.
3. The higher the probability of incarceration, the lower the net expected benefits of an offense.

Rand cites a model for incapacitation effects: 18:35

$$A/P = \frac{1}{1 + \lambda(qJS)}$$

According to Hoods and Sparks, there are two strategies to assess "dark figure" or hidden crime: 38A:11-45

1. Question general population about criminal acts they've committed.
2. Victimization surveys.

According to Barnett, Larson, and Odoni, 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 by appropriate procedure which may well vary from crime to crime. 4:3-10

The FBI's Uniform Crime Reports show tabulations to indicate the probable extent, fluctuation, and distribution of crime for the United States as a whole; geographic divisions, individual states, standard metropolitan statistical areas, and cities, towns, and counties. The measure used is a Crime Index consisting of seven important offenses which are counted as they become known to the law enforcement agencies. Crime classifications used in the index are:

75:v-189

1. Murder and nonnegligent manslaughter
2. Forcible rape
3. Robbery
4. Aggravated assault
5. Burglary-breaking or entering
6. Larceny-theft
7. Motor vehicle theft

The total number of criminal acts that occur is unknown, but those that are reported to law enforcement provide the first means of a count. Not all crimes come readily to the attention of law enforcement; not all crimes are of sufficient importance to be significant in an index; and not all important crimes occur with enough regularity to be meaningful in an index. With these considerations in mind, the above crimes were selected as a group to furnish an abbreviated and convenient measure of the crime problem.

75:48

All communities, metropolitan areas, and states are affected to a greater or lesser degree by the element of transient population.

This factor is not accounted for in crime rates since no reliable estimates by state are available nationwide. Law enforcement's performance in clearing crimes by arrest is presented by population group and geographic division. National averages are also shown indicating the type and value of the property stolen, by offense and type, and value recovered by police investigation.

75:48

The 1975 Total Crime Index shows 11,256,566 for the total United States, which is the number of index crimes known, or reported to police. The 1975 Uniform Crime Reports shows 32 percent of reported arrests for all crimes were of persons 18 years old or less.

75:49-189

32:686

Glaser enumerates three factors that bear on reporting crimes:

1. Insurance coverage
2. Relationship (victim/offender)
3. Citizen attitude toward police

In studying habitual criminals regarding self-reported crime, Rand found there may be a relationship between instances of personal violence and criminal violence and there may be found to be a "predictor of dangerousness to society." They noted two types of habitual offenders:

18:94-97

1. Intensive - prone to avoid arrest
2. Intermittant - prone to arrest

Many other efforts have touched on criminal justice issues including productivity and some variables upon which productivity depends. Included is the cohort study by Figlio, Sellin, and Wolfgang who indicated if one needs to know the probability of any child born at any given time becoming

a delinquent sometime during his life... Consider the number of children born each year - numbers which may conveniently be referred to as "generations" - and count year by year the individuals of each generation who are convicted in the courts for first offenses. When all the members of a particular generation are dead, it will then be possible to express the probability as a ratio of the total number who were convicted at least once to the total number of the generation at birth. 27:6

They cite a 1960 work by Leslie T. Wilkins entitled, Delinquent Generations, which examined and tested the theory that children born in certain years (such as war-time) are more likely to commit offenses, and concluded that, "children who reached their fifth year of age during the war (WW-2) were most crime-prone and that all born during the war had higher crime rates than expected," (but the study was widely criticized). 27:12

In the Philadelphia cohort, the rate of delinquency of birth cohort was 349.4 per 1,000 but 1862 (54 percent) committed more than one offense while only 1,613 (46 percent) were one-time offenders. They found recidivists are more likely to be nonwhites, in the lower SES, (socio-economic status) have lower IQ scores, fewer school years completed, and lower achievement levels than one-time delinquents. Nearly nine times as many index offenses were committed by recidivists (2,935) as by one-time delinquents (330). 27:65-71

In using an "offense weight" or "relative seriousness of offense" index, the mean seriousness score per offense increases as the number of offenses

committed per person increases. Onetime offenders exhibit lower seriousness scores than do recidivists. 27:74

They were also concerned with the extent to which specialization may exist in the offense histories and found knowledge of the immediately prior offense type (k - 1st) does aid in the prediction of the kth type in that there is some tendency to repeat the same type offenses. This inclination, except for theft offenses, is not very strong. Knowledge of the number and type of offenses prior to the k - 1st gives us no aid in predicting the type of the next offense. They also note that two factors, seriousness of the offense and severity of disposition, are associated with a substantial proportion of recidivism. Analysis suggests that the relationship between the average delinquent seriousness scores and background characteristics is curvilinear. 27:188-272

One of the best-known studies of delinquency, "Delinquency in a Birth Cohort," concluded in part that with the commission of each additional offense, the seriousness scores for nonindex, theft, damage, and combination offenses change negligibly, while injury seriousness scores advance dramatically at each offense rank number. Also the mean intervals follow the general expression, $\log Y = a + b \log X$, for all offense types. In addition, the mean times between offenses are similar enough for all offense times that one function, $\log Y = 1.4243 - 0.8052 \log X$, suffices to represent any type. Offense histories are compressed over a rather short period, regardless of offense type. In the same study, yearly cumulative probabilities for index offenses are presented.

Although the logistic curve [$Y = \frac{k}{1 + e^{-a - bx}}$] and the third degree curve fit the data equally well, the logistic expectancies were plotted in order to assess the extent to which the predicted values correspond to the adult data that are to be collected later.

Data used in seriousness of the offense can be extracted from the following sources:

1. Sellin - Wolfgang index of crime seriousness.
2. Maximum sentence associated with charges.
3. Criminal history of defendant.
 - a. Number of known prior arrests.
 - b. Whether arrested within past 5 years.

Some authors have commented directly on police productivity measurement.

Hatry wrote on productivity measurements for the police crime control function and cited five currently available measures:

1. Population served per police employee and per dollar.
2. Crime rates and changes in crime rates for reported crimes.
3. Clearance rates of reported crimes.
4. Arrests per police department employee and per dollar.
5. Clearance per police department employee and per dollar.

According to Barnett, Larson, and Odoni, in measuring the outputs of the CJS, four categories of measures are identified: (1) input measures, (2) intermediate measures of output, (3) crime-related measures of output and (4) "true" measures of output. Regarding inputs, it is feasible to collect information about the manpower and expenditure costs of the

CJS to practically any desirable level of detail. Regarding outputs, one can theoretically describe measures of the true effectiveness of the CJS.

Intermediate measures of output include a large class of indicators of performance, efficiency, and productivity in the CJS, connected to the true intended outputs of the CJS only through a series of logical inductions. 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. 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 allocations of resources, this type of analysis is both doable and believable.

A brief study reported by Blumstein in 1969, on a cost-effectiveness analysis in the allocation of police resources, as an example, 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. In this example, the measure of effectiveness is an intermediate output measure (response time) and the measure of cost an input measure (dollars invested).

According to Barnett, et al., the methodological foundation is already available and the data base can be created for successful applications of cost-effectiveness analyses of the type described, to the CJS,

as an aid to making decisions on how to allocate resources among the different subsystems of the CJS. 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. A detailed breakdown of the CJS into a number of interconnected constituent parts makes this model an extremely convenient tool.
4:2-25--2-26

Fisk and Winnie discuss the current status of output measurement in the United States, particularly at the local government level, with emphasis in their paper on the growing use of quantity and quality output measures by local government. The author recognizes the need to develop both the measurement and the measurement strategy, and the need to select a measurement strategy that is affordable. Quality is defined as both impact and what Hatry calls effectiveness. There is a distinction between objective-oriented measures and impact-oriented measures.
20:128-129

Mushkin and Cotton also studied the idea of characterizing the performance of public agencies through use of volume and quality indicators. They simply list a number of what they consider to be volume and quality indicators suitable for analyzing and evaluating public expenditures. The general methodology of the Muskin and Cotton approach, called PPBS, seems to lie in the maximization of the volume of output per dollar expended while maintaining the quality of output within certain limits.
20:129

Hirsh and Riccio 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. Several ratios are offered as indicators of productivity, quality of arrests, efficiency, and effectiveness. The need for applying multiple measures is recognized as a means for obtaining a more detailed picture of organizational performance.
20:126-127

Larson cites a lengthy study by Shoup and Mehay which has attempted to demonstrate the merits of the program budgeting system through application to the case of police services in the Los Angeles area.

They advocate adoption of a cost-benefit (or cost-effectiveness) approach to the allocations of police resources. He also notes a paper by Blumstein and says an extensive amount of work has been done on multi-variate regression analyses that attempt to identify statistical relationships among police inputs, crime statistics, and a host of environmental variables.
4:2-11

Lind notes performance measurements can take several major research approaches, each of which 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 Lind expects to have the widest application are: (1) Surveys designed to sample from among populations to learn how traits of interest are distributed, such as interviews, questionnaires or tests which are pretested, known for their reliability, and also known to be valid. (2) Panels or groups 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. 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. (3) Observations on organizations or groups including participant observation, sampling of members with questionnaires or interviews, the use of previously quantified data on organizational activities and the development of new activity measures. (4) Career studies to examine individuals or groups, over a course of time, which can utilize survey methods, participant observation, biographies, diaries, interviews with the subjects, interviews with those who know the subject, or data from institutions in contact with the subject. (5) Personality assessment among individuals 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. (6) 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. (7) Experimental situations to test the outcome of intervention on a given behavior sample under highly controlled conditions. (8) Public records of arrest, legislation, appropriations, and the like can be used retrospectively to evaluate impact. Currently, public records constitute the bulk of the measures of the crime problem, the operations of the justice system, and inferences about impact. These records are likely to be strongly biased by random and nonrandom errors and their use for contemporary work should be approached with caution. (9) Methodological studies concentrating on methods and statistics themselves. Sellin and Wolfgang's method for rating crime severity according to seriousness is a most careful example. (10) Organizational 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, or accident, and performance measures for personnel such as arrests or citations among police. (11) Covert and/or deceptive measures are used when one does not wish a subject or group to be aware that observations are being made. (12) Inventions or innovative responses which presume either dissatisfaction with existing methods or the growth of new ideas.

Maltz noted innovations have been and are being tried in the State and local agencies comprising the CJS, in all phases of their activities,

with many of the programs directed specifically toward the control of crime. Each of these innovations requires evaluation. Evaluation has been defined as the process of determining the value or amount of success in achieving a predetermined objective, including at least the following steps:

1. Formulation of the objective.
2. Identification of the proper criteria to be used in measuring success.
3. Determination and explanation of the degree of success.
4. Recommendations for further program activity.

The potential value of each program will not be realized if it is not evaluated in order to make the best possible decision. ^{46:4-5}

Maltz also wrote that if the program is directed at specific types of crime, the predicted number of such crimes during the study period should be determined on the basis of past data. Other statistics, such as the standard deviation of this predicted number, should also be calculated. Each program can have its own internal measures of effectiveness, based on the logical elements of which it is constituted. He goes on to say, the crime rate, the number of a specified type of crime committed per resident in a specified time period, is normally considered to be a measure of deterrence. If the crime rate decreases, it is presumed that potential offenders have modified their behavior and have committed fewer crimes, based on the supposition that the program has made the target crimes unattractive: by increasing the

actual risk of apprehension; by increasing the perceived risk of apprehension; by reducing the expected return from the crime (or the perceived return); or by making alternative forms of behavior more attractive than the target group of offenses. Most crime control programs are police-oriented and concentrate on the risk-related aspects of deterrence. Victim-oriented programs focus on reducing the expected return. ^{46:16-34}

Clearance rate is normally considered to be a measure of the ability of police to solve crimes. A cleared crime is one in which the police have identified the offender and have sufficient evidence to arrest. Maltz cites Greenwood who has identified a measure of effectiveness for detectives that appears to be more useful than clearance rate, the "Detective Arrest Index," based on many of the same considerations as clearance rate, but more specific and minimizing some of the problems. Maltz says "clearance rate can be a useful measure for determining the effectiveness of crime control programs." ^{46:34-38}

Another measure of effectiveness often used as a determinant of crime control effectiveness is the arrest rate per police officer or per resident which is not related to the total number of offenses. ^{46:38}

The "crime seriousness index" proposed by Sellin and Wolfgang included some of the major disutilities of crimes. Crimes are weighted according to the degree and nature of injury to the victims: whether they were intimidated and the nature of the intimidation, whether premises were forcibly entered, and the kind and value of property stolen, by requesting

a sample of people to estimate the relative seriousness of various crimes. All of the factors used to determine the weights are (or should be) included in offense reports, according to Maltz, and he says "perhaps a better index of the relative value of property loss to the victim would be the value of loss in relation to the amount of the individual's discretionary income." The incorporation of a modified form of the index by a police department, as the permanent legacy of an evaluation, would be a significant step toward improving crime data.
46:38-41

The perceived risk of crime is greater than the actual risk of crime, and perceived risk does not seem to be correlated with the actual crime rate. Public opinion surveys with regard to perceptions about crime and safety have been made frequently. Almost 200 surveys of crime-related topics have been identified by Biderman, et al. One study suggests that perceived and actual risk of crime are correlated, but public concern about crime is not correlated with actual risk. A side benefit of evaluation would be an estimate of the business losses suffered due to crime, as part of the total cost of crime.
46:41-42

McGuire says important qualitative inferences can be drawn concerning plausible relations between costs, incapacitation benefits, and other benefits of incarceration. Incarceration removes criminally productive individuals from contact with free society, and the incapacitation benefit is the value of crimes thus avoided. For the correctional system or institution, it is the summation of these benefits for the

respective confined population. For the law enforcement agency, it is the sum of these benefits for those individuals the agency was responsible for incarcerating. Incapacitation benefits in any period are inversely related to the magnitude of the displacement effect, a measure of the longrun supply elasticity of criminal activities.
48:13-14
The quantification of incapacitation benefits requires:

1. Projecting hypothetical criminal careers
2. Estimating social losses
3. Incorporating into the analysis estimates of the displacement effect

Regarding the projection of criminal careers, four techniques are available:
48:14

1. Compute, based upon inmate records and characteristics, a historical crimes per year function, projected over the incarceration period in order to determine crimes avoided.
2. Assume that at any point in time offenders fall into two categories: (1) incarcerated and (2) not incarcerated. Assuming that the offense rate of the former would equal that of the latter were they not incarcerated, let the latter's offense rates proxy crimes avoided.
3. From estimates of λ , the expected number of actual crimes per year, compute the estimated number of crimes precluded by incarceration.
4. Use recidivistic criminal activity as an index of crimes which would have not been committed had the individual been incarcerated.

The ratio of the victimization rate to the arrest rate has been proposed as a conversion factor relating actual criminal activities with arrests.
48:17

The primary determinants of variance in displacement effects among offense categories are the degree to which an offense is economically oriented and performable without offense-specific physical or human capital, and whether the criminal activity associated with an offense is controlled by organized crime. Economic motivation and ease of entry are both assumed to flatten the longrun supply curve; i.e., to increase the displacement effect, because the transition from legal to criminal activities is less costly in these circumstances. Control by organized crime is assumed to increase the displacement effect because of the probable ready replacement of incarcerated members of criminal organizations.

48:20

Incapacitation effects, which are criminal activities avoided because individuals are removed from contact with free society, can be approached by computing arrest per month estimates from individual inmate records, converted to institution and system level estimates based on the relevant confined populations in each offense category. If displacement effects are important, then the offense classification structure of the sample confined population suggests that the incapacitative effectiveness of the sample institutions will likely be markedly reduced.

48:22

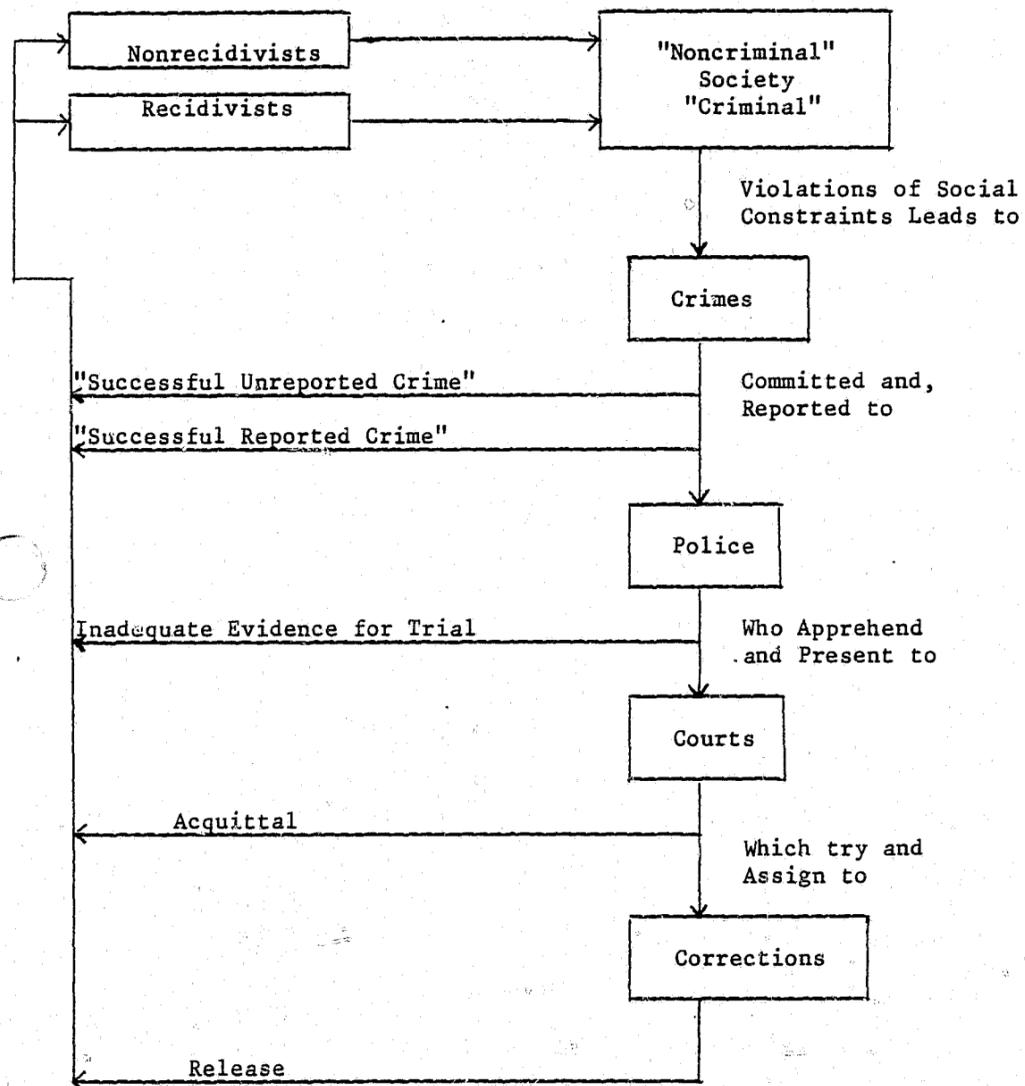
Evaluations of social programs are often thought to be akin to the award of academic grades to school students--a means by which to identify those which are "better." Evaluations should be designed to be of use to decisionmakers facing the following problems:

1. Whether to continue funding a particular program.
2. Whether technical assistance should be provided.
3. Whether funding of a proposed new program appears warranted.

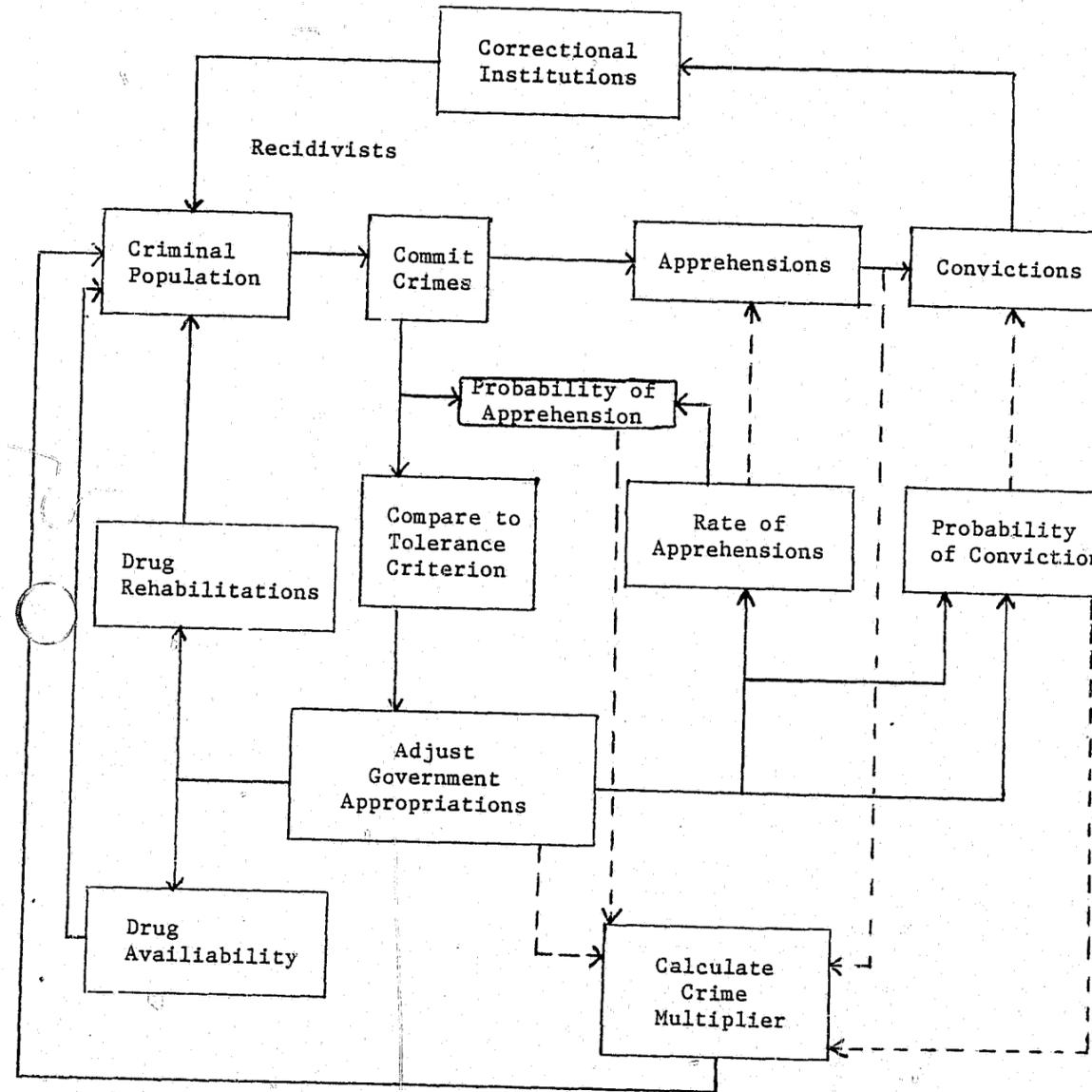
Stewart quotes Daniel Glaser as stating, "often the most effective way to reduce the extent to which people are labeled deviant is not to change their behavior but to change the labeling practices so that they are no longer considered deviant."

71:2-3

Riccio modeled the CJS as:



OVERALL CRIMINAL JUSTICE SYSTEM



CRIME CONTROL SYSTEM

In a study to describe the intractable inmate and to determine the existence of factors predictive of intractability, the intractable inmate is defined as an inmate who presents a chronic disciplinary problem within the prison. Fifty intractable and fifty tractable inmates were selected for study, and the data was analyzed to determine group differences and predictive factors. In comparison with the tractable group, the intractable inmates studied were generally nonwhite, single, not heavy users of alcohol, and they exhibited disciplinary problems before incarceration. They began their criminal history at an earlier age, more often used an alias, had a greater number of police contacts, and once incarcerated, were confined longer. Levy and Meyers concluded that the intractable inmate can be differentiated from the tractable inmate by six variables, and when properly weighted, correctly classified 78 percent of the time. Using stepwise multiple regression analysis of the 22 variables resulted in the selection of six variables which as a set, provide the best prediction for the sample studied. These six variables yield a multiple regression coefficient of .630 ($p < .001$). A coefficient of this size is not generally suitable for individual predictions.

Deutsch sees the development of the conceptual basis for performance measurement as:

1. Development of a measurement selecting decision process.
2. Developing new performance measures.
3. Development of a decision process for selecting a measurement strategy.
4. Examination of new measurement strategies.

5. Developing a methodology for design of measurement processes.
6. Examination of new measurement strategies.

He sees the identification and selection of organizational objectives as:

1. Development of decision process for selecting optimal objectives.
2. Determination of optimal response.
3. Determination of the induced objectives of the CJS and its component organizations.
4. Design of procedures for detecting inconsistency among law enforcement objectives.

Deutsch sees the determination of overall organizational performance as:

1. Design of functional models for defining the effectiveness function.
2. Design of empirical models for defining the effectiveness function.
3. Identifying external indicators of overall performance.
4. Identification of overall activity conflicts.

His structure for determination of overall CJS performance is:

1. Design of functional models for effectiveness function.
2. Design of empirical models for the effectiveness function.
3. Identifying regional and national external indicators of overall CJS performance.
4. Identification of inconsistency of objectives.
5. Analysis and quantification of costs for objective inconsistencies.

According to Deutsch, then, the 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.

Once the individual activities of each component organization within the CJS have been identified and typified as to the nature of the activity, a cross-organizational classification of activities can be performed to group all similar CJS activities. An activities matrix form can be used by arranging all the CJS activities along a linear scale. By classifying them by the characteristics of the activity alone, it is possible to arrange the activities into groupings of similar activities. A second matrix that can be utilized is the matrix showing the relationships between the component organizations and the various activities. 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 and care must be taken when proceeding to choose a resource-oriented measure of effectiveness.

A wide variety of approaches for effectiveness measurement are currently available, according to Deutsch. 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.

The first source of information about possible measures of effectiveness is the relatively large body of literature dealing with individual and organizational performance measurement. Behavioral scientists have developed some rather unusual methodologies for evaluating an organization, such as measurement of the resolution of conflict within an organization. There are two distinct types of analysis that can be performed in the field that will yield the bases for measurement approaches.

1. The analysis of the stated objectives of the organization and the activities chosen to reach those objectives.
2. The analysis of observable effects looks not at the objectives and activities, but attempts to identify observable effects of the activities.

Another source of measurement approaches is the thought process of the researcher. 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, ensuring a good coverage of possible measurement approaches. The whole point is to develop something that can serve as a valid indicator of a law enforcement organization's effectiveness.

A good deal of information is needed about a measurement approach to adequately fit the approach to an appropriate application, such as:

1. Data required.
2. Effort required for operation.
3. Nature of the output of the approach.
4. Strengths and limitations of the technique.

The specification of the "most useful applications" 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. 20:112-114

A measurement process can be designed to be self-improving, increasing the efficiency of the overall measurement process, and improving the quality and significance of results obtained. Once a particular measurement approach has been proposed as a viable way to assess the effectiveness of a certain organization's activities, there need be some validation to ensure that the approach yields accurate and reliable results, and that the results are being properly interpreted. Real validation of a measurement approach lies in the accuracy and reliability of the results its application produces. To test for accuracy, the results obtained by the first approach on the original problem can be compared to:

1. Results obtained by a different approach applied to the same activities, or
2. 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. Researchers would prefer to apply a dissimilar validation approach if possible. 20:120-121

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. 20:121

Dollars and services are both commodities, but there is no simple value relationship between any single service by itself and dollar expenditures. However, there may be some merit in setting upper and lower bounds on the economic value of a service to estimate its monetary value. Determining or estimating the effectiveness function is more applicable when there is relative homogeneity among the objectives; when the evaluators determine that an overall measure of effectiveness is needed, and when there is desire to express the results of measurements as a scalar value. 20:48

The most commonly used model for overall organizational effectiveness is an additive model using weighting factors to adjust for the difference in value of achieving the various objectives. A weighting vector w is defined such that for an n-objective organization,

$$w = (w_1, w_2, \dots, w_n),$$

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 relates to what any objective contributes to the overall effectiveness. Weights can be assigned through a number of techniques. 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.
20:48-49

A second type of model is the second and higher order model, which allows for interaction between the various effectiveness scores. This 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.
20:49

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 judgments 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, such as simple linear regression and polynomial curve-fitting methods. An interesting aspect of the search for weighting coefficients is a process by which the judgments by the evaluators of the effectiveness of the organization are combined and analyzed. The most promising

technique is known as the Delphi Method, which combines informed opinion, directed questioning, and feedback of responses to produce a convergence of opinion. 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 . This can be done even if it is decided not to attempt to estimate overall effectiveness directly, but to bracket it in a confidence interval, which might result from the diversity of opinion of the evaluators. A typical interval estimate might be specified by stating that the overall effectiveness is between 70 and 80 percent with a probability of .95.
20:50-52

A second general class of models of organizational effectiveness is the deterministic models, 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 e vector by eliminating redundant measures of effectiveness. 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. Therefore, determining the effectiveness function is a simpler task if there exists relative homogeneity of the objectives. 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. 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. 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, because, just as the component organization is a multiobjective, multidepartment organization, the CJS is a multiobjective, multicomponent entity. 20:52-53

Two basic ideas to be considered when attempting to speak of the goals, objectives and activities (GOA) of the CJS are: 20:68

1. The GOA of the CJS could be considered to be the GOA of the component organizations, conflicts, inconsistencies, and all.
2. A "top-down" concept is used if 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.

Deutsch goes on to say 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, that multiple complementary measures are better than solitary measures, and that there is the tacit recognition of the measurement process. 20:134

Ostrom developed definitions of output and efficiency and demonstrated potential output and efficiency measures and their application. He suggested 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 was 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. Ordinal rankings, although not as precise as exact measures, do provide a sufficient basis for drawing inferences regarding the performance of an evaluated agency, according to Deutsch. 20:130-131

The classification of activities is a prerequisite for the selection of the performance measure, and subsequently the measurement strategy and measurement approach. 20:131

Mantel, et al., failed to classify each and every service by preconceived definition, but 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 weighted heavily on the more "important" packages. The Delphi method was invoked to assign weighting factors to convert the computed utility factors into a quality index. Six weighting factors were used, regardless of the agency, to transform the six utility functions to a quality index. 20:132-133

Larson, et al., made two principal recommendations in 1976: 4:2-12--2-13

1. The Law Enforcement Assistance Administration (LEAA) should assume responsibility for the collection and dissemination of information on the various types of resources utilized by the CJS on an annual basis. The survey should be of a scope similar to that of the

LEAA-Census Bureau annual volume on Expenditure and Employment Data for the CJS. The survey depth should be considerably expanded and seek much more detailed information than that which is obtained currently. LEAA should publish the results of these surveys on an annual basis and in easily comprehensible form. Emphases should be on exhibiting trends in CJS expenditures and in discussing the probable underlying reasons for these trends.

2. LEAA should undertake or support a series of studies to analyze CJS expenditures and manpower data (including the examination of time-series trends for individual locations or for groups of jurisdictions and the performance of cross-sectional comparisons among individual municipalities, states, or regions with varying or similar characteristics).

Larson also lists promising research topics, each concerning one or more aspects of the CJS:
4:2-13--2-15

1. An analysis of the make-up and composition of the well-known momentous increases that police department budgets experienced during the last decade.
2. The major trends in salaries and benefits for CJS personnel and the relationship--if any--between wage gains and the various unionization movements.
3. Internal trends in CJS employment (e.g., changes in the relative proportions of uniformed and civilian employees).

4. 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.
5. The details of the allocation of police manpower among various possible functions (e.g., preventive patrol, response to calls for assistance, investigations, clerical tasks, enforcement of traffic regulations, etc.).
6. 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.
7. The relationship of "size" to CJS expenditures.
8. The budget fractions allocated to capital investments and to current expenditures in different parts of the CJS.

LEAA currently requires project grant applications to include an "evaluation component" and program/project objectives must be planned to meet goals.
26:5

According to Movahedi and Ogles, prediction in criminology, among other fields of social inquiry, is reduced to the statistical forecasting of the behavior of a class of offenders or of an individual offender. It is argued that probability is a mathematical function defined for classes of events or sequences of events in the long run and as such is not applicable to an individual instance. They define subjective probability by saying, "the numerical value of the degree of confidence or partial belief in a proposition can be estimated by the highest odds that the individual would be willing to offer on the truth of the proposition. Thus, if

the odds in favor of Brown's success on parole were 8:2, the subjective probability of Brown's success on parole would be $8/8+2$." 49:177-182

Movahedi and Ogles note that a statement ascribing a probability (in a relative frequency sense) to a single event has a fictitious meaning, but the notion of logical probability seems, on the other hand, to provide a meaningful explication of the probability of a single case. 49:186
For example:

The probability (observed) that offenders with Y and Z characteristics succeed on parole is .72. Brown has Y and Z characteristics. The conclusion that Brown has a .72 probability of success is not part of a valid conclusion and is erroneous but does represent a logical relation, since in absence of other factors, we would bet on success.

51:2-119

According to Nijmegen, et al.,

1. For a theory to be a basis for predictions, it is necessary that it be sufficiently tested. Tests of hypotheses and theories are performed by deriving predictions from them and by comparing observations from reality with the theoretically expected events or outcomes.
2. Prediction of recidivism differs from first offender prediction in the extra information that has become available.
3. The reliability of a prediction instrument is an index of its stability of judgment.
4. Regarding sampling in criminal justice studies, with increasing n, the relative frequencies converge to a value P, which is called

the statistical probability of the event, or as the sample size increases with random selection, the frequencies observed approach each other.

Gibbons studied typology as it related to criminology and stated:

1. In a very general sense, this growing interest in typological studies of causation can be likened to the search for explanations for specific patterns of physical illness rather than for a unitary theory of sickness.
2. Overly aggressive behavior is an important category of deviant conduct.

"The trend toward research evaluation of programs continues... in the past, persons have argued for this program or that one, mainly on emotional grounds rather than in terms of any conclusive evidence that the program accomplishes any significant alteration of behavior"...."as more evidence develops from studies...it will be possible to declare that program X has shown to achieve a success rate" which could be compared with the success rate of program Y. 31:22-295

Clinard and Quinney, in studying typologies as related to criminology said, "distinction can be made between a classification (composed of classes) and a typology (composed of types)."..."a typology...attempts to specify the ways in which the attributes of observable phenomena are empirically connected in the formation of particular types." They enumerated the types of criminal behavior as: 17:10

1. Violent Personal Criminal Behavior
2. Occasional Property Criminal Behavior
3. Public Order Criminal Behavior
4. Conventional Criminal Behavior
5. Political Criminal Behavior
6. Occupational Criminal Behavior
7. Corporate Criminal Behavior
8. Organized Criminal Behavior
9. Professional Criminal Behavior

When the police of a country are called upon to undertake crime prevention activities in several domains, two problems arise: 39:99

1. Priority of the various domains
2. Choice of activities

The police must use the available resources to obtain the greatest possible preventive effect. However, it is difficult to determine what constitutes the "greatest possible preventive effect," and it is equally difficult to define and measure a "preventive effect" for verification and comparison purposes. Then there are psychological or political factors; e.g., need for security, or the political considerations of those who hold the decisionmaking powers. Interpol asked the National Central Bureaus to indicate: 39:99

1. Whether or not there is an officially-compiled and applied list of priorities.
2. Whether any scientific methods have been developed to measure "preventive effect," to compare the preventive effect of different activities.

The most promising method seems to consist of comparing the relevant crime data from two separate areas, one where a specific crime prevention method has been adopted and the other where this method has not been used. The crime prevention activities incumbent upon the police are performed by various services, along with other tasks. Also there may be special services in the field of crime prevention. 39:105

The National Central Bureaus were also asked to describe the organization of the services and personnel performing crime prevention duties, and to discuss the effect this organization has on the actual crime prevention work undertaken by the police. Twenty-seven countries gave information relevant to this question, saying essentially the role of any police officer is basically - or at least to some extent - to prevent crime. 39:106

The Criminal Justice Symposium Focusing on Police Productivity in 1974 noted that the research effort is broken down into three components:

1. A comprehensive analysis of evaluation criteria of urban public safety services, directed toward the understanding of productivity and effectiveness of urban public safety services.
2. Development of a set of analytical and simulation models that should be useful as planning, research, and management tools for urban public safety systems in many cities.
3. An evaluation of the impact of new criteria, methodologies, technologies, and organizational forms on traditional crime-hazard rating schemes.

They also noted that unfortunately the most utilized "measurement" to date has been the Uniform Crime Reports. 59:17-39

One researcher has identified three distinctive police styles or strategies which he labeled:

1. Watchman style
2. Legalistic style
3. Service style

These styles reflect the relative emphasis of the department on citizen complaints for order maintenance, law enforcement, and service calls. ^{46:45}

Maltz notes that the reports and records of police departments comprise one of the primary sources of data for evaluating crime control programs, and the police are notified of the occurrence of most crimes by telephone. ^{46:58}

Data reflecting crimes and arrests in the United States generally have come from the FBI's Uniform Crime Reports.

According to the Comptroller General, prior to September 1975, the FBI had allocated its investigative resources based on the average number of cases handled by a special agent, with little attention paid to the quality, nature, or scope of the cases. Management information was limited and was primarily related to the caseload. The FBI uses accomplishment statistics in budget justifications, congressional testimony, speeches, and informational pamphlets and reports. Its accomplishments are listed in five categories: ^{62:7-27}

1. Convictions
2. Fines
3. Savings
4. Recoveries
5. Fugitive locations

In response to valid criticism, the FBI implemented on October 1, 1977, a new system, the Case Management Information System (CMIS), which utilized a revised Monthly Administrative Report Recordkeeping System (TURK) to reflect manpower allocations and costs. A new statistics letter to provide an expanded data base for measuring the results of investigative activities was also prepared. ^{62:64}

After surveying the techniques currently in use and being studied for evaluation and productivity measurement of the CJS and its component agencies, numerous problems stand out, including:

1. Definitional differences, for example, of recidivism.
2. Lack of clear CJS relationship between incapacitation and deterrence.
3. Use of models and statistical or operations research techniques not totally applicable.
4. Lack of total crime statistic or even of an agreement on how to delineate the "dark area" of unreported crime.
5. Lack of suitable input data for evaluation of the CJS as a total system, or for subsystem analysis.

The following is a survey of criticisms of the methods of evaluation and productivity measurement within the CJS. Some are self-criticisms by authors who point out flaws in their own work. Some are by authors who purport to offer something better - but do they?

One example of the problems encountered in applying a productivity measurement system which includes the feedback concept of recidivism to the CJS, is the varied definitions of the term "recidivism." Webster defines recidivism as "a tendency to relapse into a previous condition or mode of behavior," whereas Kitchener, et al., see recidivism or failure after release in the CJS as being a parole violation or any conviction of a felony or felony-like offense (including fines and probation sentences).^{42:}

The Bureau of Prisons in 1970 defined a failure after release as a parole violation or any sentence of 60 days or more (including probation).^{76:}

Gottfredson, et al., see a failure after release as a return to prison for 60 days or more.^{33:} Kassenbaum, et al., see a failure after release as being returned to prison, or 90 days jail, or felony conviction.^{41:} And Simon and Cockerham apply the constraint, "reimprisoned for any term," to define the very same phrase or word.^{69:}

According to Blumstein and Larson, at nearly every processing stage in the CJS, one of the possible alternative decisions is to dismiss the offender from further processing, making estimates of recidivism difficult by the fact that we rarely know when an individual has committed a crime. We are limited to using such probabilities as those of rearrest or reimprisonment and the observed values will depend on the definition used. In criminology, recidivism is often defined as "a falling back or relapse into prior criminal habits, especially after punishment, and this may grossly underestimate the probability of "repetition of crime," the true but unknown recidivism.^{10:213}

Williams notes three problems in defining the dependent variable of recidivism:^{81:8}

1. What event is to be considered a "failure" -- a rearrest, a reprosecution, or a reconviction?
 2. How can the seriousness of the recidivistic event be taken into account?
 3. How can the frequency of recidivistic events be accounted for?
- Whatever the method used, unknown or unmeasurable factors, such as the degree of interest taken in an offender or his own decision to

give up offending, influence outcome and interfere with the relationships which, for the population as a whole appear to exist between certain variables and reconviction.
55:30

In a classic study on the control of recidivism, according to Harris and Moitra, Martinson remarked that "with few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism." They also noted that in a 1931 keynote report, the Wickersham Commission deplored the lack of systematic, accurate, and complete statistics on crime, criminals, and CJS, and went on to say although it is commonly recognized that recidivism potential is time-dependent, such assessments always disregard the times at which events occur, and the transfer of statistical methods to the law enforcement/criminal justice environment is far from simple, since these situations are unique in operation and usually lead to nonstandard problems, such as:

1. Data collection (mostly still manual and highly unreliable).
2. Analysis and interpretation (so many complex socioeconomic issues are involved).
3. Implementation (a most political environment).
35:78-79

Van Alstyne and Gottfredson said more sophisticated technique does not improve the ability to predict parole success beyond that achieved by the simpler method, which indicates that improvement in statistical technique may not be the best means of improving predictive efficiency.

Although the use of predictions of "dangerousness" or recidivism risk as a basis for the nature or duration of state intervention is a matter of controversy, the need for more efficient statistical prediction methods as aids in theoretical tests, policy studies, and evaluation designs is of continuing concern. The development of criminological prediction, which has occurred primarily in the area of parolee risk assessment, has been concerned almost exclusively with increased sophistication of the statistical methods. Despite the clear trend in the development of statistical prediction toward more theoretically appropriate statistical models, recent evidence indicates that the more advanced statistical techniques have added little to overall predictive efficiency. One study by Wilbanks in 1972 that compared the efficiency of several techniques using the same data set and employing the requisite validation procedure found the less sophisticated techniques to perform as well as the more advanced methods in the validation samples. Data typically available for predictive analysis contains numerous errors of measurement and the more sophisticated methods are more susceptible to capitalization on chance variations in constructing the prediction equations.
77:172-173

Although Solomon has demonstrated that the log-linear technique has potential for prediction studies, the results obtained by the method must be shown to have predictive validity to have utility for policy studies or evaluation research (according to Van Alstyne and Gottfredson). It must be shown that the configuration of elements derived in constructing a prediction instrument

are also predictive of the criterion in independent samples from the population. The results of the validation study by Van Alstyne and Gottfredson indicate that, for this sample and with the attributes chosen for study, the log-linear technique achieved about the same predictive efficiency as was obtained by the more parsimonious Burgess method in both the construction and validation samples, which may have been caused, in part, by the fact that interaction effects were not important in these data for accounting for parole outcome. Much more research emphasis should be placed on the nature of the predictor and criterion variables studied. The variables included in the analysis, as well as in most contemporary prediction studies, have repeatedly been shown to predict recidivism (variously defined) at a modest level.

If predictive efficiency in this area is to be increased, it would appear that considerable attention needs to be given to the discovery of predictor candidates that can add new dimensions to explain the variance in outcome. The inability of statistical advances to increase predictive power, coupled with the multitude of policy-relevant and evaluation purposes to which statistical prediction techniques are suited in criminal justice, would seem to lend greater urgency to this need. The failure of the log-linear results to replicate, and the fact that an adequate model could be found for the construction sample but not for the validation sample suggest that great care must attend the use of the technique as a theory-testing method.^{77:176-190}

According to Stollmack the most common method for predicting incarcerated populations appears to be extrapolation of linear trends totally on past

data. This is of little use in predicting the effects of recent changes in arrest rates, court policy, release policy, etc. Total reliance on linear regression is a tacit admission that we know nothing about the process which creates the phenomenon.^{73:142}

Since this prediction of incarcerated populations is usually based on extrapolations of linear trends, sometimes using regression, and relying totally on past data, the mathematical models are always incomplete representations of reality in that they can never incorporate all factors affecting a process. The validity of projections made using current methodology is dependent on reliable estimates of the parameter values (conviction rates, incarceration rate, etc.). Some models should be used only for short-range projections, since they do not specifically account for long-term effects of factors such as changes in profiles of the population or trends in public and court attitudes.^{73:142-162}

The Stollmack and Harris study points out the risk of using statistical techniques when obtaining random samples may not be possible. Using this method, one cannot trace failure rates back to causal socioeconomic events such as riots, unemployment, etc.^{37:2}

The major supporters of the incapacitation argue there is no question about whether incapacitation works. Brounstein and Kamrass say the critical question in assessing this strategy is the number of crimes that the criminal would have committed if he were free on the streets. If the number is large, incapacitation can be very effective. But if

the number is small, we do not avert very many crimes for the \$10,000
53:5
it costs to imprison a man for a year.

Shinnar, one of the investigators who has developed the stochastic-
process model of incapacitation, has argued that a significant fraction
of the increase in crime over the 1960-70 period is attributable to
the fact that the chance of a reported serious crime resulting in im-
prisonment of a criminal has declined from about 10 percent in 1960
to about 3 percent in 1970. He argues, therefore, that we should
restore the same incapacitation rate we had in 1960, and that we would
reduce crime by as much as 50 percent through that incapacitation.
Another group argues that is is improper and probably unconstitutional
to use any prediction of a man's future criminality in deciding how to
punish him for a current offense. Their argument derives from the
illegality of punishing people for future criminality. They further
argue that violent crime is so inherently rare that anyone who tries
to predict whether an individual will commit violent crimes in the
future is almost always certain to be wrong, citing a high rate
53:5
of "false positives."

The dependent variable in recidivism research is usually some measure
of recidivism, but since people do not report when they commit a crime,
subsequent arrest is used as a proxy. Defining how far a defendant
must move into the CJS before being considered a recidivist is important.
If arrests are used, someone may be included as a recidivist who did

not actually commit another crime; if convictions are used, many persons
who did commit crimes will not be included. In order to develop a comparable
measure of the frequency of recidivistic acts, it is necessary to give
each person an equal amount of time to recidivate. There is still another
complication in giving defendants a fixed period to recidivate: each
53:6-269
person must be able to recidivate.

One reason why models are not used to a greater extent is that they
must be validated, that is, it must be demonstrated that the model will
be predictive when used on a group other than the one upon which it
was constructed. The successful validation of a model does not result
in its instant acceptance. The CJS is very protective of its domain.
Decisionmakers view with a great degree of suspicion anything they
see as an attempt to replace their professional judgment with "computer
53:277
judgment."

Basic assumptions must be met before linear regression techniques can
be applied with maximum effectiveness, and linear regression models
which predict events with only two outcomes violate two basic assumptions,
58:64
according to Palmer and Carlson.

The analytical approach behind linear regression begins with assumptions
about the manner in which two variables are related. The value of one
variable is believed to depend on the value of another. When this relation-
ship is stated in the form of an equation, the first step in constructing
a regression model has been completed. In the real world, however, a

given value of Y is rarely found to be associated with a given X value, especially in the social sciences. More likely, a particular X value is found to be associated with a range of Y values. All values of Y that are actually observed are considered to consist of two parts:

1. The result of the X variable having taken a particular value.
2. Due to the influence of the random error.

Once the regression equation is specified, four assumptions are usually made concerning the random error term. It is assumed that the error term is normally distributed, has a mean of zero, and is homoskedastic and nonautocorrelated. 58:64-65

Error terms drawn from distributions not having an identical variance are said to be heteroskedastic. The line $Y = a + bX$ describes the basic relationship believed to exist between variables Y and X. For actual values of Y observed at values of X, if the relationship were exact with no random error, all of the points would fall on the regression line, and no dispersion would be observed. The operation of the random error causes the actual observations to fall at varying distances from the line. The difference between a given point and the line represents the size of the error for that observation. Because the error term is homoskedastic, the dispersion around the regression line tends to be of the same magnitude at each of the X values. The dispersion of the error term tends to change as the value of X changes, implying that the variance of the error term is not constant, in heteroskedastic examples. 58:66

Nonautocorrelation implies that the error term at one value of X is not correlated with the error term at another value of X. If it were discovered that, for example, the error term at each value of X tends to be a fixed percentage larger than the error at the previous value of X, this assumption would be violated. The error term must always occur at random, as if drawn from a hat, and not be generated by some regular process. Three conditions apply to X, the independent variable.

1. The independent variable must be nonrandom.
2. The set of X values must be fixed between samples.
3. The X values must not all be equal to the same number and they neither grow nor decline without limit as sample size increases. 58:66-67

Linear regression analysis can provide estimates of the true parameters (a and b), and in effect minimizes the sum of the squares of all the residuals. Linear regression analysis also produces statistics which are used to measure the quality of the results. The variances of the estimated parameters are provided. These are used to construct t-statistics which in turn are used to test the hypothesis that b (or a) is equal to zero. If it is possible to reject the hypothesis that b equals zero, then \hat{b} , an estimate, is said to be statistically significant. Significance does not indicate size. T-statistics are valid only when the assumption of the normally distributed error term is satisfied. 58:67-68

Another statistic provided by this analysis is the R^2 , a number between zero and one, the percent of the variation of Y which is explained by the variation of X, but the R and R^2 may not be meaningful if the

assumption of a normally distributed error term is violated. According to Palmer and Carlson, one of the most consistent confusions about linear regression analysis in the criminal justice literature is that it is applicable only with linear relationships. Although there need not be a linear relationship between the dependent and independent variables, there must be a linear relationship between the parameters and the error term. That is why this technique is called linear regression. 58:68-69

In a regression model which has a binary dependent variable, two of the basic assumptions do not hold. If the Y variable is restricted to two values (e.g., zero and one), the error term cannot be considered normally distributed or homoskedastic. A normally distributed error term can theoretically assume any one of an infinite range of values given a particular X. Thus, in the case of a binary dependent variable, the stochastic disturbance cannot be considered normally distributed. The variance of the error term obviously depends upon the value taken by X. Because of the nonnormality of the error term, the R^2 as well as the t-statistics used to measure the significance of the coefficients are meaningless. Because of the heteroskedasticity, the model will produce "inefficient estimators and imprecise predictions." The results of a linear discriminant function are the same as the results of linear regression analysis when the dependent variables can take on only two values. 58:71-72

When a researcher is confronted with data which violate the assumptions of homoskedasticity or normality, there are various pooling or weighting techniques which can be used to eliminate or mitigate these problems. One such method is called weighted least squares, which requires running the regression twice. There are several drawbacks associated with the method of pooling of data. Pooling and weighting improve the quality of the results but neither of these two methods completely eliminates the problems inherent in using linear regression with a binary dependent variable. As an example of another problem encountered in recidivism analyses, a Michigan study does not distinguish between persons returned as parole violators and persons returned for new convictions. 58:71-72

Nonrandom samples appear in many recidivism studies, for example, a stratified rather than a random sample could include observations for only one sex or for only 1 year. There is a considerable literature dealing with what are called validation studies. These involve testing the results of estimates obtained with one data set on a second data set. Some of the techniques used are inappropriate when applied to a regression analysis which has a binary dependent variable. One method estimates coefficients with one data set and then uses these coefficients to predict the recidivism rates of another data set. The residuals, which are computed as the difference between the Y's predicted by the coefficients and the Y's observed in the second data set, are used to compute an R^2 . This R^2 is subtracted from the R^2 obtained from the original data set and the difference is the "shrinkage." But, neither R^2 has any meaning when the dependent variable is a binary. Therefore,

their difference has no meaning. Regression analysis in any form estimates the probability of parole success or failure. It does not predict success or failure per se. The logical error is in assuming that a probability of .6 is a prediction of parole success. It is not. It is simply a prediction that 60 percent of the parolees will succeed. Palmer and Carlson summarize, saying the history of the use of regression analysis to predict recidivism rates has been less than spectacular. 58:78-80

According to Harris and Moitra, there are problems of comparability and validity of program assessments. It has become clear to many concerned with the analysis of the complex data made available by law enforcement and criminal justice agencies that there is a great need for more agreement on the meaning of program results and then for much more satisfactory means to measure "recidivism." Citing numerous problems associated with classical techniques of measurement, they say a major problem is that, although it is commonly recognized that recidivism potential is time dependent, such assessments always disregard the times at which events occur. 36:194-196

Delinquency in a Birth Cohort addressed the question, "Can inferential statements be made about switching from one type offense to another or continuing with the same type offense?" or "Can we predict the kth crime type?" Two problems were cited:

1. Determination of the transition probability $P_{ij}(k)$ where k is the number of the offense in a series, by type j of the kth offense, given the type i of the k-1 st offense. 27:175
2. Comparison of transition matrices generated.

Blumstein and Larson noted, "in order for an offender to be sent to prison, at least seven actions must occur after the crime is committed":

1. The crime must be detected and/or reported to the police.
2. The offender must be arrested.
3. He must be charged with a felony.
4. The suspect must be prosecuted (the prosecutor must ask for an indictment).
5. The suspect must be brought to trial.
6. The defendant must be found guilty.
7. The convicted offender must be sentenced to a . . . correctional institution.

Thus, the probability of imprisonment, given that a crime has been committed, is:

$$P_1 \cdot P_2 \cdot P_3 \cdot P_4 \cdot P_5 \cdot P_6 \cdot P_7 = P$$

If the mean time served by those sent to prison is T years, then the expected incarceration for one offense is P(T), which might be considered by a rational individual contemplating the risk in committing a crime. Further, if for burglary the fraction of cases in which the police were not notified was 0.42, then $P_1 = 1 - .42 = .58$ and for the California model:

$(.58)(.23)(.50)(.51)(.91)(.93)(.25) = 0.007$ is the probability of adult incarceration for burglary, given a burglary committed by an adult. The average time served in prison (including parole violation time) being about 3 years, then the expected time one could be incarcerated for burglary is 3 (0.007) years or about 7.7 days. This example depicts well the obvious discrepancies in results obtained by using different definitions of recidivism.^{10:}

According to Deutsch, the controversy of statistical evaluation of deterrence effectiveness of the present CJS and of specific programs or changes implemented in the existing system has in part centered around the value of crime incidence data. He presented an example using an empirical-stochastic model developed solely from UCR data to illustrate an approach to evaluation of system effectiveness. Deutsch also noted a 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.^{21:i-1}

The use of crime-rate indices as performance indicators presumes that:

1. Crime rate data accurately reflects true victimization.
2. Changes in CJS effectiveness account for nearly all changes in victimization.

The first assumption, that official crime rates accurately reflect true victimization, is subject to great controversy. Several researchers, such as Ostrom, suggest that the FBI Crime Index is widely

regarded as being extremely unreliable, and Ostrom reports that even the President's Commission on Law Enforcement and Administration of Justice considered victimization to be best determined by citizen surveys. Certain weaknesses have been identified in the accuracy of the FBI's UCR. 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 crimes across the nation,^{21:i-7} rather than in the accuracy.

Forst reviewed and updated the empirical aspect of an analysis by Isaac Ehrlich of the effect of specific deterrents on the index crime rate. For 1960, Ehrlich estimated that a 1 percent increase in spending on police would produce, by way of an increase in the probability of punishment, a 3 percent decrease in the serious crime rate. However, Forst used data for 1970 within a similar analytical model and found the crime rate to be virtually insensitive to cross-state variation in either the probability or severity of punishment. Forst found most of the difference appeared to be due to methodological problems with Ehrlich's empirical work, and cast doubt on Ehrlich's result and on the strong policy recommendations that have followed his analysis.^{30:i}

Forst's small elasticities and t-statistics for the crime deterrence variables contrast sharply with Ehrlich's findings. Elasticity of a binary variable has no meaningful interpretation. Forst attempted to update Ehrlich's empirical findings with data that are 10 years more recent using a more exhaustive set of variables. He found punishment

variables have a smaller effect in 1970 than in 1960 using Ehrlich's model, saying the evidence presented suggests strongly that Ehrlich's crime deterrence variables are, to a large degree, substitutes for demographic factors that are real determinants of crime. Forst further noted the omitted variables problem is not new to the literature on the economics of crime and delinquency. 30:9-21

A number of problems stand in the way of partitioning the index crime rate. One is that many reported criminal episodes consist of multiple offenses, another is that the rate of nonreporting of criminal episodes varies across crime categories. Since the number of offenses is both the numerator of the crime rate and the denominator of the probability of incarceration, any error in the measurement of the number of offenses would exaggerate the estimates of the deterrence effect of the degree of "certainty" of punishment. 30: footnotes 3, 11

According to Shinnar and Shinnar, in a complex system, exact estimates are impossible, but for purposes of policymaking, a lower bound on the effects of incapacitation could be very useful. Deterrence and rehabilitation will decrease the number of criminals and/or reduce their individual crime rate and the length of their career. As long as we can assume that incapacitation does not increase these parameters, the prediction of our model will be conservative in the sense that the reduction in crime will be larger than predicted, but can we? The assumption that λ is uniform is also incorrect. Shinnar and Shinnar assume a large fraction of those entering a criminal career have a short career, and career length is exponentially distributed.

There is another assumption which is crucial to their predictions. In using the measured value of qJS and λq they made the implicit assumption that the average q for the criminal is equal to the fraction of crimes solved by conviction. They attribute the majority of unsolved crimes to criminals who are convicted at least once. This is crucial because 70 percent of all safety crimes in the United States are never solved. If most crimes are committed by criminals who are never caught, then no incapacitive policy will work until there are means to catch them at least once. Since prison sentences or other convictions may never be recorded in the convict's file, arrest records are the best personal data we have. 68:591-592

According to Shinnar and Shinnar there are two ways one can perform a study on recidivism:

1. Look at an instantaneous sample of arrested offenders.
2. Follow the career of a sample of offenders.

The second method has the advantage that it is less affected by the unsteady nature of the system, but has the disadvantage that very long times are needed to provide reasonable accuracy. 68:593

Both national and New York State prison statistics show the fraction of prisoners who had no previous commitment to any penal institution to be approximately 0.35, or a total lifetime recidivism rate of 0.65, but most first offenders are not imprisoned so second offenders enter corrections. There is a long unknown timelag between arrest and conviction (or commitment) and short time followups (less than 5 years) may be hard to interpret. 68:596

Data indicate that more than 80 percent of solved crimes are committed by recidivists. The question is, who commits the 70 percent of crimes which are never solved. The most likely possibility is that they are committed by the same group of recidivists who commit the crimes which are solved, but two other possibilities must be considered: (1) most of these crimes are committed by amateurs or one-timers, i.e., those who commit only one or two crimes in their lifetime. (2) These crimes are committed by a highly skilled group of professionals who never get caught. 68:597

Shinnar and Shinnar say that while we assume a uniform criminal there is obviously no such thing. The fraction of crimes cleared by arrest or by conviction has decreased in recent years, and large changes in qJS are therefore mainly due to changes in J, the probability of receiving a jail term once having been convicted. 68:599-602

McGuire points out, regarding the current attempts to quantify incapacitation effects, that none of the approaches typically result in quantifiable incapacitation benefits estimates, because of inadequate data relating to individual criminal careers, either known or actual. The same problems affect estimates of crime cost and displacement effects, acceptable measures of neither having been computed. The estimates however are of arrests avoided, not crimes. They thus represent a probable minimum index of criminal activity avoided, given incarceration of the sample confined population. 48:14-17

The choice between incapacitation effects measured as arrests or victimizations cannot be made on the basis of information currently available. If it is assumed that the sample population is composed of the least competent criminals, who always are apprehended, then arrest may be the proper measure. If it is assumed that the sample population is a random draw from the total criminal population, then victimization may be the appropriate measure. The truth probably lies between these poles. Regarding policy applications, this situation is unfortunate because the effectiveness of the sample institutions as producers of incapacitation effects is markedly enhanced the closer to the victimizations pole is the true state. Analyses have been conducted under the tacit assumption of no displacement effects. And entry of these into the question can only reduce the incapacitative effectiveness of institutions, and more so if the magnitude of the effects are positively correlated with offense categories strongly represented in the sample confined population. Displacement effects cannot be quantifiably entered into the analysis. 48:19-21

According to Wenk, et al., the quest for an operationally practical predictor of violence from simple classification appears to be futile, and the present state of the art holds little promise for the development of a prediction instrument that would warrant implementation in actual preventive or correctional programs. The problem is fundamentally related to the nature of the phenomenon: reported violence. Violence typically erupts out of a crisis, and the certification of the events is rare. Concern about violence will inevitably lead to the development

of special treatment programs, but the majority of persons placed in such programs must be false positives—persons who would not commit the act which the program is designed to prevent.
65:400-402

According to Gibbons, "existing typologies of criminals and delinquents are ambiguous and lacking in specificity" and an "adequate typology should provide detailed and specific indicators of the descriptive ingredients of the categories in the system in order that the claims can be checked against empirical evidence." He cites the fallacy in research on groups or types of violators which ignores that each violator is an individual.
31:39-297

Stewart notes that the evaluation researcher, monitor, and decisionmaking evaluation consumer all bring different perspectives to the conduct of an evaluation. They each adopt a certain sort of tunnel vision in which purposes are very narrowly defined.
71:25

The report of The Criminal Justice Symposium Focusing on Police Productivity noted:

1. The fragmentation of police forces means that many are unable to afford the overhead investment in new analytic talent and/or the training of existing staff necessary to rationalize operations.
2. The police are sorely in need of the concentrated analysis and productivity improvement required by all other Government operations.
3. Because assessment necessarily incorporates subjective judgments, unanimity on the exact benefit of a particular change is impossible.

Too often the response to "the law enforcement problem" is to spend more tax money and hire more police officers.

5. In most cases, unless the number of police officers is increased dramatically, the money involved in adding just a few men might be better spent in upgrading existing manpower or their equipment.
59:2-17

The Symposium report also said victimization studies have demonstrated that the UCR reflects only a percentage of those crimes committed, but the absence of other measures have elevated crime statistics to a level of importance far beyond their actual worth. Even if they were accurate, reported incidences of crime are the results of numerous and various conditions over which the police have little or no control (e.g., the proportion of low-income families in the community, the ratio of youths to the total population, the number of unemployed, the population density, and the effectiveness of courts and correctional programs). Other important factors that affect the usefulness of crime statistics are the methods by which they are collected and recorded and the consistency with which they are interpreted. Any of these factors, or several of them taken together, may have more to do with changes in crime rates than anything the police department may or may not do.
59:18

It is an unfortunate reality that we have few useful measures to assist police managers in pinpointing and managing their resource utilization problems. Police managers need information that is concerned with police performance in light of the goals and activities of police work. Also, because different crime rates represent differing opportunities for making

arrests, the apprehension productivity measure must be considered in light of changes in the crime rate, thus productivity improvements due merely to higher crime rates should be distinguished from improvements resulting from better use of the patrol force.

59:26-29

Some police departments and police administrators are evaluated on the basis of public relations. When this happens, we encounter "negative" productivity issues that can bring down a police department. Administrators in many public services have simply grown accustomed to adding personnel as the sole management response to be considered. Consequently, police, fire and emergency medical services comprise some of the most labor-intensive, undercapitalized industries in the United States today.

59:36-38

Implementation Difficulties include: (1) Ill-Defined Objectives and Constraints. A popular word in operations research, optimization, often bears little relevance to operational realities of governmental service systems, primarily because of the difficulties in defining objectives and constraints. (2) Lack of Productivity Measures. Since system objectives are poorly defined, so are measures of system productivity. (3) Internal Resistance to Innovation. Innovation is apt to be frustrated unless there are receptive personnel in key positions. (4) Resistance to Outside Technical Assistance. (5) Operational Complexity. No one has yet found a way to reliably estimate the number of crimes prevented by the police.

59:42-69

A related problem in using the conventional definition of productivity is that there is not a single output of police crime patrol. Police services have a variety of purposes and cannot be adequately reflected by any single indicator. The objectives of police crime control include: (1) arrest of offenders, (2) promotion of a feeling of security in the community, (3) improving the trust of people in their local government, (4) protecting the moral sensitivities of the community, and (5) enforcing ordinances against nuisances. Although it may be tempting to develop a weighted index of these various outputs so that a single number can be used to represent the combined output, different people will disagree as to the relative weights to put on different outputs. The use of such weights often mask the value judgments of their creators. Changes in the CJS are difficult to analyze, in part due to the absence of standards.

59:69-81

According to Greenberg, a predictive device such as the California Base Expectancy score can make two kinds of errors: (1) It can release individuals predicted not to recidivate but who in fact do so (false negatives). (2) It can fail to release individuals predicted to recidivate, but who would not recidivate if released (false positives). If the aim were simply to minimize error, the board could do better by releasing all the inmates, since it would then be wrong only 12 percent of the time. A statistician's best predictions identified 7.7 percent of the violent recidivists, with a false positive ratio of 12 to 1. He concluded pessimistically: The lack of precision of our selection process seems inherent in the limitations of the quantifiable variables we have.

34:544-548

The state of prediction is evidently rather poor, and implementation of a policy of selective confinement based on predictions of dangerousness would clearly founder on the gross inaccuracies of predictions. There are some fundamental limitations to the degree to which improvements are likely to be possible. (1) The extreme practical difficulties and high cost of improving the accuracy of data to be used in predictions. (2) The interactional nature of much recidivist crime, and in particular, of violent crimes. If a particular individual's recidivism depends not only on his or her own personal traits, but also on largely unforeseeable contingencies such as how others behave toward that individual, the information that would be essential for an accurate prediction would be omitted from the actuarial analysis that forms the basis for a prediction. Contingencies may be at least as important as biographical data in determining parole success or failure. 34:548-549

Consider a man released from prison who is now accused of another crime. No one would question that when he is tried for this new offense, the appropriate standard in reaching a verdict should be the reasonable doubt test, just as in the first trial. Suppose, however, that instead of having been accused of committing a crime that has already taken place, the man had been predicted to engage in some crime at a later date. Why would we tolerate a lesser degree of certainty with regard to the incarceration of someone predicted to engage in a crime that may not take place? If this reasoning is persuasive, the reasonable doubt criterion would have to be used for decisions involving selective incapacitation. 34:549

Most rehabilitation programs have shown no measurable effect on recidivism, and in the few cases where an effect has been shown, it was not large. Recidivism rates of released prisoners are about the same as those of matched probationers, and one recent study by Berecochea in 1973 found that length of time served in prison had no effect on recidivism. There is no compelling evidence that imprisonment substantially increases (or decreases) the likelihood of subsequent criminal involvement. 34:558

Offenses leading to imprisonment are only the tiny, most visible tip of a very large iceberg of offenses that do not lead to a police report, arrest, conviction, or imprisonment. In 1970, for example, there were 1,551,300 arrests for index crimes in the United States, but only about 77,000 persons were sentenced to prison in that year, and not all of them for index offenses. The clearance rate for index offenses was only 20percent in 1970, and this figure would be even smaller were crimes not reported to police included in the measure of likelihood that an offense will result in an arrest. Parole statistics indicating only rates of return to prison understate the amount of recidivist crime committed by parolees. There is no published information indicating what percentage of all arrests in a given year are virginal, according to Greenberg. 34:558-564

A victimization study conducted in 1965 by the National Opinion Research Center concluded that slightly fewer than half of all index crimes were reported to the police. Other studies of victimization found varying degrees of unreporting: in some areas there were 1.5 times as many crimes reported by victims as were reported to the police, while in other areas

the rate of nonreporting was 3 or more. For 1972, roughly half of crimes of violence and burglary of households were reported, about a quarter to a third of personal larcenies, and 75-80 percent of commercial thefts. 34:564-565

When criminals are taken out of circulation in substantial numbers through imprisonment, market forces may attract noncriminals into criminal activity. The magnitude of the crime prevented through incapacitation must also be reduced by the amount of crime committed as a result of imprisonment, by inmates against other inmates and guards, and by guards against inmates. 34:567-571

Two factors limit the size of the incapacitative function of imprisonment.

1. The low rate of return to serious crime among parolees, which may indicate that many inmates are nearing the end of their crime careers by the time they are sent to prison.
2. The low rate of imprisonment for index crimes, which reflects low clearance rates, prosecutorial discretion to drop charges or reduce them, and judicial reluctance to impose prison sentences.

Unlike street lighting, which may also reduce crime, imprisonment imposes very heavy costs on a limited number of individuals. 34:572-576

Malmberg and Deutsch point out that since λ_t is a measure of the free criminal's propensity to commit offenses in period t, if we knew the number of periods (n) an offender was incarcerated, the product, $\lambda_t n$, would estimate the potential savings realized by imprisonment of that individual for n periods. If we knew the number of individuals who

were incarcerated in each period, r_t , we would estimate the number of crimes averted in the future through incapacitation from prevailing policy in period k as :

$$\sum_{i=k}^{k+n} \lambda_i r_i.$$

The real hole in this idea is that nobody has even a valid estimate for λ , the criminal's rate of crime commission. 23:141

The Report of the Advisory Group on Productivity in Law Enforcement 54:7-24 noted:

1. Many of the measures currently being applied to police services do not provide managers with the information they need to help them improve operations.
2. Because the incidence of crime is a function of many factors unrelated to police activity, crime rates alone are insufficient measures.
3. The UCR documents only reported crimes.
4. One reason that existing data are not put to better use is that the police mission is complex, and specific objectives of the force are not always clear.
5. The majority of these data are not sufficiently refined to provide police managers with dependable and useful information which can lead to better performance.
6. Police both affect and are affected by other elements of the several systems of which they are a part. Effectiveness in preventing crime,

for instance, depends in part on how well the corrections agency performs in rehabilitating felons.

7. Different types of arrests have different values.

The Advisory Group, recognizing that expenditures for law enforcement were not unlimited, said the more difficult problem is how to increase the effectiveness of those available resources, and traditionally, the number of arrests has been used as an output measure. However, arrests themselves may be too easily subject to inflation and clearances may be unsuitable because crime frequently cannot be attributed accurately to offenders. 54:17-22

One difficulty in productivity measurement of a law enforcement organization is the numerous services provided by the force that do not relate to incidents of crime or suspicious activities which make up the large majority of calls for service. Another difficulty associated with assessing the relative effectiveness of special units is that the departments that use them find they may attract more capable officers, ones who normally would account for a high number of quality arrests on regular patrol. Officers with attributes which prove to be positively related to effectiveness may tend to cluster together. 54:27-35

A principal objective of the police is to prevent crime. Yet many police departments do not think positively and specifically about crime prevention. The sum efforts of the police department theoretically are geared toward deterring crime; the very existence of the department serves notice on would-be criminals that society has the means to track down and apprehend offenders. Unfortunately, the factors affecting crime prevention are

extremely difficult to isolate and measure. Difficulties must be overcome before reliable measures of productivity in crime prevention can be devised; e.g., the rate of reported crimes represents only a fraction of all crime committed. Because productivity is a comparative concept, care must be taken to avoid the "measuring of apples against oranges" and other statistical fallacies. Crime-prevention activities have, in many instances, been effective, but most have been subjected to little evaluation beyond subjective judgments or limited observations. One is unlikely to be able to make judgments about priorities and resource allocations among various crime-prevention activities unless it is known how productive or effective they are in comparison to each other. 54:37-39
Measures are simply a tool for better evaluation.

The Advisory Group also noted unless the costs to carry out these programs can be isolated, productivity improvements will be difficult, if not impossible, to measure. The results of measurement must enable police managers to distinguish between crime-prevention programs that yield a lasting improvement and those whose effect is only temporary. Measures must also be structured so as to relate the improvement directly to specific activities carried out in the program and to a determination as to whether or not the program has simply displaced crime. 54:40-41

Many questions remain unanswered in the field of crime prevention, such as:

1. What are reasonable crime-prevention goals?

2. What behavior patterns can be changed or encouraged to decrease the likelihood of crimes being committed?
3. How can the changing sociocultural profile of a community be described, and how does this affect ongoing crime-prevention programs?

Systematic research is needed to make possible the planning and design of more effective programs. Universities, research institutions, and State and Federal agencies working to prevent crime have a responsibility in this area. 54:43

With respect to individual officers the Advisory Group said no satisfactory "hard" crime-related indicators exist for measuring the performance of police personnel, nor does a single measure or index of performance exist for an individual policeman. 54:48

The Advisory Group also cited barriers to productivity improvement such as the reluctance to try new ideas. Bottom-up departments, such as those found in most police departments, tend to become closed circles in which practices pass down from one closely knit group to another as new recruits are "taught the ropes," and while the attitudes developed by an officer reinforce solidarity within the force, they also discourage openness to outside ideas. The discouragement of lateral entry deprives the department of technical skills needed to select and evaluate innovations. The bottom-up structure places in leadership positions men who may have demonstrated excellent operating skills and abilities, but does not necessarily also insure skill in management. Too few police organizations have innovative

leadership. Innovativeness of departmental leadership can also be highly dependent on political considerations. 54:65-66

Once adopted, innovative programs may have difficulty in surviving. Political realities make it difficult for police management to support programs other than those that appear successful soon after their adoption. The lack of technical skills among most department personnel often makes it difficult to carry out innovative programs, and difficulties in evaluating programs make it difficult to "prove" the effectiveness. Innovative new programs are especially susceptible to rejection if they are felt to detract from the functions regularly expected of police departments. Another factor affecting success or failure is the productivity of the resources invested in a new program, and the delay before an innovation becomes productive. Any action that would promote the acceptance of innovation by police departments at the expense of organizational cohesion and morale could well be counterproductive. 54:67-68

Traditionally, the FBI has managed its investigative staff on the basis of caseload and accomplishments, not on the nature and importance of its investigations. Realizing the limitations of this method, they implemented a new approach called quality over quantity, but a clear definition of a quality case or priority area was lacking. Their accomplishment statistics have been misleading. For example: the format did not explain what the statistics meant and how dollar values had been determined. Caseload alone is not a good indicator of staffing

needs, because investigations vary in their nature, complexity, and importance.
62:1

Statistics may continue to be misleading, due partially to a lack of criteria that clearly state how accomplishments are to be claimed. The FBI had certain problems that inhibited effective implementations of productivity measurements:

1. The FBI had not clearly defined what constitutes a priority investigative area or established criteria for identifying quality cases versus cases of marginal importance.
2. The FBI had not developed sufficient management information for implementing and measuring the effectiveness of the quality over quantity approach. Existing information was primarily caseload related. Information on the results of investigations and on the application of resources did not interrelate. Information on investigative results was limited to a few categories of accomplishment statistics that were misleading because of the way they were presented.
3. The FBI and U.S. attorneys generally were neither coordinating the selection of criminal security problems for priority investigative and prosecutive attention, nor developing prosecutive and investigative guidelines for violations not normally prosecuted.

Also according to a 1978 Comptroller General report, no attempt has been made to establish a precise definition of quality that will apply to every investigation to determine whether it fits predetermined criteria.
62:35-65

CONTINUED

2 OF 4

Deutsch points out some problems with current law enforcement productivity measurements, such as improper use of the statistical methodology underlying control charts which is not directly applicable to detecting shifts in time series data since the monthly occurrences of a particular type of crime have been shown to be correlated. He points out, relative to the study of effectiveness of the Massachusetts 1975 Gun Control Law, the use of poor input data, such as the murder and nonnegligent manslaughter classification which also includes those homicides that resulted from knives or cutting instruments as well as other dangerous weapons, and the armed robbery classification, which had been used when the robber was armed with any weapon, not merely a firearm. It should also be noted that the same study ignored external factors such as the Federal Gun Control Act of 1968 which was implemented during the data years the study was based on.

Riccio points out that some models are not intended to be an accurate representation of criminal and crime abatement activities, being almost entirely constructed from intuitive analysis and not real world data.

According to Chapman, regarding the studies of Sutherland, Becker, Ehrlich, and Bonger, most of the empirical testing of the hypotheses has either ignored or submerged other basic parts of the CJS. Early studies have been criticized because they neglected to consider the possibility that the neighborhoods may have acted as collectors of criminals rather than having acted as a corrupting influence upon the potential lawbreakers. The statistical techniques employed have not

always been entirely appropriate, and crucial parts of the theoretical development were lacking in some studies. In both of the Quinney studies, the emphasis is on correlation analysis, usually between crime and one factor, when a far better technique that could be used is some form of multiple regression analysis which can hold specific variables constant.

Arrest rates have been considered an output of police and thus the relevant dependent variable in a police production function. This is an oversimplification of police output, which includes far more activities. Studies which attempt economic "rational criminal" analyses contain potential problems. The city's loss is not necessarily the criminal's gain, since it is unlikely that the criminal will realize full value of the goods that are stolen.

There is no real theory of which environmental variables should influence the violent crime rate, and punishment is not included in some models, reflecting a lack of understanding of the entire CJS.

Maltz cites several reasons for poor evaluations in the CJS:

1. Many evaluations are based on insufficient data sources.
2. The nature of the political process.
3. The lack of expertise of those called upon to perform the evaluation.
4. It is almost always easier to describe the problem than it is to prescribe a viable solution.

The type of evaluation used most frequently has its roots in experimental research and seeks to determine the relationship between two variables in which it implicitly assumed that the dependent variable does not affect the independent variable. However there is no "standard" population. In a crime control program, it may be impossible to classify variables as dependent and independent; they may all affect and be affected by each other.
46:9-10

Relating the actions taken during a program to the final results is not a simple matter. Statistics cannot and do not substitute for a logical connection between the effect produced and the conditions which produced it. Finding the logical connections between cause and effect in crime control programs is made more difficult by the elusive nature of the population being "treated": the offenders. If no control area is used in the evaluation, there is an implicit assumption that future crime rates can be reliably predicted from past crime data, however, a significant change may be instituted during the evaluation which materially affects the crime rate.
46:11-16

In many cases where crime reductions have been measured and attributed to programs, it is unclear whether there has been an actual reduction in crime or whether the crime has been displaced. The amount of displacement depends to an extent on the characteristics of the offender, according to Maltz. The categorization of differential effects of deterrents can be broadened to include the type of crime as well as the characteristics of the offender. Deterrents may have little effect on

perpetrators of "expressive" crimes, whereas deterrents may have a strong effect on "instrumental" crimes. Deterrence may produce a diversion to legal alternatives to crime; it also may cause displacement to illegal alternatives, to other forms of crime, to other tactics and targets, and to other areas. Evaluation is also made complex because offenders can change their manner of committing a crime when a new program is established to counter their activities.
46:20-21

Maltz says crime data are far from perfect, and quotes Sir Josiah Stamp, "The Government is very keen on amassing statistics. They collect them, add them, refer them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that everyone of these figures comes in the first instance from the . . . (village watchman), who just puts down what he damn pleases."
46:27

A dominant factor in the way crimes are categorized is the legal definition of the criminal acts, giving rise to a number of artificial and illogical complexities (e.g., the difference between classifying a purse-snatcher as a robbery or a larceny depending upon how hard the thief yanked the purse, whether he approached from the front or rear, the victim's perception of the situation, and the fear engendered in the victim. Many UCR categories are too broad for research purposes and some have arbitrary limits put on them. Stranger-to-stranger crime is an example which fits no single category. The Uniform Crime Reports are based on data voluntarily furnished by state and local law enforcement agencies, and were not designed for research purposes. The UCR statistics

are based on crimes reported to the police, and many crimes go unreported. Victimization studies are best suited to determining long-term effects, but not that well suited to most crime control program evaluations. ^{46:27-29}

Maltz notes inaccuracies in reported crime can result because it is assumed that the ratio of unreported to reported crime stays about the same from year to year in each category, and similarly assumed that the definitions of categories remain the same from year to year. The crime rates, as presently calculated, do not reflect the true situation. It is difficult but useful to distinguish between actual deterrence (due to an actual increase in risk) and deterrence that is purely psychological in nature (due to a perceived increase in risk). Another difficulty with the use of arrest rates stems from the operation of the rest of the CJS. The use of the arrest rate by itself, therefore, does not appear to be appropriate as a measure of effectiveness for most crime control programs. ^{46:30-39}

One of the most difficult aspects of an evaluation may be getting the police officers to fill out different or new reports for collecting evaluation data. Police data are normally not sufficient for an evaluation. They are collected by police departments for police purposes, not research purposes. Different programs will require differing kinds of supplemental data, such as citizens surveys. The information in CJS records should not be considered "hard" just because of its apparent specificity. ^{46:44-62}

Another evaluation problem, time lag, is pointed out by Avi-Itzhak and Shinnar who say that in case the probability of an offender surviving

incarceration as a person who will again commit crimes is reduced by 50 percent, the decline in crime level will not be immediate; rather it will be a gradual decline taking many years before the new crime level is achieved. The system is not stationary and the real problem, which is not easily solvable, is to quantitatively identify the parameter changes contributing to the sharp rise in crime rates. ^{3:203}

Fisher, et al., in Predictive Sentencing note the quest for rationality in sentencing is stymied at the outset by the essentially discretionary role in which the judge functions when sentencing and, correspondingly, without the benefit of efficacious sentencing criteria. At most, the sentencing process, unlike the adjudicating process of determining innocence or guilt, largely involves unfettered discretion in the selection of specific sanctions within the predetermined range of legislatively authorized sanctions that can be imposed for the commission of an offense. In sentencing, the average judge will find neither the comfort of a precise body of interpretable sentencing standards nor a body of knowledge or methodology comparable to that in the law when adjudicating. His concern now is with the behavior of the offender in the future, and advances in psychology, together with the increasing awareness of the effect of changes in social and economic well being on the behavior of people, thrust the contemporary judge into an alien sentencing environment. Why do legislators fail to give guidance to judges by addressing themselves in their legislation to the purpose and function of sentencing? Why do they fail to develop adequate sentencing criteria

and sanctions to implement this policy? The sentence of a particular judge on a particular day will reflect his own, rather than society's perception of the justification for punishment? Large-scale disparity in sentencing was noted, not only from state-to-state, but frequently within the confines of a single jurisdiction. 8:6-7

Each year judges impose roughly two million sentences with little or no objective information on the effect of those sentences on subsequent behavior. Little study has been done relative to predictive sentencing and the effects of the sentence on subsequent offender behavior. The legal path through experimentation, privacy, and altering behavior is either left largely uncharted in the judicial decisions or is a thicket of concept and doctrine that can hardly be separated into manageable components and then synthesized to form meaningful conclusions from which to evaluate the legal implications of the project. 28:9-119

Other problems occur after sentencing. With respect to the stepwise multiple regression analysis which yielded six variables, discussed by Myers and Levy, the six variables predictive of intractable behavior before incarceration would have classified correctly 78 percent of the inmates studied. However, 11 percent of the tractable inmates would have been incorrectly classified as intractable if such a classification system were used, which could result in two problems: 44:226

1. The label would be incorrect and unfair, "as a matter of justice we should never take power over the convicted criminal on the basis of unreliable predictions of his dangerousness."

2. The false label could lead to a self-fulfilling prophesy so that the "labeling process of classification . . . may change perception of the person by others, and through this his own self-image."

According to Deutsch, the popularity of performance measurement in 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. 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 should 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. 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 law enforcement agency. 20:1

A failure of the measurement strategy can lead to inconsistent or unreliable results, therefore care must be taken in selection of a measurement as in choosing the measure of performance. The single element of the measurement strategy that has thus far received the

greatest attention is the type of data, and there is great controversy as to whether UCR data is the best indicator of criminal activity or victimization surveys are best. 20:31

Deutsch notes 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. For example, a paper by Hirsch and Riccio 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?

No attention has been given to the more important considerations of determining how the overall measurement process will fit in with police activities. 20:36

Deutsch goes on to cite other work, such as that done by Larson, Avi-Itzhak and Shinnar, Hirsch, Zacker, Reppetto, and others which 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 measurement. Accurate data is required for accurate results, and data collection methods must be well defined to avoid introducing artifacts of the collection method into the data. Also 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. 20:36-37

In many cases, the cost effectiveness index is not valid for use as a measure of effectiveness, and has been improperly applied. The difficulty lies in the attempt to correlate expenditure of resources with benefits produced, tending to overly compress available information into a single ratio. It is difficult to find single objective organizations, particularly in the CJS, and attempting to express noneconomic values in monetary terms for the convenience of the evaluator is a dubious endeavor, a practice all too common among social research. 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 noneconomic objective that does not depend on the degree to which other objectives are satisfied. 20:44-47

Even though its component organizations have formal communications and assist one another, there is no real chain of command in the CJS. Regardless of the lack of a chain-of-command, there are available

approaches to determining an overall measure of CJS effectiveness. One 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 the 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.

20:54

According to Duetsch, another drawback 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. 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 systemwide GOA, there is likely to be some relation to the GOA of the component organization. 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 systemwide set of GOA ensures that there will be inconsistencies.

23:54-68

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 exists. One strategy for finding those relationships 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, but 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 police sciences also are often quite low in explanatory 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.

20:74

Deutsch notes that 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 is that the results must be available in a understandable form. Human elements within the evaluated organization will hold the results suspect unless it is clear just how the measurement approach works, which shows the need for avoidance of difficult and cumbersome procedure in the measurement process itself, and the need for results specified in workable terms. The problem for the researcher designing a package 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, and managerial cooperation is essential to any internal evaluative effort, as well as being of great value to external evaluation groups.
20:122

There is a great deal of current literature available on performance measurements for the social services, but in general, current literature is flawed by the lack of the proper preparatory work in understanding the meaning of measurements, and many efforts totally ignore the difference between measures of effectiveness, measurement strategies, and measurement processes. Deutsch says 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. 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, and several of the measures of effectiveness offered contain terms representing behavior not in control of the measured organization.
20:124-125

Deutsch noted that Holzer presented a management-oriented productivity measurement paper for application at the urban police force level, and said that although the paper had a number of interesting points, such as a recognition of the need for a conceptual framework for measurement, there were a number of serious problems with the paper. The first and most basic was 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, yet there is a great deal of evidence to the contrary. Within the paper the word "non-deterrence" appears in relation to effectiveness measurement. The two implications of the use of such a word are astounding: (1) That police control victimization levels, and that greater police effectiveness reduces these levels, which is not generally the case. (2) Until police experiment approaches are fully developed and tested in CJS applications, there can be no certain analytical basis for attempting to attribute the nonoccurrence of certain events to changes in police behavior.
20:125

The problem lies in determining the degree to which the changed behavior actually changed reported crime. If there is a variation in reported crime, finding the part of the variance due to a change in police behavior is the problem. 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. The paper is oriented toward maximizing productivity at the component level, and optimizing performance at the component level does not always lead to optimal system performance.
20:126

PPBS is inappropriate for CJS applications, according to Deutsch. PPBS as described by Mushkin and Cotton attempts to characterize public agency performance by using volume and quality indicators.
20:130

According to Deutsch, among the shortcomings of the Hirsch and Riccio effort is the primary assumption that productivity is a good measure of police performance. Productivity is quite limited in its meaning for applications in the CJS, and productivity measurements rapidly develop into unpopular quotas. Another major difficulty apparent in the paper is that no groundwork is laid for selection of performance measures, no consideration is given to measurement strategies or the design of a general measurement process, and many of the measures proposed are contaminated by factors outside the control of police. 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. No logical decision process was developed for selecting performance measures. 20:130-131

Deutsch criticized Ostrom's paper citing the lack of a well-defined decision process for selecting performance measures, measurement strategies, and measurement processes which has led to the rather arbitrary selection of a performance measure and measurement study. 20:131

Regarding the work of Mantel, et al., Deutsch says reported efforts have been limited in scope and rather superficial in their treatment of the problems associated with performance measurements for public services. The researchers made an attempt to classify each and every service (activity) of the agencies by preconceived definitions, but failed. 20:132

Deutsch goes on to note a basic problem 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. A good deal of existing research into performance measurements for CJS applications has been shown to be fragmented and superficial, and 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. Too little effort has been expended toward developing a truly universal foundation for the measurement of organizational behavior. The real shortcoming of current evaluative efforts is the lack of a coherent set of measurement principals. 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. There are many diffuse and disparate research efforts which have been able to expose a few of the relevant performance measures, but the total results fail largely due to fragmentation. 20:134-137

According to Barnett, Larson, and Odoni, pitfalls in the processing of information are due primarily to two potential sources of error: (1) excessive degree of aggregation of the data as they appear in the surveys; (2) differences from place to place in the statutes regarding the functions of police departments. They note 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

in the CJS, and the information already available from concluded surveys has not been analyzed in a way to shed light on system-level resource allocations.
4:1-5--2-2

Expenditure and Employment Data for the Criminal Justice System, issued jointly by the U.S. Department of Justice (LEAA), and the U.S. Department of Commerce (Bureau of the Census), an annual survey, covers all facets of the CJS. Unfortunately, this information suffers from an excessive degree of aggregation: the details of the allocation of resources within each of the CJS subsystems are not dealt with. Consequently, the value of the data for resource allocating decisions is limited. 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.
4:2-6

Larson, et al., indicate that a brief review of existing literature on police expenditures and resource usage illustrates what they believe to be common characteristics of similar past work on other aspects of the CJS as well:
4:2-12

1. Studies have to rely on incomplete data bases.
2. 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.
3. The few studies that exist offer little in terms of long-term perspectives on developments in the CJS area.

4. Few comparisons are provided with parallel developments in other areas of public or private activity.

The 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, the true objective of the CJS 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. But, 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." Blumstein has suggested a measure called "social disruption" as a high-level indicator of output for the CJS, but its complexity and its many unquantifiable features preclude operating with it analytically at this time. 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 make evaluation difficult.
4:2-18--2-19

Crime-related measures of output suffer from two major deficiencies:

1. They are only partly measurable in many instances. Surveys clearly indicate that crime is grossly underreported but it is difficult to determine just by how much.
2. Great difficulty is involved in predicting the effects of resource allocation decisions on crime-related statistics.
4:2-19--2-20

Larson, et al., 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, largely because of three problems: (1) accuracy of much of the data. (2) analyses of the data are often incomplete or inaccurate, (3) popular indices about crime calculated from raw data are often inherently inappropriate--opague at best and misleading at worst. A problem of numerous crime-related analyses is that the conclusions are inconsistent with the data that supposedly spawned them.

For homicide, the key problem is not the accuracy of statistics, but of understanding their implications. With other violent felonies, the situation is somewhat reversed. 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.

The deterrent effect of particular measures against crime is a subject of bitter and widespread controversy. Two major concerns are: (1) How much time incarcerated is actually meted out to people who commit a given crime? (2) How much deterrence is associated with a given sentence?

Regarding recidivism, Barnett, Larson, and Odoni note 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.

For recidivism, an annual rate seems more useful than a repeating-probability, yet the probability is the widely-quoted figure. For something as serious as crime, ease of calculation might not be the best criterion for choosing statistical measures. Recidivism rates are but a means to a statistical end.

Larson, et al., quote Suchman, " 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." Another problem, pointed out by Larson, et al, is 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.

According to Lind, the problem of using the impact on crime to measure the effectiveness of police actions is that it is exceptionally difficult to determine what crime levels would be with and without them, which 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. 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. Measures based on public opinion are subject to many of the same difficulties, since many factors other than what our criminal justice agencies do will affect the ratings. 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 CJS as a whole. 45:8-15

Lind, in his study, reaches these conclusions: (1) Even if we had an acceptable measure of the total performance of the CJS, 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. This is because the funding of our criminal justice institutions is fragmented between local, State, and Federal agencies and different units of government have responsibility for different institutions and activities within the criminal justice. 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. (2) 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 workload of other sectors of the system. 45:16-17

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 on measured performance. Several major problems with the crime data that we now have make it almost useless for evaluative purposes and for the purpose of doing basic research on the factors that influence crime: (1) Most of our crime statistics are based on inaccurate and biased data collected and prepared by police departments on the basis of reports to them. For purposes of evaluation, it is simply unacceptable to have agencies that may be evaluated control the source of data. (2) The problems of under-reporting which are well known. (3) 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. 45:18-21

Lind points out that our existing data is not adequate for the job, and the answer is not to try to bludgeon the police into keeping more and better records. 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. 45:22

Lind notes additional problems such as recent data that suggest that variations in correctional experience (length of sentence, treatment, vocational training, etc.) make not the slightest impact on recidivism rates when one controls for offender characteristics. 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. 45:25; App. A, 20

Missing are studies of what really happens in recordkeeping. Well-kept records can be worse than poorly kept ones if they take time for no known purpose. Missing are studies commissioned to survey dark number crimes. Missing, too, might be appraisals of departmental performance, offered by known offenders on the streets, and by samples of citizens. It is the measure of performance by others' standards which is lacking, lacking in the justice system just as it is almost anywhere else. Yet, it is clearly too much to expect organizations to criticize themselves in any serious way, and customer satisfaction or consumerism has not yet reached the justice system. 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. Yet, how do we question these persons and how much credibility do we assign to their testimony? 45:App. A, 17-19

We cannot be sure that the present focus of criminal justice operations is optimal. The direct enterprise of justice is aimed at offenders, their apprehension, adjudication, and disposition. Evaluation is usually

linked to objectives, yet objectives and goals are often multiple and/or unclear. 45:App. A, 10-11

Lind says effectiveness depends on the official mandate to each component of the system and relies on the component's own traditional measure of its product. Each component can claim that its part of the system "works" even though each process need not be in harmony with the other. 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 problem of underinclusiveness. 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. Attributing to a police department a burglary clearance rate without knowledge of the processes leading to the figures is an example of overinclusiveness, or just plain not knowing enough about what is going on within an institution and about an evaluation measure. Underinclusiveness occurs when one fails to include enough in defining an interest area or seeking to comprehend the events leading to the results or interpretation of a measure. 45: App. A, 4-7

Evaluation based on standards which are or are seen as threatening to others can hardly expect to resolve the problems of the CJS if the diversity of interests and views is conceded to be a part of the problem. 45:App. A, 3

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, and a threat arises insofar as evaluation findings may imply the need for specific changes. Resistance both to evaluation and to innovation in public governmental institutions can be extreme. 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. A consistent problem in evaluation is priority setting in resource allocation. 45:App. A, 1-32

According to Lind, our CJS should be viewed as an integrated whole--even if it is not in practice. "Corrections does not correct." Treatment is aimed at the offender, while many of the causes of his crime may be in his environment, which is left untouched. Probation or parole is often a joke. Our recidivism statistics, which are inadequate because they depend on catching an offender an additional time, indicate a measure of our failure. 45:App. B, 12

52:118-120

Nishimura cited three problems with prediction methods:

1. Prediction items become obsolete.
2. Projective techniques are difficult for police.
3. Results sometimes did not agree with experienced judgment.

Regarding the use of failure-rate measures of recidivism, Harris and Moitra point out that, "a good many failure processes are not going to be constant over time." 36:199

The article, Comments on "The Mathematics of Behavioral Change," reviewed the Maltz-McCleary article which questioned whether there are persons who

would never recidivate, asked what does "never" mean in this context, and asked "Do criminals wear out gradually, or do they quit all at once"?

Hirsh discussed the problem of "false positives" - those mistakenly predicted to engage in criminal activity. Criminal conduct tends to have two characteristics which make it resistant to accurate prediction:

1. It is comparatively rare, the more dangerous, the rarer.
2. It has no known, clearly identifiable symptoms.

"Prediction therefore becomes a matter of developing statistical correlations between observed characteristics of offenders and subsequent criminal conduct." Preventive confinement was examined and found to be unsuitable for current use. 78:730-758

According to Nijmegen and Zwaneburg "stability of prediction generally conflicts with four other criteria which prediction instruments must meet: 51:26

1. Simplicity.
2. Efficiency.
3. Reliability.
4. Validity.

which are, too, not very much in line with each other." It might be argued that the concept of probability is in itself not suited for individual prediction. They also cite four problems in the theory of measurement: 51:63-109

1. Representation.
2. Uniqueness.
3. Meaningfulness.
4. Scaling problem.

The statement "the probability that John is going to be delinquent is .75" is not well formulated John is a member of a class of people from which, according to our experience, 75 percent are going to be delinquent means the expectation is that $\frac{3}{4}$ of the total number of people in that class will become delinquent.
51:130-131

"One of the most important practical requirements of prediction instruments is appeal. Appeal can be increased by using potential causes as predictive variables." "Very low probabilities can explain, whereas they cannot predict."

"The use of prediction instruments in order to get information about the possible future delinquency of an individual is incompatible with such phenomenalism (statistical association or empirical data). For this, it is necessary to have a fundamental insight . . . into the social and psychic processes leading to deviancy."
51:173-175

Movahedi and Ogles note prediction in criminology, among other fields of social inquiry, is reduced to the statistical forecasting of the behavior of a class of offenders or of an individual offender and say, "it is argued that probability is a mathematical function defined for classes of events or sequences of events in the long run and as such is not applicable to an individual instance." Subjective probability is "the numerical value of the degree of confidence or partial belief in a proposition estimated by the highest odds that the individual would be willing to offer on the truth of the proposition. Thus, if the odds in favor of Brown's success on parole were 8:2, the subjective probability of Brown's success on parole would be $\frac{8}{8+2}$."
49:177-182

According to Movahedi and Ogles, "A statement ascribing a probability (in a relative frequency sense) to a single event has a fictitious meaning", but "the notion of logical probability seems, on the other hand, to provide a meaningful explication of the probability of a single case," for example: The probability (observed) that offenders with Y and Z characteristics succeed on parole is .72, and Brown has Y and Z characteristics. The conclusion that Brown has a .72 probability of success is not part of a valid conclusion and is erroneous, but does represent a logical relation, since in absence of other factors, we would bet on success.
49:186

An LEAA study noted "one of the major deficiencies in police administration is the lack of adequate performance measures. The need to develop sound productivity measures is becoming crucial with today's budget constraints."

Deutsch and Richards compared victimization and reported crime rates for certain offenses graphically showing the disparate relations as:
22:

inherent criminality and trends in police propensity to arrest people from a particular demographic group. A "problem in measuring the volume of 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 the amount of the behavior."^{7:4-32}

Delinquency in A Birth Cohort noted that total crime is difficult to determine and "at present, we have no satisfactory way of estimating this number with confidence" because of (among other reasons):

1. Crimes which the violator does not know are forbidden
2. Victim unwilling to report
3. Police take remedial action with no report (or some other authority such as mental or social agency takes action and makes no court record)

Also cited is the fallacy of using police reports due to arrests of nonviolators or innocent people.

McClintock noted that as early as the 1830's Bulwar recognized the defect of using reported arrests to measure criminality and Rawson wrote, "crime may abound most where arrests are least numerous." The starting point to study the "dark figure" should be "known crime," since:

1. Committed crimes leads to:
2. Known crime leads to:
3. Crimes with arrest lead to:
4. Crimes with conviction.

Reportable crimes and detectable crimes are more likely reported or known.^{47:}

Biderman noted that in 1966 the President's Commission on Law Enforcement and Administration of Justice cited surveys which "show that the actual amount of crime in the United States today is several times that reported in the UCR." Respondents in a national survey report being victimized by crime more than twice as frequently as would be expected on the basis of the UCR. A Michigan study recorded transactions between the police and citizens, with interviews of a sample of the same citizens several months later. Since over 20 percent of the citizens failed to report the recorded incident, severe underreporting would be indicated by the survey method and that method may have to be regarded "as dipping only shallowly and perhaps inaccurately into 'dark figure' crimes." Victimization rates from the survey appear to be higher than UCR rates for two reasons:^{8:17-32}

1. Failure of citizens to report to police.
2. Failure of police to report in UCR.

Krohn, Waldo, and Chiricos reported regarding the utility of self-reported crime data, that Gold used informants as an external check of self-reported criminality and found 72 percent of his sample of subjects could be considered "truthtellers." Clark and Tift used polygraph examination as an external validity check on questionnaire data, and found "that all respondents underreported the frequency of at least one behavior, and one-half of the respondents overreported on at least one behavioral item. However, the overall proportion of correct answers was 81.5 percent." Defleur criticized their methodology, reducing the magnitude of their findings. In interviews, responses are affected by:

1. Friendliness of interviewer.

2. Social distance between interviewer and respondent.
3. Threat which the interviewer represents to the respondent.

For seven of eight offenses, checklist self-reports elicited a higher rate of admitted delinquency than interview self-reports. 43:546-550

Hood and Sparks, noted two strategies to assess "dark figure" or hidden crime: 38

1. Question the general population about criminal acts they've committed. According to Dentler "it should be discontinued because the method seems too shaky and the results too equivocal to deserve further effort."
2. Victimization surveys.

Both are "liable to distorted and untruthful answers."

Chaiken, et al., cited data problems saying although examples were found of failures to implement a program, project or model, because the user agency was unable to understand the programming language or the conceptual foundations of the model, the main model attribute that proved to be an obstacle to implementation was a requirement for data that was unavailable to the agency. The agency characteristics found to be obstacles to implementation were as follows: 15:xii

1. The introduction of a model is generally not undertaken in response to some pressing need or problem to be solved, and other matters considered of greater importance can divert resources or personnel from development and use of the model.
2. Very often a single advocate in the user agency saw the need for a model, conducted a search for the appropriate one, sponsored his choice before agency administrators, and pursued implementation. Progress then depended

on the advocate's judgment, continued attention, and political skills. Vulnerability to changes in personnel increases as time elapsed on a project increases.

3. The lack of professionalization among the planners. The agency's personnel did not have advanced training, a tradition of using any kind of analytical techniques, or a world view that extended beyond the immediate organization. This problem is a far-reaching one, extending beyond modeling per se, and touching on the current capabilities of criminal justice agencies to support a competent planning process.

The potential value of models to indicate the types of information and data that are needed for management purposes is not being fulfilled, because

the implications of models for management information systems has not been summarized in a form accessible to the designers of such systems.

Many criminal justice planners and operating agencies are uncertain about the circumstances under which models can be useful, whether an appropriate model already exists for handling a particular problem, and, if so, which one would be best suited to their needs. As is the case with most models designed for governmental planning purposes, criminal justice models have not been used to as great an extent as the model builder might have hoped. 15:xiv-1

Patrick V. Murphy, former police commissioner of New York City, said "because there is very little exchange of experience and ideas among departments, standards of appropriate performance are slow to emerge." He went on to say "it is a misuse of UCR figures to draw from them implications about the productivity of a police department," and "UCR data do not accurately

portray either the nature or the extent of the crime problem in any given municipality." The Police Foundation's 1975 Readings on Productivity in Policing cited measures requiring significant additional data-gathering: 61:37-98

1. Crime rates including estimates of unreported crimes based on victimization studies.
2. Clearance rates including estimates of unreported crimes based on victimization studies.
3. Percent of felony arrests that "survive" preliminary hearings in courts of limited jurisdiction.
4. Percent of arrests that lead to convictions.
5. Average response times for calls for service.
6. Percent of crimes solved in less than "x" days.
7. Percent of population indicating a lack of feeling of security.
8. Percent of population expressing dissatisfaction with police services.

The same "Readings" noted problems with existing measures and ways to alleviate these problems with newer measurements. Crime prevention-deterrence measures are the first concern. Communities hope crime prevention and deterrence are a major impact of police crime control activities. It is extremely difficult to determine how many crimes police activity has prevented. What is done as a practical matter is to measure the number of crimes ~~that~~ have not been deterred. Victimization surveys rely on memories, and willingness to respond, of those sampled, and are therefore subject to errors. The data from a victimization survey can be used directly to give an estimate of total victimization, but only an estimate. Critics have proposed various weighting schemes, such as the

Sellin-Wolfgang weights, or some modifications of them. The ideal measure of crime prevention productivity, "the number of crimes prevented per man-year," is not feasible given the current state of the measurement art. 61:98-101

Apprehension of offenders measures are the second type considered. The measures of output on apprehension customarily used are the number of arrests and clearance rates. Important problems exist with the current definitions and data collection procedures used for each measure. The number of arrests per police man-year at first glance may seem to be a very attractive measure of police productivity, but the mere fact that an arrest was made does not mean that the person committing the crime was successfully brought to justice. As a step towards evaluating the quality and effectiveness of arrests, and at the same time to reduce the likelihood of encouraging undesirable arrests, the use of the measure "number or percent of arrests that pass the first judicial screening," a productivity measure in the classical form could be used. Also, the procedures on how to count reduced charges and multiple charges on an arrest need to be specified. Measures reflecting arrest dispositions have current drawbacks. First, the data are not currently generally available to police agencies. A second problem is that the reasons for arrests not surviving the first judicial screening 61:102-104 can be quite diverse, and many of them may be unrelated to police actions.

Regarding the use of clearance rates, concerns have been expressed at various times about such problems as variations in what constitutes a clearance and the variability that can occur when a police department emphasizes or does not emphasize exceptional clearances. As with arrest rates, the counts of clearance used in most jurisdictions include incidents for which an arrest

was made, regardless of whether the arrests survive the initial judicial screening. A clearance (as defined by the FBI) is recorded by the arrest of any one of several criminals committing a crime even if the others are never apprehended. There is currently no measure which indicates the success of the police in identifying and apprehending each of the offenders involved in a single crime. Two variations of this measure seem appropriate. One would emphasize the total solution rate: the "total number of man-crimes for which someone was successfully brought to justice." The second variation would focus on the risk to the criminal. As practical matter, for many crimes, such as burglaries, it is not possible to know how many offenders actually participated. Police departments and their communities currently do not have full information on the number of successful apprehensions 61:105-106 relative to the total number of those who should be apprehended.

Productivity measurement in such terms as "percent of police time spent on productive activities" is likely to be a measure of major concern to internal police management, but difficulties abound in defining what are, and what are not, productive activities. There are also problems in calculating the amount of input: 61:109-110

1. A major question is what to do about police resources applied to noncrime services, especially in the common situation where the same police officers carry out both crime and noncrime functions.
2. A similar problem exists for providing data on specific crime control activities, such as specific police units. The same police employees may routinely switch from one crime to the other.
3. Overhead, support costs, and costs of equipment are also problems.

Supervision, employee fringe benefits, vehicle maintenance, and equipment and vehicle purchases need to be considered when measuring output against total cost.

Sellin and Wolfgang stated, "the UCR method provides no solution for the problem of how to deal statistically with a complex of offenses or with simple offenses that vary in seriousness but carry the same legal title," but "information possessed by the police about violations of the criminal law remains the best source of data for index purposes." 67:294

According to Bottoms and Nilsson, information is one of the most valuable commodities in an organization since quality decisions are crucially dependent on it. Needed is a requirement study so that the kind of information for operational and management decisions can be provided in a timely manner. 12:26

Deutsch, et al., noted the use of crime-rate indices as performance indicators presumes that:

1. Crime rate data accurately reflects true victimization.
2. Changes in CJS effectiveness account for nearly all changes in victimization.

It can be shown that available research does not support these assumptions. Larson, et al., said "A panoply (an array) of statistics about the operations of police and courts are floating about; perhaps too many for citizens to absorb." 4:3-17

But how good are these statistics? The FBI, in its Uniform Crime Reports 75:43 admits that crime based on police detection is understated.

The Rand Corporation made a study involving self-reported criminality which involved a small sample and may have been a poor prediction of the entire criminal population. Their sample was stratified as to:

1. Age.
2. Geographical history of respondents.
3. Losers (people in prison for at least 2nd time).
4. Sex (all males).
5. Present reason for confinement.

The interview questions may be such that the sample with mean education levels ≤ 8 years could not understand. Also the questionnaire may have been (2 hours) too long to keep the respondents' attention. And then does the assumption that a criminal is a rational being really apply?
18:

The Rand report stated that, "even in an offender sample as small and select as this, the dominant finding was diversity," their "data emphasizes that arrest records do not suffice in distinguishing among the more serious and the less serious habitual offenders" and demonstrates "poor correlation between offenders' actual behavior and their arrest records" making the predictive value (of record alone) weak. The report also noted that official data is biased because:
18:114-156

1. The lack of uniform methods of reporting and recording crime data.
2. The unevenness of law enforcement with respect to different racial and socioeconomic groups and geographic regions.
3. Their being limited to offenders who become involved in the legal-reactive process.

Beck and Hoffman found that data did not support the theory that longer prison terms will result in dramatically higher recidivism rates, and noted:
6:203

1. All actuarial devices predict outcomes for groups and not for individuals.
2. Actuarial devices may overlook other elements such as attitude.
3. Actuarial devices are based primarily on information found in the inmates' institutional files, often found to contain inaccurate or even contradictory information.

A major obstacle in trying to model crime systems is the lack of accessible detailed data for validation of the basic assumptions, according to Avi-Itzkak and Shinnar, while Riccio noted "whether or not hard and fast laws of human behavior even exist is a matter of debate", and it is difficult to isolate phenomenon to be studied from interactions and interferences.
3:212

Former Attorney General Ramsey Clark said, "to think of controlling street crime while organized crime flourishes is to ignore their clear connections," and "most crime is never reported to the police. And much crime is inaccurately reported. Erroneous crime statistics are often used to create the impression that the new chief is doing a good job, or to support a movement to add more police. Frequently an apparent increase in crime really reflects on improving effectiveness in law enforcement, or in the reporting of crime itself." "Efforts must be made not only to secure full reporting of crime, but to be sure reports are truthful and accurate."
Clark also pointed out that to "assume only half of all violent crime comes to the attention of the police is a reasonable assumption."
16:37-49

Showing the lack of data, the National Opinion Research Center (NORC) 1965 survey showed more than half of all crimes and 38 percent of index crimes were unreported. Glaser said both the President's Commission on Law Enforcement and the Administration of Justice in 1967 and the National Commission on Law Observance and Enforcement (the Wickersham Commission) in 1931 noted the difficulties in measuring the costs of crime. One major difficulty is that crime costs, although real, are implicit. The Wickersham Commission in 1931 stated, a just summary for today -- "we do not today have adequate criminal statistics" . . . and pointed out "even where crime is suspected one cannot say legally that a crime has been committed until the existence of a delictual fact has been passed upon by a court or jury," even further clouding the issues of crime data, crime measurement, and criminal justice agency and system evaluation.

32:686,1056

79:88-155

As previously noted, the lack of sound, adequate performance measures is a major deficiency in the CJS. Purported measures such as arrest rates and reported crime figures have proved inadequate for management purposes.

63:3

Research reveals several recommendations or proposals to improve evaluation and productivity measurement in the CJS, some of course having more merit or utility than others.

4:2-26--3-18

For example, Barnett, Larson and Odoni made the following recommendations:

1. The LEAA should sponsor studies and activities aimed at:
 - a. Identifying useful intermediate output measures such as 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.
2. The LEAA 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.

Devising and calculating systemwide performance measures should be a high priority for the LEAA.

Larson, et al., go on to argue the need for certain programs and believe that LEAA should adopt as long-range goals:

4:3-21--3-22

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 correction systems. 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.
3. The development of a set of statistical testing procedures to measure quantifiable effects of innovations in the area of criminal justice.

45:35-39

Lind proposed that:

1. LEAA should review the expenditure of funds to determine what proportion of the funds spent in support of the CJS are spent on sophisticated research and evaluation, compared to the research expenditures as a function of the support to and cost of other human and social problems (e.g., national defense, health, or education). The findings may suggest new allocation priorities.
2. 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.
3. LEAA should review its expenditures and sample representatively from among agencies to identify the conditions which characterize successful versus unsuccessful programs.

4. With respect to the objectives of the CJS, a study should be undertaken which seeks to array the objectives which exist for the various components in the system, including formal and inferred goals.
5. With respect to conventional output measures of the system used to infer performance, 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.
6. 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.
7. Work should be done to develop methods for finding and eliciting reports from victims, both those experiencing dark number crimes and reporting victims so that 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.
8. 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, capable of routine use to assess the dark number, police-court-corrections adequacy, and to contribute to recommendations for improved crime protection and efficiency in the justice system.

9. With regard to the description and definition of organizations, events, people and processes which interrelate with the governmental institutions ordinarily defined as the CJS and to influence outcome measures, an expanded mapping is in order (including internal studies of the process through which people enter the system and studies of the events and outcomes associated with sideways referrals out of the system through diversion).
10. 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 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 nonpublic institutions.
11. Past evaluations have produced important findings with regard to the effectiveness of system components, and suggest that experimentation must take place within traditional components of the system which provides for the evaluation of major innovations.
12. Methodological studies are in order to consider the yield of various statistical tests, and to show intercorrelations among various operating and outcome measures. 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 to compare different research methods for their yield, and to introduce alternative product measures.

45:39

A commitment to evaluation is a commitment to genuine research. Maltz wrote on program evaluation that one of the most important determinants of the objectivity of the evaluation is the attitude of the heads of the agency running the program and an outside evaluator is usually considered more impartial than one coming from within the agency. One of the primary roles of an administrator is to evaluate the efforts of his agency. The goals of the program determine the criteria which are used to measure its effectiveness, but a program might be beneficial in some unforeseen way, wholly outside the original criteria. 46:14-33

According to Maltz, programs aimed at controlling crime should not be evaluated solely for their effect on crime, and most programs cannot, by their very nature, focus on one specific objective alone. Evaluating how well a program achieved its goals is not the only purpose of an evaluation; how and why the results were achieved are of equal importance. External measures relate to the former evaluation; internal measures are concerned with the latter, and many crime control programs are dependent on good public or community relations in order to achieve their goals. 46:33

The evaluator's style relates to the need to maintain liaison with the persons involved in the program's operation. The police department's style affects the transferability of the program. The extent of program transferability can be determined to some extent by the validity of the assumptions which were made to justify the program. An evaluation should not be conducted at arm's length from the agency or program, or from the vantage point of an ivory tower. Evaluations can fail when the evaluation team does not

maintain a strong and continuing liaison with the agency running the program. In programs conducted in police departments, the support of the police chief is vital to the success of the program and the evaluation. Agency administrators may look upon evaluation efforts with suspicion, concerned that someone is checking up on them; thus the maintenance of strong and continuing liaison with the agency administrators is a necessity to ensure a viable program and evaluation. Agency coordination should not be restricted to the top levels. The patrolman who implements the program should be asked his views on its effectiveness, as should the field supervisor. 46:43

Maltz recommended that the program planning include a description of the assumptions and logic underlying the choice of the program. During the course of the evaluation these assumptions should be tested and verified. Other similar problems in statistically relating cause to effect can be described and the evaluator should strive to identify the mechanism which relates the two and should explain discrepancies in the logic underlying the program. 46:46

Maltz proposed a program of general evaluation framework: 46:47-48

1. Develop the program rationale.
2. Select the evaluation team.
3. Select areas for implementing the program and for control.
4. Choose external measures of effectiveness and internal measures of effectiveness.
5. Determine data requirements (with quality control checks) for the measures of effectiveness and the displacement effects, compensating for the inadequacies of official crime data.

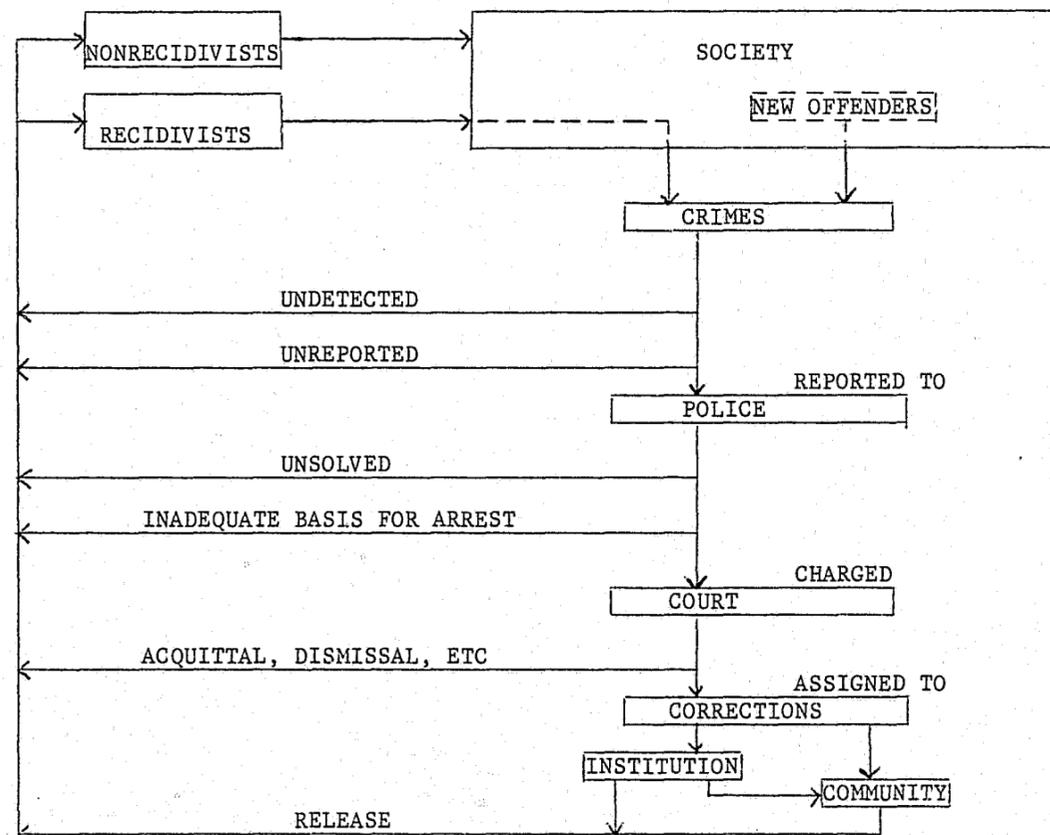
6. Develop baseline data and information for the experimental, control, and boundary areas.
7. Collect and analyze data after a short period of operation, and develop preliminary results.
8. Modify the program, the assumptions and rationale, the data collection procedures, and the measures of effectiveness, as necessary.
9. Complete the collection and analysis of data and information, and develop and interpret the results.
10. Verify the program rationale in light of the findings.
11. Describe the permanent changes that have resulted from the program.
12. Determine the transferability of the program and recommend the best means to effect the transfer.

Maltz concluded that deficiencies in the available data present some significant problems in crime control evaluations, but they are not insurmountable. Monitoring the data quality, more careful analysis of the data, and the collection of additional data will minimize the problems, and achieve maximum utility. Among the more important considerations in conducting an evaluation is the need to maintain strong liaison with groups within the police department which are affected by the program. Assumptions and logic which were initially used to justify the program should be tested and verified during the course of the evaluation. 46:57

An Interpol study showed that research into evaluation methods has played hardly any part in determining crime prevention priorities and that it has produced no results which can be applied in practice. Crime prevention work with a reliable scientific basis would be more likely to succeed and

to achieve better results from limited resources than work based on plausible assumptions and practical experience at best or - at worst - on vague beliefs and administrative traditions that have never been questioned. 39:102

Blumstein and Larson advocated modeling criminal justice as a total system and in their analysis, considered sensitivity analysis as well as other subroutines. They said, "one central problem in improving law enforcement is the need to examine the total CJS, comprising police, prosecution, courts, and correction agencies, in an integrated way," and saw the total Criminal Justice System as: 10:234 9:vii-4



Chapman says that what is needed is both an integration of the theories of crime and police response into a coherent model and an empirical testing of that model. 14:48

Nijmegen and Zwaneburg define a model as a device to represent a large and complex set of objects, their attributes, and interrelations in an ordered pattern with the primary function of simplifying reality in order to make it accessible for scientific study; also a system, the elements of which represent empirical objects. Qualitative models are verbal representations of reality, whereas the globe or the maps and diagrams can also be seen as quantitative, if their representing qualities are such that there exists a scale relationship between them and reality, and a mathematical model is a representing system, the elements of which are of a mathematical nature. 51:105-107

They define measurement as allocating numbers to aspects of things which are in themselves nonnumerical. Events or outcomes can be said to have a probability of occurrence. Some events out of the universe of events may form an equivalence class. The probability of this equivalence class or subset of the universe of discourse is then some function of the probabilities of the elements of that particular subset. Such make some form of prediction possible. 51:108-115

Chaiken, et al., in discussing models and their use, say the heart of any attempt to analyze a situation or issue and make a rational decision is the existence or creation of a device or procedure to provide insight

into the consequences of any decision that might be contemplated. That device or procedure, termed a "model," is a simplified representation of whatever part of the real world is important to the issue under study, one that can be manipulated to forecast or at least give some clue as to the outcome that is likely to follow a particular action. A model might be purely verbal or a simple diagram. No matter how it is represented, a model is designed to help a decisionmaker make a better decision than in its absence he might otherwise make. The adequacy of a quantitative model depends on how it captures the essence of the issues and how well the numerical values it requires can be estimated. He cautions against oversimplifying or overcomplicating a model.
15:3-5

A decisionmaker faced with a problem needs to develop a good idea of his objective or what it is he wants to accomplish and, if others are involved, communicate it correctly, and to seek out various alternatives, options, or actions that appear to offer some possibility for attaining the objective. To forecast or estimate consequences, models are used. Additional models may also be used to compare and rank alternatives, although this is often done intuitively.
15:5

According to Chaiken, et al., in the physical sciences, models can be made to describe the problem under investigation so faithfully that the results obtained from them can be accepted as completely valid for all practical purposes; but where behavioral, political, and social factors play a large role, we have to base our calculations on, and

supplement our model results with, a great deal of judgment. Models and model building provide guidance for that judgment, and reliance on judgment and intuition is crucial to every decision. A great virtue of models and model building is that they provide a systematic, explicit, and efficient way to focus the required judgment and intuition, particularly that of experts and specialists on whom analysts must usually depend for practical knowledge and experience.
15:6

An explicit model, quantitative or not, introduces structure into a problem, enabling involved decisions to be broken into constituent parts, that can often be considered one at a time. The model provides an effective means of communication, and feedback to guide the participants in refining their earlier judgments. By "exercising" the model and testing for sensitivity, information can be generated that may lead the users to alter their original judgment, and even to intuit a solution in spite of deficiencies in the calculations. In an area such as criminal justice, the model builder is likely to find a situation where the relationships between the elements are very imprecisely known and little data exists for determining them. His approach is to select certain elements as being relevant to the problem under consideration; to make explicit, where known, the relationships between the elements selected, and to conjecture the nature of other relationships that he judges significant. His model is thus likely to be ad hoc and tentative, subject to modification and improvement as new information and insight become available. We should not look at a model merely as a "black box," a device to provide a route from

a set of hypotheses to a prediction about the real world. So narrow a view ignores a most important product of the modeling process: the insight into the problem it can provide. 15:6-7

For public policy problems the models most used, on the whole the most useful, and most often the only sort even considered by analysts, are quantitative models that resemble the "scientific" models developed in the physical sciences, consisting of a system of logical relationships that attempt to express the processes that determine the outcome of alternative actions by means of a set of mathematical equations and/or computer programs. Quantitative models divide into two categories:

1. Analytic models, the outcome or solution is extracted from the model by mathematical analysis.
2. Simulations, the outcome is estimated by means of a series of imaginary experiments on the model.

A model would be strictly quantitative if the situation or activity under investigation was represented by that model so faithfully that a decision could be made solely on the basis of the results obtained from the model. Few real-world issues are susceptible to resolution by such a completely quantitative treatment; almost always, judgment will be needed at the end as well as earlier. Unfortunately, many criminal justice problems cannot be handled satisfactorily or even approached sensibly by means of quantitative models. At its simplest, an analytic quantitative model, once set up, may involve no more effort than the substitution of numerical values in a mathematical expression

or formula and a little arithmetic. In somewhat more complicated analytic models, the form of the desired equation is established by the model builder, but the equation includes some constants whose value is not known in advance. These constants, called parameters, may vary and are determined from appropriate data. Another type of analytic model, called an optimization model, is still more complicated, but the user does not have to try every possible value of the decision variables to see which results look best to him. 15:7-8

Simulation is the term applied to the process of modeling the essential features of a situation, and then predicting what is likely to happen by operating with the model case-by-case; i.e., by estimating the results of proposed actions from a series of imaginary experiments. A great advantage of computer simulation for investigating complex problems is that a digital computer can be used to represent, with precision, processes for which satisfactory analytic approximations do not exist. Typically, a real system is subject to chance elements; these can be taken into account in the computer program by the use of random numbers. Simulation with a high-speed digital computer is a powerful technique, but the ease with which a simulation can be put together makes it tempting to employ the technique where insufficient data exist to justify such a model. 15:8-9

An operational game is a simulation involving human participants acting as simulators for at least some aspect of the problem, and is an outgrowth of military war gaming. Gaming is now widespread, but the extension

to the investigation of public policy problems is in its infancy. Gaming was originally developed to investigate the problems of a decisionmaker whose actions might be countered by those of one or more intelligent opponents, and is an approach one can use to tackle problems of that wide class for which no satisfactory quantitative model can be constructed. By allowing for the introduction of judgment at every step, a game provides the opportunity to take into account intangible factors often considered completely beyond the scope of analysis - courage, cooperation, commitment, and morale, for instance. 15:9-11

The use of a committee or panel to provide advice on a decision or policy is a time-honored, well-established, and much used procedure, open to a number of well-known objections and it often leads to very biased and ill-considered recommendations. A number of ways to improve the procedure by structuring the discussion have been suggested, the most promising of which, other than gaming, appears to be the Delphi approach, according to Chaiken, et al. Delphi is a procedure for arriving at a forecast or estimate by eliciting and refining the opinions of a group of people by means of a series of individual interrogations. Since it can serve the same roles as a model, providing insights into or predictions about a contemplated action, a Delphi procedure can be considered an extended form or at least a replacement for the standard representative model. The accuracy of Delphi estimates and predictions is generally greater than that obtained from unstructured committee discussion, but Delphi is not a substitute for an analytic model or simulation unless

one feels so little confidence in their validity that he is willing to depend on committee judgment instead. The Delphi approach is characterized by three simple ideas:

1. ANONYMITY. Originally, it was by a written questionnaire but recently, with increasing frequency, by online computer console. In determining an estimate or predictions, the responses are not matched with the respondents.
2. ITERATION AND CONTROLLED FEEDBACK. After each round of questionnaire, all or part of the information generated in previous stages is fed back to the participants in order that they may use it to revise their earlier answers.
3. STATISTICAL GROUP RESPONSE. Although the group opinion tends to converge with feedback, the normal outcome is a spread of opinion even after several iterations. Rather than making an attempt to force unanimity, some form of summary statistic, usually the median, is used to represent the group response.

Delphi is not an opinion polling technique. Its purpose is not to furnish the investigator with data about the respondents but, rather, to estimate the answers to questions for which there is no well-defined way to find a definitive answer at the time of the exercise. Delphi techniques offer a way to introduce a systematic approach to problems where conventional models cannot be formulated. 15:11-14

Models may be strong in some aspects, weak in others, useful for one policy question but totally irrelevant for a closely related one.

They can be appraised or compared only in the context of a particular policy decision and a host of other considerations. No hard and fast rules can be given for appraising models, only questions that should be raised and judgments that should be made. The first step in acquiring a model is to identify the policy issue to be addressed by using the model. Once the policy issue has been identified, the following types of questions should be asked:

1. Is this problem of interest to my agency or to some group that is benefitted by my agency?
2. Why is it of interest?
3. Is the problem formulated properly, or is it just a symptom of a much larger and deeper problem?
4. Why has the problem not been tackled or solved before?
5. Is there any reason not to tackle the problem?
6. If the analysis can be carried out successfully and advice provided, what will be done with the results?
7. Will anybody be able to act on its recommendations?
8. Is the inquiry politically sensitive?
9. Is it likely to commit the agency to continuing support?

Essentially, one wants to find out, before the tedious task of working through an elaborate model to determine if its predictions are correct, whether the results are likely to be worth the effort, and if they are, whether anyone is willing and in a position to do anything with them. ^{15:14-15}

The specificity of the problem definition has much to do with the desired scope of the model. If the problem is vaguely stated, then the model should be flexible and have multiple capabilities. The nature of the decisions to be made with the model should also be considered, and it is important to distinguish between one-shot and recurring decisions. In the case of recurring decisions, it is important to consider how often the model is to be used. In some cases the decisionmaker knows not only the problem to be addressed but also the solution he plans to propose. He anticipates that a model will confirm his decision and assist him in persuading others to adopt his plan. This is a legitimate use of models, but one should be prepared for the possibility that the outcome will not be as expected, and the model's results should persuade the decisionmaker to adopt some alternative solution. Designing a new model can be a time consuming process. Even with existing models, months or more may pass before the program works properly and appropriate data have been collected. ^{15:15}

Model builders are likely to be concerned with issues of technical quality that are of little interest to an administrator. The model should be non-trivial, powerful, and efficient. For optimization models, there is also a technical distinction between algorithms that are guaranteed to find the optimal value of the objective function and heuristic algorithms, which yield good, but not necessarily optimal, solutions. Every model goes through a period in its development when programming bugs or other errors cause the model to malfunction. As long as there is some means for

correcting errors when they do occur, the user should not judge the overall technical quality of the model by such isolated occurrences. 15:16

Verification and validation are two aspects of quality that a potential user can check for himself, and should. A model is said to be verified if it does what the model builder intended for it to do. Typically, a model is verified by testing it with sample data that correspond to known output, by setting some of the data input to extreme values, or by holding some of the variables constant to determine whether the output changes in anticipated ways as the other variables are changed. In many cases a model is verified by checking its output against results provided by previously verified models. Validation refers to checking that the outputs of the model conform to reality. In some cases the model's output can be compared to data from historical sources or from an experiment conducted for validation purposes, but easy validation 15:16-17 is the exception rather than the rule.

Models differ greatly in the amount and level of detail of data required. If two models are equally satisfactory for answering the policy issue at hand, and one requires less data or more readily available data than the other, then it is to be preferred. In many instances, one of the most useful functions of a model is to focus an agency's attention on the types of data that will best serve subsequent management purposes. The questions of cost arise at several levels:

1. Designing or acquiring the model.
2. Collecting data.

3. Operating the model on a computer system.
4. Analyzing the output of the model.
5. Implementing the decisions arising out of the analysis.

Only fairly complex simulation models entail computer costs large enough to be a factor in whether or not to use the model. Some models can be operated by persons having little or no technical training, but others require the assistance of specialists in a particular programming language or statistical technique. 15:17-18

Chaiken, et al., noted that some criminal justice models apply to only one of the major components of the CJS: police, courts, or corrections. A model of the overall CJS must integrate the three components, not only modeling the behavior of each component itself but also including the interaction among them. Interactions among the elements of the CJS are really of many types, but models have focused on offender flow, and the primary interactions here consist of down-stream effects and feedback of recidivists. Overall CJS models are useful tools for planners even though there is currently no organizational structure with single management control over a total CJS. The President's Commission on Law Enforcement and Administration of Justice recommended that a closer relationship be developed among the elements of the CJS, and funds to stimulate this were provided by the Omnibus Safe Streets and Crime Control Act of 1968. 15:20

According to Chaiken, et al., the term "model" refers to a device or procedure for providing insight into the consequences of a decision, and models have failed to achieve the level of use for policy decisions that was intended by the model builders and those who funded them in the CJS. While, in principle, models can be designed to assist policy-makers in nearly any kind of decision, in practice no one would take the effort to use a model unless the decision presents difficulties such as one of the following:

1. So many alternatives are available that it is not practical to consider each one before selecting the best.
2. The consequences of each alternative are too complex to be anticipated with assurance.
3. Numerous tedious calculations must be performed to evaluate each alternative.
4. The decision must be performed rapidly following specified rules.

In such situations a model can provide vital information that otherwise would not be available for making the decision and produce clear documentation of the decision process that can help persuade others of the correctness of the policymaker's position.^{15:v}

No models can tell a policymaker exactly what decision he should make in a given situation. Models must be used with common sense, good judgment, and understanding of political and budgetary constraints to make decisions. Some models, descriptive in nature, do not even pretend to suggest any policy recommendations; they simply provide

a tool for anticipating the consequences of policy changes invented by the user. Other models prescribe a "best" solution to a specified problem, but even here the user often has a choice of how the term "best" is to be defined, and he always has to use his own judgment in weighing performance characteristics not encompassed by the model builder's definition of "best." Policymakers must know how to appraise particular models to determine whether they are suitable. Most important is the match between the model and the policy issue to be addressed.^{15:v-vi}

Technical quality of a model is often difficult for a policymaker to judge, but evidence that verification and validation of the model have been conducted should serve as adequate assurances of quality. A validated model is definitely to be preferred over an unvalidated one.

Another important characteristic for appraising models is the amount and nature of data required. The cost of a model is important in terms of the types of personnel needed to use the model and the length of time they will have to work with it before decisions can be made from the output. The mode of operation of the model is often considered to be important. Some models are interactive, meaning that the user sits at a terminal and enters information directly into the model via his keyboard; the output appears immediately at the same terminal. Others operate in batch mode, whereby instructions to the program are prepared on cards or a similar input medium and the output emerges later on a high-speed printer. Examples of previous implementation and use of a model are also helpful in appraising it.^{15:vi}

The basic types of models of interest in the criminal justice field are as follows:
15:vii

1. Analytic models determine an outcome or solution from mathematical analysis, such as solving a set of equations. Generally, many features of the system to be studied are ignored or simplified in an analytic model, but the results may nonetheless be accurate enough for policy decisions. One type of analytic model is an optimization model.
2. Computer simulations imitate the operations of a system so as to produce the same statistical behavior as found in the real world. Simulation models can, in general, capture more detail of actual operations than can analytic models, but they may be more expensive to use, and data collection may be more difficult. Simulation models are always descriptive; they tell the policymaker what will happen if he makes a certain decision but do not suggest any decisions to be considered.
3. Operational gaming is a form of simulation in which human participants imitate some aspect of the real world.
4. Group judgment models are structured procedures for obtaining forecasts or estimates from a group of people. An example, called Delphi, involves using anonymous feedback of statistical information about the previous estimates provided by the group, until a consensus or firm disagreement is reached.

Models of the entire CJS have been developed as part of, and as a consequence of, the work of the President's Crime Commission in the

mid-1960's, focused on the flow of offenders through the various components of the CJS: police, prosecution, courts, corrections, and parole. These models have been useful to planners for anticipating the effects of policy changes in one part of the system on later changes elsewhere.
15:viii

One model, called JUSSIM, has been the central development in this field, having spawned a number of variants with other names. Individuals, both recidivists and new offenders in society, perpetrate crimes. Some are detected, some not; some reported, and some not. Reported crimes are processed by the police, arrests are made, and some of the arrestees are charged with a crime. These arrests become cases to be processed by the courts, and those convicted may be sentenced to the corrections subsystem. Parole and eventual release return individuals to society. A fraction of these, plus a fraction of those released from other parts of the CJS, inevitably commit crimes again. The model considers groups of these individuals, distinguished perhaps by crime type, age, sex, or other characteristics relevant to how they will be processed by the system, and calculates and displays new measures of workload and other information that permits the decisionmaker to anticipate the consequences of the proposed policy change.
15:viii

One overall CJS model, DOTSIM, is a case-by-case simulation that follows each individual offender through the system, but it has not been accepted and used to the same degree as JUSSIM and its descendants. The primary value of overall CJS models to date has been to train planners to understand the interactions among different parts of the system and to

focus their data collection efforts on information having clear value
15:viii-ix
for management purposes.

Nearly all models for police applications have been directed at patrol
forces, and can be used to analyze policy issues of the following types: 15:ix

1. Determining the total number of patrol officers a department should have (e.g., during budget preparation).
2. Allocating a fixed total number of officers among geographical commands.
3. Determining how many officers in a command should work each tour or shift.
4. Determining the hours at which tours or shifts should begin.

One of the earlier models, LEMRAS, provided the user with the capability to predict how many calls for service would be received at different times of day from various locations, whereas Rand's patrol car allocation model (PCAM) has both descriptive and prescriptive capabilities. In descriptive mode it calculates performance measures for any allocation proposed by either the user or the program itself. In prescriptive mode, PCAM can specify the minimum number of patrol cars that must be on duty to meet standards of performance established by the department. 15:ix

Most court and corrections models are similar to the models of the overall CJS; that is, they estimate characteristics of cases or offenders moving through various stages of processing, or they calculate data needed to estimate offender flow characteristics, but ordinarily have a greater level of detail for the relevant subsystem than would an overall CJS model. 15:xi

Chaiken, et al., conclude, in general, criminal justice models have failed to achieve any notable level of use for policy decisions. The explanation for this discouraging history lies only partly with the characteristics of the models themselves; primarily it rests with characteristics of user agencies and the interactions between model builders and user agencies. They recommend a Federal center be established for the purpose of making documentation and computer programs readily available. The personnel of such a center should be capable of identifying which models (if any) meet requester's needs, and a list of organizations that have already used each model could be maintained by the center to provide a starting point for further inquiries. 15:xii - xiv

Indicators of successful model implementation were found by Chaiken, et al., to be:

1. A clear and realistic understanding at the start of the project of the policy issues to be addressed and the time frame over which results would be obtained from the model.
2. The availability of suitable written documentation of the model oriented to the user.
3. A direct personal contact between agency personnel and the model builder or one of his associates.

They note criminal justice modeling is a young field that has demonstrated value for training planners to understand agency operations and interactions with other agencies, but has had little impact on policy decisions

to date, and propose a reasonably sustained effort to encourage implementation of existing high-quality models over a several-year period which should give a clear indication of whether models can serve a useful function in the criminal justice planning process. The development of new models should be encouraged and an effort should be made to institute some form of peer review in the model-funding process. Funding agencies should concentrate some efforts on testing models in a variety of jurisdictions and developing clear documentation in the form of user's manuals and case studies of implementations that failed. 15:xiii

Barnett, Larson, and Odoni observed that in performing an experiment in part of the CJS, the entire jurisdiction under consideration serves, in effect, as a model or laboratory and the number of "actors" and resources used is usually large. The experiments tend to be extremely expensive and time-consuming, and are nonrepeatable under identical conditions. The quantitative models developed in recent years can play an important role in assisting in every phase of experiments, but 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. Quantitative models provide a unique tool for producing in detail the a priori (based on theory) predictions necessary for the successful conduct of the experiment, for example, multiple offenses can be modeled as a Markov process to predict the future criminal career profiles of offenders. 4:4-2--4-4

Barnett, et al., postulated 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 framework would focus on three distinct phases of the experiment:

1. Design.
2. Execution.
3. Evaluation.

Some of the insights gained from the quantitative modeling component of the evaluation could shed light on results of the qualitative evaluation. 4:4-8--4-13

They also recommend developing a methodology to incorporate feedback from the experimental environment to affect the experimental design, using tools from operations research, such as dynamic programming, decision tree analysis, and Markov decision processes, to plan for contingencies prior to implementation of the experiment and to assist in structuring an evaluation plan that anticipates adaptive changes in the experimental design. The resulting experimental design would be a matrix of contingency plans. 4:4-15

Some uses of quantitative models in day-to-day evaluation, could be: 4:4-16

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.
3. Evaluating the utilization of the jury pool called up in a particular month.

The Criminal Justice Symposium focusing on police productivity pointed out that the absence of complete measures of output does not eliminate the possibility of effective productivity improvement, and that the first prerequisite for effective work on productivity is the development of data on workload and current operations.
59:5

The answer to the problems of the economic crunch and the development of better measurements can be the concept of productivity. Police officials who use productivity measurements properly as an in-house management tool should get maximum return from their resources. Productivity measurement in police work deals mostly with service rather than products, and should assist officials in determining, and then presenting for budgetary purposes, actual needs after the "fat" has been trimmed away. Concern for productivity based on a defensive managerial response will not have the same impact as programs developed through affirmative action.
59:18

Managers should reach out for employees and union involvement and cooperation in improving productivity. If productivity improvement programs are explained to employees and they are allowed, and even encouraged, to participate in their formulation they stand a better chance of succeeding. Participation can serve as a safeguard against boredom and apathy on the part of police officers. The handling of mostly quality cases can serve as a motivator for the officer as he grows both in skill and in his feeling of accomplishment; increasing productivity, but how much? The goals of the police department, its

managers, and the men on the street can be integrated in productivity programs.
59:21

Productivity programs offer the Police Manager the opportunity to assure maximum utilization of tax dollars, and the key element in such programs is the development of measurements. The concept of productivity and its reliance on measurement might well be the stimulus that causes police officials to act like real managers and become more concerned with an analytical approach to their job. The management analyst provides management with the necessary information it needs to make decisions concerning resource utilization. The analyst performs the necessary conceptual groundwork to establish what it is that should be measured and reported, and develops the systems to gather the required data and to monitor and report system performance. With such information, police managers will be better able to make decisions concerning the use of police manpower and equipment.
59:22-26

In order to improve productivity, the criminal justice manager must first be able to measure it. Productivity is concerned with the generation of valuable outputs via the application of manpower and equipment. In diagnosing a productivity problem, other measures will be required in addition to productivity measures such as: efficiency, effectiveness, and quality. Efficiency measures demonstrate how completely resources dedicated for a specific task are being used in the performance of that task, whereas effectiveness is concerned with the degree of fulfillment of an objective. Effectiveness measures demonstrate how good an activity is at generating useful, productive outputs. Quality measures are often

needed to help guard against reductions in the quality of police service
59:26-28
that might result from efforts to improve productivity.

The mix of services actually delivered by a criminal justice agency is
a function of local objectives and priorities, and usually includes
crime-directed and noncrime services. There are three basic objectives
59:28
of the police force:

1. Deterrence of Crime.
2. Apprehension of Criminal Offenders.
3. Satisfactory Provision of Noncrime Services.

Poor productivity is caused by the inefficient performance of an activity
or the performance of an activity which is intrinsically ineffective.

Efficiency and effectiveness measures are particularly useful in facili-
tating diagnosis of productivity problems. Geometrical probability
techniques are important in planning situations in which an administrator
examines how alternative numbers and positionings of units in the field
affect the performance of the system. A queuing situation evolves when
a population places excessive demands on a limited-capacity service system,
and in recent years, interest has been focused on network problems and
59:31-47
algorithms based on mathematical programming techniques.

When complex combinations of policy alternatives are being contemplated
in an actual urban environment, analytical models are used first to
achieve certain insights and to indicate important unresolved problems;
then simulation models are used to examine the policy alternatives in
detail. Simulation of urban public safety systems presents many new

problems not ordinarily faced in more usual situations. To be effective,
such a simulation must be structured to reflect fully the spatial rela-
tionships inherent in the operations, as well as the sequential time nature
of events common to many systems. The simulation model is constructed
to allow users to replicate to a very great extent the actual operations
of most urban police departments, providing a tool to assist in answering
a wide range of allocation questions. Police administrators should find
59:47-48
simulation models valuable for the following purposes:

1. They facilitate detailed investigations of operations.
2. They provide a consistent framework for estimating the
value of new technologies.
3. They serve as training tools to increase awareness of the
system interactions and consequences resulting from every-
day policy decisions.
4. They suggest new criteria for monitoring and evaluating
actual operating systems.

The simulation program can tabulate statistics on any algebraically
defined variable. Some variables that have been recorded most often
59:51-52
in the author's studies are:

1. Total time required to service an incident, that is travel time
plus time at the scene.
2. Workload of each patrol unit, measured in total job assignments
and in time spent on jobs.
3. Fraction of services preempted.
4. Amount of preventive patrol.
5. Travel time of a unit to reach the scene of the incident.

6. Dispatcher queue length.
7. Dispatcher queue wait.
8. The number of intersector dispatches.
9. The fraction of dispatcher and/or reassignment decisions for which the car position was estimated, rather than known exactly.
10. The fraction of dispatch decisions which were nonoptional in the sense that there was at least one available unit closer to the scene of the incident.
11. The extra distance traveled as the result of a nonoptional dispatch assignment.

An example of quantitatively based objectives providing the goals of a reallocation plan includes:

1. Provide immediate response.
2. Approximately equalize workload per car.
3. Provide about 50 percent of street time for patrol.

Constructing a model is one way to simplify complex operations and to show the relationships between activities that are combined to perform an operation. One step in model development is to identify and describe each activity that is performed to attain the goal. Activities are performed which are intended to lead to results, and the performance measure for each activity can be defined as information. The apprehension system can be conceptualized and modeled as an information processing system that performs activities which acquire, store, process, and transmit information that is needed in order to identify and arrest suspects.

The apprehension system, in terms of information processing might

be stated as follows: In investigating a crime, the investigator engages in acquisition of information. The goal of the activity is to obtain information that describes the perpetrator. If this goal is reached, the next goal is to locate the specific individual who matches the information. This is done by locating a source of information that connects information about identity with known individuals. Analysis of the number of times each activity was performed or the hours spent on each activity could lead to "efficiency" or "effectiveness" measures. The principle of relative effectiveness can also be applied to items of information. It is not sufficient to enumerate outputs; their quality must also be considered, for example, it is not sufficient to consider the number of arrests without also considering the quality of the arrests.

It seems appropriate to consider multiple measures of the output of police crime control, including quality of outputs produced.

59:63-69

The effect of noncrime services performed by criminal enforcement agencies has not been fully investigated, for example, the "number of miles patrolled," or "doors checked," or "lectures given" can be easily counted, but it is not clear as to how much they contribute to the desired prevention and apprehension objectives. Efforts should be made to ensure that output measures represent progress toward end objectives of the services or useful outcomes.

59:70

The Symposium report summarized, police productivity should be measured by a set of output, effectiveness, and quality indicators to be viewed relative to inputs in terms of both dollars and manhours. Productivity

measures may be compared over time, and may be used to evaluate specific programs as well as for forecasting police resource requirements. 59:70-71

Effectiveness measures reflect progress toward meeting objectives of the service, along with quality of service measures. These reflect the overall outputs of the service. Effectiveness measures should be distinguished from workload measures. Workload measures can be divided into two types:

1. Workload accomplished.
2. Workload presented.

Often there is a hidden workload that police departments face in the various noncrime service functions. Some are:

1. Prevent crimes.
2. Minimize casualties and property loss from crime.
3. Speedy apprehension of offenders.
4. Promote feeling of security from crime.
5. Protect civil liberties.
6. Provide service in a prompt, fair, honest, and courteous manner. 59:74-75
7. Provide service to the satisfaction of the public.

59:76

The Symposium selected these effectiveness and productivity measures:

1. Reported crime rates.
2. Victimization rates.
3. Net property loss per 1000 population.
4. Physical casualties per 1000 population.
5. Clearance rates.
6. Percent of arrests surviving preliminary court hearings.

7. Percent of arrests leading to conviction.
8. Response time to calls, by type.
9. Citizen feeling of security.
10. Citizen satisfaction with service.

These measures could be viewed relative to police service input expenditures and manpower. They also cited additional examples of productivity measures: 59:77

1. Population served per police employee.
2. Population served per \$1000 expenditures.
3. "Quality" arrests per police employee and per \$1000 expenditures.
4. Arrests per police employee and per \$1000 expenditures.

These measures could be viewed relative to the appropriate quality of service measures.

The quality of arrest might be indicated by the percent that survive preliminary court hearing, and the percent resulting in conviction. To measure productivity, effectiveness and efficiency measurements could be analyzed together. The Symposium concluded there are three main points to consider when establishing a productivity measurement system: 59:78-82

1. Multiple measures are needed.
2. Quality and not just quantity of outputs should be considered.
3. It is crucial to provide for the analysis of the data from the beginning of the project.

A 1975 Police Foundation publication edited by Wolfe and Heaphy addressed productivity in law enforcement saying, effectiveness generally refers

to achieving certain defined results or outcomes without regard to the cost of achieving them; that is, if you are in fact ever confronted with an objective to be achieved at all costs, then the standard by which your performance is being measured is an effectiveness standard. Efficiency, on the other hand refers to achieving any given result with the minimum expenditure of effort required to achieve that result. Productivity is a combination of the effectiveness and efficiency concepts. Productivity asks both whether a desired result was achieved (the effectiveness question) and what resources were consumed to achieve it (the efficiency question). Two common ways of dealing with management problems, reducing service or increasing expenditures, are often ways of decreasing, not increasing, productivity. Reducing service is reducing effectiveness. Increasing expenditures can signal decreasing efficiency. The manager who says that he needs more resources to be more productive may not really understand what productivity is. Productivity improvement usually entails different ways of doing things with concomitant cost of change.

61:3-4

Achieving productivity growth is a long-term, tedious, and unglamorous task, and productivity growth varies from sector to sector over both the short and the long term. There are five major sources of productivity growth:

61:7

1. Application of knowledge.
2. More capital per worker.
3. Higher quality of labor.
4. Improved allocation of labor.
5. Economies of scale.

Within an organization, the common element of all these factors is that they are within the purview of management responsibility.

Productivity in public services follows the same general pattern as productivity anywhere else in the economy. The concepts, aimed at improving efficiency and productivity, have the effect of determining accountability with great precision. Finding ways to measure the whole gamut of police activities will be much more difficult than determining say, the number of television sets manufactured by a company. Human interactions, which is what police work is all about, present new problems when it comes to definition and measurement.

61:7-43

What also should be developed are new measurements of overall departmental effectiveness. What we measure, and how we measure it, is only half of the productivity question. The other half is how we improve with the resources we have. Productivity improvement is a continuing process. It is a way of managing an organization by relating measurements of efficiency and effectiveness to the achievement of clearly stated organizational objectives. Every manager in the department must be thinking in terms of the productivity of his or her own unit. The productivity improvement process begins when police managers begin asking better questions than they have in the past.

61:43-44

There is a pressing need for a more rational establishment of enforcement policies and close coordination with other criminal justice agencies. Accountability is a political concept. Under our system of government, the public, acting through its elected representatives, has delegated

powers and responsibilities to various public agencies, and the public has the right to know what its agencies are doing and to hold them accountable. As the police develop the kinds of specific information that will permit the public to assess what their departments are doing for them, as they work in partnership with other agencies, and as they maximize the use of their resources, they will be taking important steps in the direction of increased accountability, productivity, and professionalism. 61:45-46

Two approaches seem appropriate to prevent intentional or unintentional error, to reinforce the credibility of the data they are providing, and to avoid the scorn of the inevitable skeptics of the world.

1. Wherever possible, provide for collection of data by an independent, disinterested source.
2. Provide external auditing of data.

Such data as obtained through a survey of citizens would have the advantage of being objectively gathered. 61:122

Measuring police crime control productivity is now a very unsatisfying activity, but substantial improvements over current general practices seem possible. Indicators of the quality of arrests seem highly desirable, and man-crimes should be accounted for. Consideration should be given to external and noncontrollable circumstances, such as varying demographic and economic characteristics, and measurements should be subject to periodic auditing, both to ensure the quality and to reinforce public credibility. 61:122-124

Before productivity can be improved, data must be collected to show current baseline productivity levels. The determination of current productivity levels will also show where improvements need to be made. Productivity measurements should not be used to compare one police department with another, but should be considered only as in-house management tools because political, demographic, and geographic differences will affect even jurisdictions of similar size in different ways. Traditional manpower utilization studies are concerned primarily with the output, as compared to the resources, but functional analysis in police work must go beyond such typical measurements and look at relationships between parts of a system. Monitoring each project in a productivity program is necessary to insure that fluctuating resources are maximized to meet changing needs, and the purpose of each project is not just to generate statistics but to get the job done better. 61:140

Belkin, Blumstein, Cassidy and Cohen, discussing the measurement of service effectiveness in the CJS called for a concept introduced to measure the contribution of law enforcement in the service areas such as:

1. Calls answered.
2. Documents/information furnished.
3. Traces completed.
4. Weapons or explosives destroyed or confiscated.

Since the primary goal of the CJS is crime reduction, a possible goal could be: minimize the sum of the social cost of crime, plus the social cost of crime control, less the service benefits, equals a total measure

of the effectiveness of the CJS. But social costs, largely implicit, include the public's fear of crime which can well lead to behavior that is itself crime - generating, such as frightened citizens being more likely to purchase weapons, legally or illegally, and these weapons could be misused either in an accident or in some criminal act. In developing effectiveness measurements, an important national need is the provision of evaluative information from jurisdictions in which evaluation is conducted to other jurisdictions, or a national data or method bank. 7:iv-40

The 1973 report of the Advisory Group on Productivity in Law Enforcement said increasing productivity in police services might be considered in four ways: 54:2-3

1. Increasing police productivity means improving current police practices to the best level known, to get better performance without proportionate increase in cost.
2. Increasing police productivity means allocating resources to activities which give the highest return for each additional dollar spent, and requires asking not just whether the force is doing things right, but also whether it is doing the right things.
3. Given the uncertainties of police work, increasing productivity means increasing the probability that a given objective will be met.
4. Increasing productivity in police work means making the most of the talents of police personnel, since economy requires they be used more effectively.

According to the Advisory Group, for any police activity, productivity must be considered in relation to effectiveness. The two concepts are

closely related and at times may be difficult to differentiate.

Effectiveness is a measure of the extent to which a goal is achieved, whereas productivity includes not just what was accomplished but what resources were required to accomplish it, and does not necessarily indicate the extent to which the result actually accomplished a given goal. The Advisory Group has used the term productivity in a way that implies a greater concern for effectiveness and quality or value of service than is usually associated with the term "efficiency," while others concerned with this concept may define productivity as effectiveness over input, or a cost-effectiveness ratio. One thing that is always common to both productivity and effectiveness is "output," or results, and better productivity assessment is an important step in the process of productivity improvement. 54:3-4

Getting a greater return for the dollar spent is not a "one shot" activity, but is an ongoing, long-term process that should be an integral part of police management. The Advisory Group has identified a five-stage process as one approach to productivity improvement: 54:4-5

1. Establishment of Objectives. Ideally each police department establishes its goals, then proceeds to identify intermediate objectives, the achievement of which will contribute to the attainment of the broader goals.
2. Systematic Assessment of Progress. Most managers have some judgment on how their police force is doing, but often these are "gut reactions" based on little more than intuition and informal evaluation. Effective assessment requires more precise measurement, for without more precise

measures it is difficult to determine how much better or worse a particular unit, strategy, or piece of equipment works.

3. Search for Improved Operating Methods. Many improved operating methods, types of equipment, and ideas being used in certain police agencies could and should be made known to and be applied in other jurisdictions. Police managers ought to play an active role in searching for new and better methods.
4. Experimentation. "Innovation" is a luxury many police departments feel they cannot afford. However, neither can they afford to hold to the status quo while conditions around them change. Clearly a prescription for a balanced approach to risk-taking is needed.
5. Implementation. The sense of caution and resistance to change that an innovation might have met from department leadership extend throughout the department, the government, and the citizenry as well. Overcoming this resistance requires involvement of those people at the experimentation stage, as well as through preparation, patience, cooperation, close monitoring of the innovation, and clear accountability.

A first step to improving measurement is to understand how the various functions of police work relate to the broader mission of the department and the goals of government. Although overall police performance may be judged by the general public on the basis of crime prevention or some perceived level of public security, the police are also responsible for non-crime-related and nonemergency services, such as:

1. Maintenance of public order.
2. Emergency response.

3. Community relations.
4. Nonemergency general services.

Measures of police performance must also take into account the other system components that affect the outputs of police work, but measuring police activity need not await, nor depend upon, a final resolution of the "proper" police role and responsibility. Certain goals can be agreed upon, and certain activities clearly are important enough that measurement of them can proceed. It may be that careful measurement and analysis of specific police activities will gradually produce a clearer understanding of their relationship to broader police and community goals. 54:10-11

The principal purpose of measurement is to provide sufficiently precise information to enable police managers to:

1. Evaluate their department's performance.
2. Identify and diagnose problem areas.
3. Design solutions.

Measures frequently stimulate constructive thinking, increasing the understanding of police activity, and also may provide a means for linking one activity to another, or one part of the management process to another. Measurement is not a substitute for sound professional judgment; it is meant to assist the manager, not dictate actions, and care must be taken to guard against measures that provoke negative activity. Some measures may require data gathering that is more expensive than their value, and consequently should be avoided. 54:11

The identification of intermediate activities and objectives requires great care and constant evaluation to assure that they do, in fact, contribute to higher departmental goals. There are two fundamental types of measures: (1) measures of results, and (2) measures of resources used. Police departments, as is true of most public services, traditionally have been more concerned with measures of resources than with measures of results, since results are generally more difficult to define and measure. A comparison between results achieved and results intended could give a simple measure of effectiveness, and result and resource measures could be compared to indicate productivity.

54:11-12

Police departments can be more productive in meeting noncrime service objectives if they carefully analyze what is required to provide these services. Measures are useful only when they lead to analysis and improvement of police operations. For a department concerned about crime prevention, the requisite activities need visibility, emphasis, competent direction, and commitment. The effectiveness of such activities suffer when they are performed as an adjunct to other activities, and while the problems of measuring crime prevention are formidable, the opportunities for developing new programs geared specifically toward the prevention of crime, nevertheless appear to be great. There is great potential for diverting departmental resources from marginally productive activities to higher leverage programs of active and anticipatory crime prevention.

54:28-45

The Advisory Group made these suggestions for overcoming impediments to innovation:

54:69-70

1. Greater emphasis on incremental changes.
2. An independent evaluation agency.
3. A regular survey of police innovations.
4. Joint funding of specialist positions.
5. More extensive use of interdepartment personnel rotation programs for periods of up to 2 years.
6. Those departments that tend to lead in the adoption of new ideas should be identified.
7. Training programs on the evaluation of new ideas, the management of change, and the concept of planned change.
8. Grants encouraging and supporting increased police productivity.

There are two distinct governmental roles in fostering police innovation:

1. To identify and disseminate promising new innovations.
2. To recognize and revise the organizational, personnel, and financial policies and procedures that choke attempts to change.

The Advisory Group made these suggestions intended to increase the likelihood that new programs will succeed:

54:70

1. Select programs that at least seem to be more productive in achieving a department's objectives than the normal police activities.
2. Have a realistic idea of the time delays before new programs can become productive.
3. Public must be educated about innovative programs.
4. Personnel must also be educated about the goals of innovative programs.

5. New programs should be carefully designed and evaluated so that, whether they are permanently adopted or dropped, the department will still have gained some knowledge from the experience.

Lind noted in the case of the CJS and its components, one can, through 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. He also felt it is critical to keep in mind the connection between the operational, process-oriented measures of performance and the higher level objective. ^{45:8-9}

One approach to performance measurement and evaluation of the criminal justice system and of its component institutions is to ask what we are trying to achieve and to push the question "Why?" until the answer is either that "It is something we value in itself" or "Whatever my ultimate goal, I am sure that this is something I want to increase or decrease." If we can develop a procedure for assigning numbers to situations so that higher numbers correspond to higher level of achievement, then we have developed an ordinal scale on which to evaluate performance with regard to that objective. ^{45:9-10}

Many measures for evaluative purposes do exist and are being used in some form and 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. Another approach to performance measurement and evaluation is to go directly to people in the community and 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, the marketing or survey approach. The ability of citizens to make informed decisions with regard to expenditures on criminal justice will probably not be great, but there are two groups of citizens who do have firsthand knowledge of the system: (1) victims of crime, and (2) those who have been arrested. These groups are potential sources of valuable information on how the system is performing and their opinions should be surveyed. ^{45:10-11}

Lind suggests 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. Survey techniques can be used to determine the attributes that people value from the service they receive from criminal justice institutions, and can help us identify the people's objectives for the system. They 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. Survey data combined with scaling techniques may help us around the problem of defining an objective function in multi-attribute decision problems. Such techniques can be used to weight crimes by their seriousness, for example the Sellin and Wolfgang crime seriousness index, or show how various police activities are weighted in value by

citizens in a community. Not only can such techniques be used to identify objectives, but to weight them as well, and it provides another method of getting a crude measure of how the CJS and its major components are performing. It provides an alternative barometer to the standard barometer, the crime rate. The public evaluation of the system is important just as is the level of crime and must be considered.

45:12-13

Two of the primary reasons for measuring performance and evaluating programs on the basis of performance are:

1. To enable us to better allocate our resources.
2. To better manage the CJS.

The decisions with regard to allocations and management determine the requirements for measurement and evaluation.

45:13

The question of resource allocation within the justice producing system can be analyzed in terms of three separate but mutually interdependent decisions:

1. Decisions about how much of the society's resources are to be devoted to the justice system.
2. Decisions about what proportion of the total expenditure will go to each of the production units.
3. Decisions within production units about the allocation of funds among specific production tasks.

The optimal expenditure for justice will be reached when the total spent has been increased to the point where the incremental increase in justice is valued equally to the incremental costs of obtaining it. There will

be very different effects in the production of justice depending on where funds are allocated within the system, and on how these funds are used within the agencies to which they are allocated.

45:13-14

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: (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? The manager thus needs measures of performance that are task-oriented and measures of effectiveness related to basic objectives. To evaluate programs he needs a measure to tell him if the task was performed well, one to tell him if it made a difference, and a task measure as a means of monitoring and controlling the operation of the organization.

45:15-16

According to Lind, for the purpose of allocating resources among different branches of the CJS, it is vitally important that we develop measures of workloads and of performance for each of the component parts. It is also necessary 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 order to effect a better balance within the CJS. This is demonstrated in the system modeling done by Blumstein, Larson, and others. We must also analyze allocation decisions to keep the higher level objectives of the system in mind. 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. Crime rates and the level of crime by type are and will remain an important element in criminal justice evaluation, but 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 CJS and appropriate for research on the factors that determine the level of crime in a given community. ^{45:17-19}

Lind proposed these program elements:

1. Research is needed to promote the more effective use of evaluation of performance to provide guidance as to how to make criminal justice evaluation more effective and to implement better decision making based on performance measurement within various parts of the CJS.
 - a. Studies of our major criminal justice institutions such as the police, would be undertaken to analyze what decisions they make, how they make those decisions, and whether they use performance measures to assess current practice and whether it could be improved by improved measurement or by the use of better evaluation techniques. This would include a review of the use of traditional measures in evaluation and assessment of 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. Envisioned 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. ^{45:19-20}

- b. A better balance of expenditure between the various branches of the CJS is needed. There is an apparent significant misallocation of resources among different criminal justice institutions. In some instances the system appears to have broken down, in that one part cannot handle the work generated by 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. Based on such an appraisal, ways of reallocating resources, or at least providing supplemental resources, could be found, to help those parts of the system that constitute a bottleneck to 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. We could also 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 ^{45:20} might be effected.
2. Studies are needed of crime for evaluative purposes to aid the public and our criminal justice agencies 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.

- a. A study should be undertaken to design a research methodology and the supporting data base that would be required to assess the degree to which various socioeconomic and CJS factors influence the crime rate. 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. What is needed is for someone with strong methodological skills to develop a simultaneous equation model of crime including both socioeconomic variables and criminal justice variables, and assess the data requirements of using such a model and develop a strategy for collecting the data. Such a study would be to design the idea analysis of the factors that influence the rate of crime, which would be useful for the purposes of criminal justice evaluation in trying to separate 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.^{45:21-22}
- b. It is proposed that studies be undertaken to determine whether it would be possible to develop survey instruments that could 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, and for the total CJS. A 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. One 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, so it is recommended that the Office of Evaluation (LEAA) 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.^{45:23}

- c. 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 crime, organized crime, and transnational crime. 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 we are not allocating the resources to enforcement in these areas that is justified.

3. Exploration is needed of the potential of survey and attitudinal scaling techniques to measure Criminal Justice outputs. There may be significant potential not only from using survey techniques, but also in the use of modern techniques of attitudinal scaling to develop measures of

performance based on citizen perception and citizen values. These techniques have promise for developing overall measures of performance for 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.

a. It is proposed that the Office of Evaluation study the potential of these techniques for criminal justice measurement and evaluation. 45:24-25

4. Methodological developments for criminal justice research and evaluation to address the problem will be relatively inexpensive and will have wide application, in most cases.

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 uses. Many statistics on crime and the performance of criminal justice institutions are misleading now 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. 45:25

b. Wherever it is possible to measure the benefits of a program in dollar terms, benefit-cost analysis can then be applied. Utility ranking might also be possible. 45:25

5. Continued basic research is needed on the operation of the CJS to enhance our ability to analyze the impact of programs on measures of performance which is perhaps the greatest obstacle to better evaluation.

a. The Office of Evaluation should support systems studies of the CJS that show the interrelation between the operation of various parts of the system and their effects on other parts of the system and how they operate. The work is basic in making predictions about what the effects of programs and various parts of the system will be on other parts of the system.

b. Many possible influences or alternatives to criminal justice action lie in the private sector, which we should consider.

c. Basic work on such things as recordkeeping, basic indicators of performance with respect to tasks within the system or basic conditions within the system should all continue to be supported. 45:26

6. Evaluation study of LEAA's own program, both retrospectively and prospectively, is needed because LEAA provides funding and has leverage to get the agencies which it funds to do evaluations.

a. It is proposed that a project be funded to design an evaluation strategy for LEAA to 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 projects evaluation and to evaluate projects. 45:26-27

Lind concludes that while the subject matter of any research program is important, good research is critically dependent on having first rate researchers and that only with development of a strong research community and the development of a receptive clientele of practitioners can the state of criminal justice evaluation be advanced. 45:27

Deutsch defines the goal of an organization as the formal reason for the existence of the organization; the overall mission of an organization. It is the 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. 20:14

A measurement strategy, according to Deutsch, 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. 20:23

A measure of 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. The basis for any measure depends on the basic process:

CONTINUED

3 OF 4

Stimulus → Response

This process describes the observable behavior of an organization.
20:23
What is not observed is not measured.

Measurement strategies are the means by which the quantitative values of performance measures are determined, and measurement strategy in the 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 environment. Measurement strategies are methodologies for applying performance measures, which in turn quantify some aspect of behavior or the impacts of behavior.
20:23

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.
20:23-24

All measures can be classified as either being absolute or relative. 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 police 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.
20:24

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 are 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 environment's response to that performance that can be shown to be correlated to the aspect of performance to be evaluated.
20:24

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. These factors must be recognized in the determination of an organization's responses to stimuli, including:
20:25

1. Goals, objectives, and activities of the organization.
2. Relationship to other CJS agencies.
3. Technology and environment of the organization.
4. Internal structure.
5. Internal workflow.
6. Internal flow of resources.

If the process Deutsch noted (by which the organization produces a response to a stimulus) is recognized, performance measures fall into three categories:

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 measures of the environment's response to the activities of the organization.

Another 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. Measures of performance can be divided in two other categories:

1. Measures which are primarily concerned with the expenditure of resources.
2. Measures concerned primarily with the achievement of objectives.

The first is resource-oriented and the second is objective-oriented. When used together, each measure tends to enhance the meaning and significance of the other. ^{20:26-27}

Deutsch summarizes, saying performance measures can be typified in at least six ways:

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, and it is often useful to apply two or more measures together in order to increase the significance of the results obtained. ^{20:27}

Deutsch cites Ostrom who wrote that the difficulty in evaluating police performance is that police provide a multiplicity of services which require the use of multiple measures, addressing both efficiency and effectiveness. ^{20:28}

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 reasonable to expect Government agencies to attempt to provide maximum services at minimum cost, but very little consideration has been given to development of measurement strategies that do not require all possible data be gathered. The theory of "sampling" is well developed and is currently applied in many areas by private industry, and could be applied in the CJS. Typically performance measures have turned to quotas and evaluative programs have been so poorly designed as to be seen as obstacles by many involved in the CJS. 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. The process of data collection is really sampling, and the measurement obtained by applying a measure of performance with a measurement strategy are really estimates obtained by sampling, thus advantage can be taken of statistical theory when developing a measurement strategy. 20:29-33

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, which consists of: 20:33-34

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.
6. Improvement of the measure and/or measurement strategy.

According to Deutsch, what should be done by those seeking to evaluate social services such as law enforcement, is to develop 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: 20:38

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.
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.

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.
3. The data gathering capability of the evaluators.

The activity may be economic in nature, which would call for a resource-oriented measure, but if there are certain noneconomic features, an objective-oriented measure may be appropriate. If the evaluators cannot afford direct measurement, surrogate measures may be chosen. 20:38-39

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.
3. The costs for data gathering.

Each type of behavior suggests different measurement strategies. 20:39

The overall measurement process is designed in accordance with:

1. The chosen measure.
2. The chosen measurement strategy.
3. The purposes of the measurement.
4. The identity of the evaluators.

The process must complement the measure and strategy selected, and needs to be designed with its purpose in mind. There are a number of factors at play in determining a best measure, strategy, and process to best accommodate the purposes of the measurement within the restrictions imposed by these factors.

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, but despite the difficulty inherent in this determination, it should be attempted. Deutsch goes on to say "there is no excuse for not developing this decision methodology to CJS applications."

Effectiveness can be defined as validity or the degree to which objectives are met, and 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. The performance of an organization can be compared to that of another organization giving relative effectiveness. A cost effectiveness index 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.

Efficiency, according to Deutsch, is the output of resources from the organization divided by the input of resources to the organization.

Efficiency is also a resource-oriented measure of effectiveness.

Efficiency would be, for example, 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.

In order to fully express each of the important aspects of an organization's behavior, it is beneficial to develop a multidimensional measure of performance. 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 and can be included in a multidimensional measure of overall performance in order to provide a more complete picture of organizational behavior than could be had through use of a simple cost effectiveness index or other single performance measure.

There is no overriding need to attempt to compress all aspects of organizational performance into a single performance index. Any multidimensional measure developed can be conveniently represented with simple vector notation.

The overall effectiveness of an organization depends on how well the activities serve the objectives and in turn, the relationship between the objectives.

The overall effectiveness of an organization is also 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:

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

can be constructed. The variable e_i represents the degree to which the i th 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 \vec{e} into a scalar value E . With respect to overall effectiveness of the CJS, 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. Assume that the courts, however, do not sentence offenders to incarceration of any appreciable duration. Under these circumstances, Avi-Itzhak and Shinnar 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 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. Quantitative modeling offers the specific advantage of being able to pinpoint weak links in the CJS, and it is much easier and less costly to manipulate a model than to experiment in the field. 20:45-55

Deutsch outlines an approach for determining a general quantitative relationship, either empirically or functionally based, between Agency effectiveness and CJS effectiveness, suggesting that it would be best to:

1. Make a quantitative evaluation of the effectiveness of each component agency.
2. Study the relationships between the component organizations to determine the influence of each agency upon the behavior of the CJS.
3. Use the knowledge gained above to derive a general relationship between the effectiveness of each component organization and overall CJS effectiveness.

Some mathematical expression might be developed from the data to yield the overall effectiveness figure. Several authors such as Hatry have come to recognize the importance of interaction between component organizations. 20:56

Deutsch goes on to recommend 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. To provide the specificity necessary for identifying difficulties within the system, this will not be subject to unproven assumptions about the cause-effect nature of crime such as is the practice of selecting some single indicator of overall performance, such as UCR trends. When the general nature of the organization, its behavior, and its environment is known, evaluators can then proceed to select a measurement approach. 20:56-62

The overall development of the measurement approach calls for the determination of the effectiveness measure, then the measurement strategy, and finally the measurement process. 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:

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 classified as to their suitable applications. Finally, the measurement approach is selected and applied. Some means for validation of both are:

1. The applicability of the measurement approach.
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
20:62
part of the feedback loop.

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,
20:63
these must be identified so that the evaluators know what is being evaluated.

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 first examines the CJS from a systems perspective, and then proceeds to consider the member agencies and their interrelationships in detail, as well as

the environment of each organization. During these analyses, the criteria for selection of a measurement approach are implicitly stated.

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. Classification is performed after the measures, strategies, and processes have been analyzed relative to their applicability to different types of activities. The measure, strategy, or process is considered in light of its characteristics and an assessment is made as to its
20:63-64
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. The grouping is based on the objectives toward which each activity is directed and the types of tasks involved. 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
20:64
only three ways:

1. Results can be analyzed to see if they are intuitively reasonable.
2. Results can be compared to results obtained by different approaches applied to the same activities.
3. By different approaches applied to similar activities in another organization.

A determination of the goals, objectives, and activities 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. The objectives must also be checked for consistency. The activities engaged in by the individual law enforcement organizations in question need to be thoroughly identified and in some way classified. Deutsch says only when these steps have been completed, will sufficient information be available to begin to identify potential measures of effectiveness. 20:65

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, individual activities throughout the system can be classified by the nature of the activity and the objective toward which the activity is directed. 20:65-66

The Criminal Justice System is a multicomponent conceptual entity engaged in the production of law enforcement services, described by its component organizations and the relationships between them. The CJS is divided into three main levels, Federal, State and local, and five main technologies: 20:66

1. Police and Investigative.
2. Prosecutorial.
3. Judicial.

4. Detention.
5. Planning and Intelligence.

The relationships between the component organizations 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 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. In the context of a systems behavior, both the composite GOA's of agencies in 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. 20:66-69

A determination of the goals of the CJS or a component organization within the CJS must be made prior to choosing an effectiveness measure. 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. The goals must be examined for consistency, rationality, desirability, and equity, to insure that any effectiveness measurements made are meaningful. 20:69-70

The objectives of the CJS and its component organizations also need to be identified and analyzed before choosing a measure of effectiveness. The objectives must be checked for consistency, rationality, desirability,

and equity just as the goals were checked, and for the same reason. One additional check must be made; it must be shown that the objectives of the organization are consistent with the goals of the organization. 20:71

The activities of the CJS are those activities engaged in by its component organizations. Generally, it is these activities for which a performance evaluation is desired. 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. 20:72

The need for identification and classification of law enforcement activities as a prerequisite to finding or applying a measure of effectiveness has been established. 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. Goals, objectives and activities.
2. Relationship to other CJS agencies.
3. Technology and environment of the organization.
4. Internal structure.
5. Internal workflow.
6. Internal flow of resources.

Knowledge of the GOA alone is insufficient for selecting a measurement approach, because performance depends on both the organization and its environment. 20:72-73

Deutsch recognized that there are several approaches to quantifying the relationships between different law enforcement organizations, but noted it is doubtful that the relationship between any two law enforcement agencies can ever be entirely known. However, 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. Certain applications of modern multivariate analysis in determining organizational interrelationships prove useful; as well as constructing functional models of structure of the organizations and the flow of resources, information, or command between them, which generally relies upon using flow elements that can be readily measured, such as crimes, criminals, or monetary resources. 20:73-75

The availability of digital computers and matrix-handling algorithms enable evaluators to make precise definition of the characteristics of an activity because a larger number of characteristics could be utilized without sacrifice of ease in computation. When no "best" measurement approach is available for a particular set of characteristics describing an activity, the available measurement approaches could be evaluated, using matrix techniques, and the approach with the best "fit" selected, perhaps by first inserting some sort of weighting factors. The weighting factor for each characteristic would represent the degree to which that characteristic described the activity. 20:119

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. Mantel, et al., developed the concept of a service "package," with the agency rated on how well it delivered the package, with the results weighted for package importance. Deutsch said 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 a 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. There must be concern for maintaining a uniformity and completeness of thought. There must also 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. A great deal of performance measurement hinges on organizational and CJS objectives. There is also a requirement for specific studies in this area.

20:126-135

Deutsch stated that it is time to execute a truly comprehensive study of performance including:

1. The theoretical basis for measurement.
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. 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 would be developed. A general framework could be structured from the behavioral characteristics of the variables involved in the measurement process. The framework would encompass the entire concept of measurement: the measure, the strategy, and the process. Needed principles will be developed, such as the concept of measurability. There will be a definite effort to determine what can be measured, and to what degree of accuracy. 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. A uniform set of principles could be the coherent basis for measurement, expressed in terms that lend themselves to application within the CJS.

20:141-142

Deutsch noted there are a number of possible measures that may be applicable to a given type of data. The first step in designing the measurement approach should be the selection of a measure, but 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 and should be able to consistently select least cost performance measures that are strategy and process compatible.
20:142-143

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 reason is that currently available measures of performance are being applied in situations and to behavior they were not designed to measure. 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, including new developments in the police 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. According to Deutsch, no single performance measure completely describes an activity, but a pair or group of measures that complement each other can 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. Given an adequate effort including behavioral modeling, studies into organizational dynamics, field work, and other areas, a multiplicity of 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. Complementary performance measures can give extended significance to results obtained from performance evaluations of a law enforcement agency. Within this context, Deutsch notes 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. There is also a significant synergistic effect involved when 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 towards discovering the behavioral principles involved in determining whether or not a particular application will yield complementary response-impact measures.
20:143-145

There are likely to be many evaluative applications in the CJS for the induced objective concept. The definition of the induced objective relates it to behavior that can be observed, and therefore measured. With regard to effectiveness measurement, identifying induced objectives 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:

1. The mere identification of induced objectives provides evidence about the agency's performance and its environment. Evaluators can compare

induced objectives to formal objectives to point out any variance between desired results and actual input.

2. Evaluators may be able to base effectiveness measurement on induced objectives where there is reason to believe the stated objective 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.

Many aspects of an organization's performance are quite difficult or costly to measure directly, and in these cases, surrogate or "proxy" measures may yield significant and representative results at reasonable cost and effort. However, 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. From correlative studies and other analysis of investigations involving field observations a general method for selecting surrogate measures could be developed, and a multiplicity of such measures useful in CJS applications could be developed, applied, and evaluated to ensure that the general method for selecting surrogate measures is valid.

There are many situations where performance is not adequately represented by a single measure, and it is highly questionable to compress different measures into a single common denominator such as dollar value. It is often quite useful to employ a multidimensional measure of effectiveness to properly represent the observed behavior, but most existing research and most applications seek only to measure output, efficiency, and some

measure of quality, which is directly converted to dollars or a utility value. Needed is development of multidimensional measures of performance for CJS applications, which might incorporate several groups of complementary measures. The measures would no doubt be varied in orientation, but all chosen to describe the same behavior. Another purpose is to show how multidimensional measures can be applied in evaluating CJS performance.

While much effort has been expended debating which source of data best suits CJS measurement needs, little attention has been directed to the development of an overall measurement strategy. Deutsch notes the need for a research effort directed toward charting the relationships between behavior and the measurement strategies best suited for performance

measurement of that behavior, to build a decision process for selecting measurement strategies.

There is a limit to the value of data and the value of compiling additional data depends on the marginal advance toward achieving organizational objectives that would be made possible by the additional information. A study into determination of the value of crime-related data should encompass both determination of how an organization in the CJS uses information and what that information costs, allowing some quantification of the marginal advantage gained by collection of additional crime-related data by a CJS organization or group of organizations to be developed.

The art of sampling is well developed and there is a statistical foundation for selecting sample sizes that minimize cost and maintain or maximize

accuracy. Measurement can be equated to sampling, so that the determination of measurement effort involves setting sample sizes. Sampling in the CJS is an example of using a new technology. 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. It is time to pull together all that is known and produce a coherent set of principles to:

20:148-149

1. Govern the current use of computers in the CJS.
2. Direct new attempted applications.

An enormous number of promising new measurement strategies are available in theory, but they have not been applied to performance measurement within the CJS. Deutsch proposed that a search for new measurement strategies be initiated, for purposes of:

20:149-150

1. Identifying a general classification scheme for all measurement strategies.
2. Surveying measurement strategies employed in non-CJS applications.
3. Applying the principles of measurement to design a multiplicity of new and useful measurement strategies.

Many of the principles of sampling correlate with or should be incorporated with the principles of measurement. Deutsch seeks to determine the involvement of sampling principles in measurement, and to develop a number of sampling-oriented measurement strategies designed such that the basic concept of sampling, selecting a representative subset of all available data, directs their execution. He also notes that time series methods stand to make a significant contribution to the performance measurement of CJS organizations, and points out the need to develop new time series measurement strategies.

This would provide a new approach to performance measurement, particularly in conjunction with the complementary response-impact measures to be developed.

20:150

Evaluators, in their haste to find performance measures 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 too little effort made to apply modern systems thinking to the performance measurement process in the CJS. According to Deutsch, we need to explore the factors that determine what composes a best measurement process, to develop a decision process for CJS evaluators to apply in selecting such a measurement process, and to consider the different process requirements of the CJS activities where measurement is applied. The decision process developed would be structured to allow the flexibility necessary to accommodate the great variety of CJS applications.

20:151

There is significant difference between the requirements of measurements carried out for internal control and those performance measurements executed for the benefit of elements external to the organization. These need to be analyzed. For example, simple enumeration might be of significant value in adapting the general measurement process to a particular internal application, but of little value in external evaluation. There are new evaluative processes being developed, a number of which could be synthesized, particularly for application at the organization level.

20:151-152

One of the basic premises of performance measurement is that the behavior in question can be identified. We need to develop a number of methodologies

for the identification and classification of law enforcement activities to produce classifications that allow for and ensure the selection of the measurement approach most appropriate to the behavior being evaluated. The methodologies need be universal in scope so that all CJS activities can be classified by functional, rather than jurisdictional characteristics, because many measurement approaches are activity-specific. Different types of behavior call for different types of measurement approaches. 20:153

The confusion of infeasible and conflicting objectives pursued by CJS component agencies and the lack of a uniform and coherent set of systemwide objectives visible to the public lead to the notion of induced objectives, because of the necessity of identifying or selecting objectives prior to measurement. Deutsch advocates 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 systemwide objectives. He also proposes a process for selecting systemwide objectives for the CJS as a complement to the organizational objective/selection process. 20:154-155

The desirability of having some estimate of the value of a public service is obvious and 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 scalar region on which value is defined. Once a sufficient number of iterations have been made, or a sufficient number of bounds applied, the result is an interval estimate of the value

of the service in question. The consideration of the value of a public in the setting of objectives also calls for the recognition of time dynamicity. 20:156

The individual CJS agencies select objectives that best suit their own needs, regardless of system needs. Overall performance of an organization or system depends on how well it performs each of its activities. We need to develop an overall set of principles for relating performance levels for the individual activities of an organization to an overall measure of performance. Functional models relating flow of resources, information, and casework through a CJS agency and its environment to overall organizational performance should be developed, expressed in terms of flow so that their contribution to overall performance could be quantified. 20:158

Deutsch goes on to recommend flow graphs which have been useful in determining the relationships between variables in complex feedback models; economic interpretations of which can be made for the attainment of various organizational objectives, and incorporated into an overall model of effectiveness; and empirical models for relating the effectiveness of an agency in its individual activities to overall effectiveness. He would also seek to apply several of the currently available correlative methods to the building of an empirical model for estimating the effectiveness function in which 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. He also seeks to identify some of the observable behavior internal to the organization that can be correlated with its overall effectiveness, such as "resolution of conflict" and "distribution of authority" fitted into an overall effectiveness function.

The empirical modeling approach as a means for determining the effectiveness function requires several subjective evaluations of overall performance. 20:159-160

Measures of performance based on the effects on the environment of the organization's operation cause one to seek to identify external surrogate measures. 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 could be used for the instances where the impact of certain aspects of organizational behavior cannot be directly measured. There is currently great interest shown in citizen surveys as data collection means. This could enable evaluators to build a behavioral model to judge the accuracy of their results and the limitations on the meaning of data from such surveys. 20:160-161

With emphasis on the activity-objective interaction, a general method for detecting conflicts and some theoretical adaptations for inclusion in the models of overall performance are needed. Development efforts should be undertaken for purposes of describing the relationship between the achievement of CJS objectives and for purposes of allocating resources among the various CJS component organizations and for identifying conflicting sets of objectives. Information about the contribution that individual organizations make to overall performance will assist in evaluating proposed policy changes. 20:161

Determination of overall CJS performance includes developing the general principles underlying systemwide performance measurement, and analyzing 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 accurate results in its application. 20:161-162

Deutsch also says flow graph models could be devised to account for information, resource, and benefit flows both within component organizations and between them with the objective to develop an adequate definition of the effectiveness function of the entire CJS. It could also be that the entire CJS falls within the scope of an economic model, with the objective to develop an effectiveness function from the economic considerations involved in the operation of the CJS. Empirical models for the effectiveness function might also develop a relationship between performance at the component level and overall CJS performance. Correlative analysis, as a means of describing the relationship, could serve as an empirically-derived effectiveness function. Another method might be exploring the possible application of the Delphi Method and similar methods in making an empirical description of the effectiveness function for the CJS, for use in making subjective evaluations of overall performance. 20:162-163

There is a cost to the relative lack of coordination among CJS component organizations and the lack of a chain of command ensures that conflicting objectives are designated. Quantification of the costs of the conflicts would certainly prove useful in both the determination of overall CJS performance and in the selection of CJS objectives. The costs developed may be expressed in economic or noneconomic terms, so long as the overall magnitude of the problem is truly reflected, based upon the functional interrelationships of the component organizations. 20:164-165

Deutsch summarizes noting a profound need for the development of concepts of measurement strategies and measurement process for CJS application.

There must be some effort made to develop models relating effectiveness

achieved in relation to individual objectives to overall organizational effectiveness. 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. Research can lead to new and more useful measurement approaches that depart from the rather superficial treatment previously afforded CJS applications. 20:165

A logical approach to examining the performance of an organization which underlies the philosophy of measurement is through behavioral analysis. Its major advantage is that it deals only with observable effects and an organization is characterized by its behavior which consists of the observable activities engaged in by its components. 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 effectiveness contribution the organization makes toward the overall effectiveness of the CJS. The technology of an organization must also 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. 20:13-86

Internal work flow in a law enforcement organization is the flow of information and individuals to be processed through the organization and can be diagrammed using flow charts. The internal flow of resources is the movement and expenditure of capital, man hours, and equipment hours. To trace the flow the: (1) "bottom up approach" traces individual transactions through the organization and records estimated resources expended at each step and the summation provides an estimate of organization variable

cost, or (2) the "top down approach" uses the budget to allocate for the CJS as a whole and considers quantifying the internal resources flow by: 20:91-98

1. Identifying the resource flow to each component organization.
2. Determining the resource flow relationships between each of the component organizations.

Many studies in the CJS have been of groups, such as Rand's group of California recidivists. Stewart says that if a design is dependent on some sort of treatment and comparison groups, their existence should be confirmed prior to implementation of the design. He also points out that administrative record systems are not constructed and maintained with the evaluator in mind, and the weakness of many data systems is the reason for many interview or survey type studies. He warns with respect to CJS data:

1. Do not assume information exists.
2. Do not assume existing information is readily available.
3. Do not assume definitions are consistent.
4. Do not assume information systems are compatible.
5. Analyze the process of inclusion and exclusion for possible effects on your purposes.

He also warns that even if one has some data in hand and those data have been analyzed by someone competent, interpretation of the numbers is no simple procedure. "Obvious" interpretations may prove false. 71:11-16

Stewart defines a variable as something we observe and for which we can characterize differences or variations. The simplest sort of variable is a "dichotomous attribute" composed of only two categories, for example, a prison releasee either does or does not recidivate within 6 months of

release. Explanatory variables are those which are used as the basis for developing an explanation of the variability of other variables. An independent variable (which may also be termed a predictor variable) is the basic type of explanatory variable. Dependent variables represent the phenomena to be explained by the explanatory variables. A variable is constituted of some number of categories or values such that, for any given observation, one and only one of the categories is appropriate. Ideally, the categories are exhaustive and mutually exclusive over some class of observations.
71:32-35

A great deal of evaluation work involves relating two or more variables. In the strict, technical sense, the term "correlation" refers to a limited set of statistical measures or coefficients. More generally, two variables are correlated if they show some association. Interaction effects are frequently encountered in the conduct of evaluation research. They occur when the joint effect of two or more explanatory variables is other than the simple sum of their individual effects upon the dependent variable.
71:34-37

Stewart reminds the reader that "correlation is not causation." Most measures of association are symmetric in that, if A is related to B, then it is also the case that B is related to A. However, we tend to think of causation as nonsymmetric, but if there is a causal link, we expect a correlation as well. Errors of measurement are significant in interpreting correlations because if the errors of measurement in two variables are uncorrelated, the correlation of the two variables will be diminished. Spurious correlation is a term which reminds us again that

correlation is not causation. The standard notion of spurious correlation is that two variables (X and Y) are correlated because they are both the effects of the third, common variable (Z). More generally, the effect of the third variable is to modify the correlation which would "otherwise" occur between the two primary variables. Partial correlation has been employed. Ecological correlations in their most simple form are correlations based on "collective" or "areal" units. The real concern is with the ecological fallacy which involves attributing ecological level findings to individual units, for example, if neighborhoods can be said to vary in terms of their tolerance for various forms of deviance, then we would expect ecological correlations among the several types of deviance to be positive. Ecological correlation does not indicate that one form of deviance affects another form of deviance.

The Pearson product moment correlation coefficient is the measure typically assumed when the word "correlation" is used in its technical sense. The mathematical qualities of this coefficient are those of a simple model which assumes various things about the variables and their relationship:
71:38-39

1. The relationship is linear.
2. The measurement scales of the variables are "equal interval."
3. The errors of measurement in the two variables are uncorrelated.

This report now deviates from its survey form to make two recommendations based on research observations:

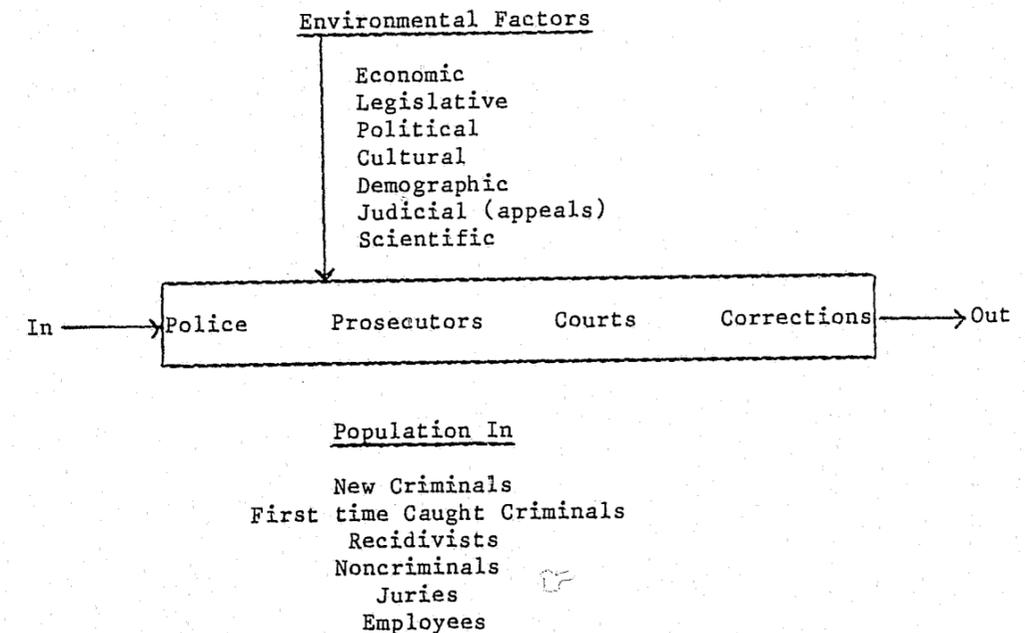
1. Criminal justice in the United States should be modeled in total or as a system. The model should include environmental factors as well as the interactions which occur between and among agencies.
2. Relative seriousness should be assigned to each crime type for which activities are performed by each criminal justice agency. Total agency or program utility could then be a function of time or effort applied to the various activities performed regarding the relatively weighted crimes. Noncrime activities could either be weighted separately by opinion survey, or allocated to the crime types in a manner similar to which overhead is allocated by accountants to production areas in an industrial application.

Modeling the United States CJS as a whole should include the Federal, State, county and municipal systems. These systems affect other agencies and the overall system, as well as affecting and being affected by environmental and economic factors. For example, consider the substitution and displacement effects on a jurisdiction when a neighboring jurisdiction implements a policy which criminals perceive as extremely hard on crime and criminal activities. A major benefit of total CJS modeling would be the understanding and quantification of these and other effects. Consider, too, the possibility of examination of each subsystem alone, after total system relationships are understood and quantified.

To model the CJS as a whole requires the inclusion of at least:

1. Law Enforcement.
2. Prosecution.
3. Courts.
4. Corrections.

These might be analogous to the internal parts of a machine operating in a factory, with the environmental and input variables applicable to the system. An oversimplified model might be:



One possible method of approaching the overall CJS model is utilization of the Delphi method of opinion convergence. Everyone agrees that there is an interjurisdictional or displacement effect. But how much? By surveying or empaneling experts and using the Delphi method, these effects could be modeled. Periodic validations and improvements to the model could refine the system over time.

Consider modeling the overall system as a map of the United States.

Ask the experts:

1. What effect does a 10 percent change in Virginia criminal justice effectiveness have on North Carolina? On California?
2. What effect does a 10 percent change in one jurisdiction's criminal justice effectiveness have on the State? On adjacent jurisdictions? On distant jurisdictions?
3. What effect does a 10 percent change in one component's (such as prosecution's) effectiveness have on the other CJS components within the same jurisdiction?

The resulting model for United States criminal justice will be large and complex, but it is needed as a starting point in evaluation. If the model was designed to be used in the interactive mode, its value as a training tool for CJS management would be enhanced. Its value in allocating scarce resources is obvious - invest where the resources will be most beneficial to society as a whole.

Modeling the CJS graphically and mathematically will result in a conceptualization and a quantification of how the component agencies interact with each other, the system inputs, and the environmental variables. A result would be the ability to determine what functional utility an agency adds to overall system productivity, given certain conditions or information. Service overlaps would also be clearly shown. Sensitivity analysis would also be feasible to determine what effect on agency contribution changes in the other variables will make. For example, one could investigate the effect on other agencies of changes in the output or capacity of a given agency.

A CJS agency should not be evaluated in isolation or in a vacuum until the agency's interrelationships and environmental effects are modeled, understood, and quantified to the point that agency contributions can be realistically examined with respect to what it should be under given conditions. For example, what change in the efficiency of ATF (enforcing firearms laws) could be expected if a law was enacted prohibiting the possession of any firearm outside one's own dwelling? What changes in effectiveness? Productivity? How would such a law affect other agencies?

To emphasize the need for total system evaluation, an analogy to single agency evaluation might be evaluating a freight car in the middle of a long train without considering any other cars, the engine, the engineer, a snowstorm, the uphill climb, and the broken tracks.

Another analogy might be a criminal conspiracy case wherein all the facts and circumstances together may prove something in total that the parts could not, taken separately.

A law enforcement agency would contribute little to society if there were no prosecution, no courts, no correction agencies; but on the other hand, what contribution could a jail facility make without a law enforcement agency to make arrests or courts to sentence offenders? The interagency dependency and effects can also be observed when an understaffed prosecutor declines to prosecute cases recommended by a law enforcement agency and the fact exists that a crime has been committed and there is evidence or probable cause to believe the suspect committed it. How much relative harm is done

by failing to prosecute a criminal? Is this type of inaction encouraging crime?

To evaluate the CJS, agency, or program implies that their products have value, utility, usefulness, or worth. Although agreement can be easily reached that a given law enforcement program has value, the exact value is not easily derived. For example, of what value is removing one criminal tool such as a firearm from criminal hands? What do we make this value in terms of? Dollars? Utility? Persons likely to be shot? Robbed?

40:
According to Jung, the flow, and velocity of flow of firearms may be analogous to the flow and velocity of flow of money, well understood by economists. Both have initial supply and both have turnovers. To evaluate firearms flow in this manner, a sample would be required to analyze historically. Then additional analysis might take the form of sensitivity analysis to determine what effect each of the variables have on each other in such a relationship. Additional analysis might also take the form of the more traditional turnover in the accounting sense, as usually applied in inventory analysis. This analogy points out that it may be possible to determine the average street life of a firearm (as well as velocity and turnover estimates). One must bear in mind that these numbers likely will show great dispersion around any mean and are likely to be vastly different for different type firearms such as handguns as opposed to long guns. Such estimates may vary area to area and over time. If these numbers are found to change over time, then it might also be feasible to calculate the rate of change equation. Such investigations may produce a value in

terms of utility or if the total number of crimes and total cost of crime could be determined, in terms of dollars, but do not appear feasible at this time with available data.

Considering the constraints imposed by current data limitations, a seriousness or utility index for evaluating all law enforcement and criminal justice services seems more feasible. Such an index could, by surveying an appropriate population, establish a relative value for each crime and noncrime service provided. According to Wolfgang, such a relative value index is feasible for even the large number of Federal statutes. His recent work, "National Survey of Crime Severity," reports such an index for 204 crime items as determined by survey incorporated in the LEAA and Bureau of Census, "Victimization Surveys." If the cited survey is limited to 60,000 households in 13 U.S. cities, it may not be appropriate for a seriousness or utility survey of Federal crimes. For example, the firearms laws, forever controversial, may have far different value to rural Americans as opposed to city dwellers. A proper survey would give every citizen an equal chance to be a respondent. Each criminal justice agency has priorities, expressed or implied, which suggest that certain of their services are deemed to have greater value than others. Since there is no structure or chain of command for the CJS as a whole, the priorities or values are left up to each agency. These agencies naturally establish self-serving priorities largely ignoring the needs of the overall system. Given this propensity for the Criminal Justice Agency to adopt priorities and therefore goals, objectives, and activities which tend to maximize some internally generated criteria, it is easily

seen why the overall CJS productivity is not even considered, much less measured.

The proposed index of utility, value, or seriousness would rank crime types. For example, in the Federal sector, all Federal crimes would be ranked in a numerical weight order similar to the manner in which Wolfgang ranked 204 crimes using the "Victimization Surveys" to elicit public opinion. It would then be possible for each criminal justice agency to simply apply its time or effort expended to the index weight for the subject crime their activity applies to, relating a comparative total value function.

For example, let the crime, "bombing resulting in death," take on the arbitrary relative value 98.0 on a 100 scale. The law enforcement agency investigating the bombing applies its time, say 100 man-days as:

$$100 \text{ man-days} \times \frac{98.0 \text{ units of value}}{\text{man-day}} = 9800 \text{ units of value}$$

The prosecutor trying the suspect applies his time as say, 10 man-days, as: $10 \text{ man-days} \times \frac{98.0 \text{ units of value}}{\text{man-day}} = 980 \text{ units of value}$

Similar calculations could be made in court and corrections applications. As an illustration, take a simple prison with two prisoners, one of whom is our bombing suspect, from above, having been convicted and incarcerated.

Let the other be a forger whose crime was valued at
20.0 units of value/man-day

For the simple prison illustration the sum of the utility for the day would be

Prisoner #1	98.0
Prisoner #2	$\frac{20.0}{118.0}$

If the prison staff required to house these two inmates is three men, then the utility function would be

$$3 \text{ man-days} \times \frac{118.0 \text{ units of value}}{\text{man-day}} = 354 \text{ units of value}$$

The astute reader at this point has noted that the foregoing may address effectiveness but certainly not efficiency. One method to complete the productivity measure which must include efficiency in addition to effectiveness would be to develop a relationship of "units of value" to some input variable, such as "dollars" or "number of employees," with dollars of course being the more standardizing and useful comparison device.

To further illustrate the use of a relative value index in CJS evaluation, consider the following examples using 1977 total employment and expenditure data for the system components:

1. Law Enforcement.
2. Judicial.
3. Corrections.
4. Prosecution.

1. Law Enforcement

Full-time equivalent employees	645,015
If an employee has 250 workdays/year, employee days available	161,250,000

If the employees applied their time to tasks regarding crimes having a mean relative value of 20.0 units/day,

Then units of utility/year 3,225,000,000

Enforcement cost/year \$11,864,870,000

It then follows

$$\frac{3,225,000,000 \text{ units}}{\$11,864,875,000} = .2714 \text{ units}/\$$$

2. Judicial

Full-time equivalent employees	150,546
If an employee works 250 days/year, employee days available	<u>37,636,500</u>

If these employees applied their time to tasks regarding crimes having a mean relative value of 20.0 units/day

Then units of utility/year 752,730,000

Judicial cost/year \$2,638,251,000

It then follows

$$\frac{752,730,000 \text{ units}}{\$2,638,251,000} = .2843 \text{ units}/\$$$

3. Corrections

Full-time equivalent employees	255,008
If an employee works 250 days/year employee days available	63,752,000
(to exclude parole functions multiply by a factor of .86)	54,826,720

If these employees applied their time to tasks regarding inmates whose convictions were for crimes having a mean relative value of 20.0 units/day

Then units of utility/year 1,096,534,400

Corrections cost/year (excluding parole at above ratio) \$4,243,297,600

It then follows

$$\frac{1,096,534,400 \text{ units}}{\$4,243,297,600} = .2569 \text{ units}/\$$$

4. Prosecution

Full-time equivalent employees	63,902
If an employee works 250 days/year employee days available	15,975,500

If these employees applied their time to tasks regarding crimes having a mean relative value of 20.0 units of utility/day

Then units of utility/year 319,510,000

Prosecution cost/year \$1,225,344,000

It then follows

$$\frac{319,510,000 \text{ units}}{\$1,225,344,000} = .2612 \text{ units}/\$$$

In looking over the foregoing examples, the reader likely was surprised that the comparative utilities were so close to the same value. Recall, however, that these figures are totals for the respective system components and that the system works currently to a degree. Also recall that the 20.0 utility units/day was arbitrarily, but uniformly assigned and that within each CJS component and each agency, this figure is likely to show great variation or dispersion about any component or system mean.

The necessity may arise to tackle the problem of allocating indirect labor and overhead costs to the appropriate production function. This is not an insurmountable difficulty. Industrial entities utilize well-established accounting techniques and similar procedures could be developed for CJS applications. For example, police dispatcher efforts could be applied at the mean crime relative value on a daily, weekly, monthly or yearly basis. Real or implicit rentals could be applied on the basis of square feet of buildings used by these respective functions. For interagency comparability, uniformity in definition, application of measures, and reporting is necessary.

The system of weighting activities by relative value of the crimes they pertain to is not meant to imply that no time or effort should be expended on low relative value activities. It is meant to show that emphasis could be placed on activities associated with crimes of higher relative value. For example, if traffic enforcement in a city is deemed to require a minimum of 200 man-days, then evaluation can be made with that minimum application accounted for.

Another criticism which might arise after viewing these recommendations is, "this addresses the crime, but not the criminal," or "what about the career criminal?" Traditionally, in the American judicial system, a suspect may only be convicted of crimes after guilt is proven beyond a reasonable doubt. Of course, there is a precedent for considering past record in sentencing an offender, but that is not predictive sentencing, which would be meting out penalty based on what crimes the suspect is likely to commit later. The constitutional implications of such actions are obviously prohibitive. Predicting the type and frequency of crimes to be committed by a particular suspect is not feasible at this writing. Attempts to incarcerate those criminals deemed prone to violent crime have typically failed due to the high numbers of false-positives, that is, the high numbers predicted to commit additional violent crimes, who do not. In fact, the way to minimize forecasting error may be to turn all the inmates loose.

If an individual is suspected of having committed a crime, then certainly the investigation warranted should be applied at the seriousness index rate for that crime, but if an individual with a severe criminal record is incidentally caught for some minor offense, it is the minor offense rate at which the time should be applied.

There may be some merit in developing a criminal history factor for the index utilizing the same survey as is used to derive the index. For example, respondent opinion could be solicited to determine what factor to multiply current crime index rate by, for persons having one

prior conviction, two, three, and then perhaps, four or more. More specificity might be desired, utilizing categories or types of prior convictions, such as felony versus nonfelony or violent versus non-violent. This position might be supported by reasoning that persons with prior offenses may require disproportionate CJS effort, especially to apprehend and incarcerate.

Multiple violation crimes must also be addressed to implement the recommendations. The envisioned relative value index could be cumulative with respect to multiple offense crimes by simply summing the relative value index rate for each crime type. For example, in an armed robbery which also resulted in a murder, the respective relative value index rates could be added together and effort applied as a function of the summed rate.

It should be pointed out that the tendency toward reporting what is important to management is recognized as a potential pitfall in any system of measurement. Independent audit, that is, examination of reported results by impartial persons from outside the agency, could be used to verify stated results, much in the same manner that a certified public accountant's opinion of a corporation's stated operating results and financial position indicates those statements are made according to generally accepted accounting principles applied on a basis consistent with past reports. For agencies in the Federal Sector, this audit could be done in the normal course of business by GAO or OMB, who likely would be thrilled to have some objective basis on which to audit social service organizations.

It should be pointed out here that the purpose of these recommendations is not necessarily to find where to cut expenditures, but to find which parts of the system need additional input or resources to become more system objective oriented; to find which areas need to be strengthened in order to reach an optimum or a balanced condition for maximum system productivity. These recommendations are meant solely as a management tool. Neither this nor any quantitative method should be applied without regard to experienced managerial judgment.

The advantages of having a uniform productivity measurement technique resulting from uniform definitions and data, relating program or agency value in terms of relative utility, are numerous both with respect to the management functions of planning, allocating, evaluating, and controlling within the agencies; and for the external evaluation and allocation decisions with respect to resources to be divided among the agencies.

No single agency should be expected to accomplish such broad improvements in criminal justice evaluation, but each agency should be expected to contribute to the project. An agency such as LEAA, whose name implies assistance to law enforcement, should be expected to support and even lead in such an effort.

For the Federal executive who feels that productivity measurement is not important, the following excerpts from a statement made to the House Budget Committee by Comptroller General Elmer Staats, in 1978, point out that agency budgets may soon be tied to productivity.

Productivity data is broadly defined to include all measures of efficiency to determine how well an organization is using available resources to produce the required goods and services for a constant level of quality. But efficiency measures do not comment on the appropriateness of an activity itself. Both efficiency and effectiveness measures are necessary to get a true picture of an organization's performance.

GAO, in collaboration with the Civil Service Commission and the Office of Management and Budget, in 1973, concluded that the most important use of productivity measures is in analyzing the causes of change to ascertain what action management can take to influence future trends.

Five basic benefits accrue when productivity data is used in the budgeting and, in a larger sense, the management process:

1. Agency managers will place greater emphasis on improving productivity if they believe productivity data will be used for formulating, reviewing, and executing budgets.
2. Using productivity data in conjunction with specific program objectives contributes to better agency projection of resource needs. The capability of OMB and the Congress to review those needs is also improved.
3. Budget estimates are more credible when they are supported with productivity measures because reliance on judgmental estimates is minimized.
4. Clearly presented and meaningful productivity data will make agencies more accountable to OMB, the President, and the Congress.

5. Because the need for reallocating resources occurs at various times during budget review and execution, the availability of productivity data enhances managers' ability to react accurately and expediently.

The Federal budget continues to increase dramatically causing the taxpayers increasing concern over how their tax dollars are being spent. The Congress, by emphasizing its interest and concern for productivity through appropriation hearings, can have a significant impact on Federal productivity improvement.

Budget requests should be based on reliable estimates of personnel requirements. Productivity data is necessary for management to accurately determine the personnel and funds required to meet an organization's goals. Emphasizing productivity during budget preparation and approval can help ensure more efficient use of the Federal budget dollar. The use of productivity data in the budget process is now limited. Larger potential exists for increasing the use of the data in formulating and executing budgets, especially in labor-intensive agencies. A top-down, across-the-board emphasis for the Congress, OMB and agency top management is the best way to increase Federal managers' use of productivity data, thereby increasing productivity itself.

When reviewing agency budgets, oversight committees and appropriations subcommittees can further encourage use of productivity data by:

1. Requesting productivity data to support agency requests for staffing increases.
2. Requesting concise statements on the status of agency or department productivity improvement programs, work measurement systems, and the extent to which budgets are based on productivity data.
3. Creating an atmosphere of positive reinforcement for using productivity data through the use of budgetary and organizational incentives.
4. Encouraging agencies to identify major productivity improvements possible through investment in capital equipment.

The March 1977 issue of the LEAA Newsletter advised that perhaps the greatest need in the police area is to assist administrators to develop the effective management systems required to facilitate program evaluation and program direction. The lack of an effective program performance evaluation system was reported by the 1973 National Advisory Commission on Criminal Justice Standards and Goals. It recommended that "a national study be undertaken to determine methods to evaluate and measure the effectiveness of individual police agencies in performing their crime control functions." Developing a system to enable police administrators and others to better evaluate the effectiveness of police operations is a primary police concern.
57:1-9

According to Stewart, the mandate, if not demand, for quantitative program evaluations is widespread within the criminal justice area.
71:1

One of the major deficiencies in police administration is the lack of adequate performance measures. The need to develop sound productivity measures is becoming crucial with today's budget constraints.
63:3

According to the American Justice Institute, the consequences of the inability of police to measure effectiveness and productivity well are damaging; the inability to measure how well objectives are achieved impairs evaluation of performance, effectiveness, and productivity; the ability of police to remain accountable is impaired; the ability of the police to perform an extensive number of critical management functions is impaired; and the willingness of police to submit to rational evaluation

is questioned, thereby inducing or reinforcing credibility problems. This nation's political system demands that police be accountable for the degree to which objectives are achieved. Because police do not measure effectiveness and productivity adequately, they cannot satisfy their legal and moral obligation to remain accountable.

56:3

Blumstein and Larson noted the GJS "has remained remarkably unchanged throughout the significant social, technological, and managerial changes of recent decades... partly from the insularity of these institutions and their relative freedom from external examination and influence" ... and also partly due to the "independence of the individual components of the system, each of which operates within prescribed rules to attain its own suboptimized objective."

10:1

The human behavioral aspect of evaluation is often overlooked in the quest of a method or data, but Blum addressed evaluation from the social/behavioral sciences point of view. Evaluation is a necessary activity to accomplish rational public administration and public policy formulation. One basic aspect of evaluation planning is the examination of the likely sources of resistance and another is provision for facilitating innovations once recommended.

45:App. A, 1

Evaluation implies that justifications for the existence of criminal justice institutions, procedures, objectives, and endorsements are being questioned. The conduct of an evaluation requires that it be

guided by a strategy. A strategy means that alternatives will be considered, costs and benefits weighed, value positions considered, and resources allocated. The strategy should acknowledge the interests of those authorizing the evaluation.

45:App. A, 1-4

Blum said the term "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 constitute flows of and blocks to the processing of the daily business; people, papers, and objects. Of interest also are the flow and counterflow of emotions, ideas, and power. One encounters the problem of under-inclusiveness here. The "system" is ordinarily conceived according to its formal aspects, but no systems investigator is so naive as to presume that these are the only elements. These are but aspects or emanations of daily work. 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. One could select 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 maintaining peace, preventing crime, general deterrence, reinforcing the individual conscience through public displays of punishment, satisfying vengeance, substituting controlled community retaliation for the vagaries of individual or family retaliation, reassuring citizens by the display of lawful authority, or even that lofty correctional goal of penance, reform, and rehabilitation. But it is

quite likely, that most "crime prevention" occurs when families or well-adjusted parents rear law-abiding children who grow up in wholesome peer groups. It is also likely that general deterrence operates as part of the general fabric of morality, reinforced by individual conscience and social norms. ^{45:App. A, 5-9}

An overall strategy for evaluating the existing formal justice system, the goal of identifying objectives, and finding measures which help us learn to what extent these are being met requires that we learn what people expect from the CJS, 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. We need also learn which groups are not "served" by the system. ^{45:App A, 12}

One approach is to interview samples of people participating in the system. Another method is to observe those participating to see how their careers are affected by their exposure to the system. ^{45:App. A, 13}

Victims can be a very useful source of performance data about the system. Why not ask a random sample of victims one day after they have reported their victimization what it was they wanted from the police? Then ask these victims how they feel about the police and courts 2 weeks later, and then 6 months and a year later. A wide-based population survey and victim followup, and police response to victim data illustrate quite different approaches to evaluation. Victim data surveying is a feedback 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 crime victims are clients of the system whose evaluations are often overlooked. ^{45:App. A, 14-15}

Blum also noted that offenders are the system's products whose views are remarkably ignored. 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, such as:

1. Was the apprehension efficient?
2. Was there brutality?
3. How many apprehensions were missed when they could have occurred?
4. Were the police charitable, drunk, or simply too busy to make other missed arrests?

One of the functions of external evaluation is to determine if managers share values with the citizenry and, whether they do or not, to see if they can assess the state of their organization. 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. ^{45: App. A, 16-21}

The design of evaluation studies which embrace the awareness of dynamic effects depends on the discipline from which the evaluator comes and the kind of concerns which he, along with administrators and policymakers have about the operation of an agency or other component. 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 measures of shifts away from or towards additional goals, short and long-term, in the system component. The role of evaluation in the consistently difficult problem of priority setting for resource allocation in systems with limited capabilities and the description of capabilities within a given agency become important evaluation activities.
45: App. A, 22

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 is sufficiently important then one must expect to do research which isolates the variables which influence rates measuring attainment. Variables which are under administrative control and 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. We must also ask what impact change in the system from the outside may have on CJS objectives. For a given sector of the system, one does have measures which assist in defining objectives; describing those features within the system that do and do not bear on these, and identifying additional influences outside of the system which also bear on results. One can develop improved means for measuring objective attainment. The system, overall, must "work" because it exists. Existence and support prove

efficacy, even if it is only the efficacy of tradition and virtue.

A second conclusion is that the system obviously does not "work."

Surveys easily document dissatisfaction and the widespread conviction that the goals of crime prevention, detection, deterrence, corrections and justice are not being met. 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.
45:App. A, 24-28

Meanwhile, in the social sciences, preoccupation grows with the sophistication of measurement, and a number of reasonable tools now exist which can assist in the evaluation of the CJS. 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.
45:App. A, 29

Sellin and Wolfgang pointed out that "conventional criminal statistics are designed to show both the frequency and the degree of seriousness of violations of the criminal law." "The frequency is shown by counting the offenses committed; the seriousness, by grouping the offenses counted in categories according to legal definitions and arranging these categories in an order with the one containing the most serious offenses at the top and listing the rest in a decreasing order of seriousness."
67:292

The Office of Management and Budget, Evaluation and Program Implementation Division, surveyed agency evaluation practices. Noted in these survey results was a large number of agencies which have an office of planning and evaluation. Two examples were:
74:ii

1. The Department of Commerce, Office of Planning and Evaluation,
2. The Department of Justice has Offices of Planning and Evaluation in the FBI, DEA, and INS.

These phenomena imply recognition that the planning group should also evaluate.

Maltz said implementation of any new program is bound to surface unforeseen
46:18 problems. But Bottoms and Nilsson pointed out management sciences offer potentially valuable assistance to the police administrator at every level of command. The usefulness of the management sciences - operations research, systems analysis, and the applied computer sciences can attack planning problems, procedural problems, and problems of crime control.

Policy constraints - internal and external - also affect the operational feasibility of some of the action suggestions. Some benefits of using technology are:

1. Improved control over the allocation of financial, equipment, and personnel resources.
2. Identification and evaluation of alternatives in action, deployments, and in the purchase of new hardware.
3. Improved response to citizen calls for service.
4. Improved effectiveness in the preventive patrol function.
5. Enhanced police-community relationships arising from the apparent increased professionalism.

These benefits can lead to reduced burdens on the decisionmaker and
12:22 improved morale in the street.

No police agency in this country, according to Bottoms and Nilsson, can look forward to meeting increased demands with proportionate increases in manpower. The management sciences, particularly operations research, can show how to capture the manpower leverage offered by science and technology at acceptable cost. The understanding of police-community goals generated by a conceptual systems analysis allows the police administrator to recognize the negative benefits of courses of action which might recommend themselves because of low cost or convenience. Operations research and systems analysis assist in identifying the conceptual and quantitative relationships that exist in complex operations such as those of the police department. The computer sciences provide techniques that can sift and analyze a vast amount of information from a variety of sources to give the police administrators timely information on which they can base decisions. Judgment as to the relative weight that must be assigned to various kinds of information still resides with the experienced police officer. Operations research can offer valuable assistance to the police administrator, help provide information on the effectiveness of alternative ways of deploying resources, help evaluate proposed technology,
12:22-26 and provide new concepts to providing police services.

There are many challenges ahead in applying the techniques of operations research in society. Probably no area is under greater scrutiny than the field of criminal justice, and in no other field does the risk of failure or error carry so heavy a burden. Operations research can help to provide some of the answers to these problems.

George Hanson reported to the Symposium Focusing on Police Productivity that the majority of law enforcement agencies are faced with an expanding workload. Resources, however, are increasingly difficult to secure, requiring the best use of available resources through the application of advanced management methods. Use of our most valuable resource, people, (participation of all personnel) was encouraged by the use of the task force approach to the solution of problems and development of systems. If goal setting is to result in realistic expectation, information must be available in a format that permits a thorough analysis, current data must be placed in a historical context, and there must be a means to react to changes that affect performance.

59:61-62

During the Symposium, Philip S. Schaenman said that improving our ability to measure police productivity is an important part of the efforts needed to improve that productivity. These needs include productivity measures that can be used for routine, periodic feedback to police, management, and the citizenry. Productivity refers to the efficiency of using resources to produce desired products or services. Real output is often measured in dollars; however, productivity is usually stated as an index, e.g., change from some base year. For government services such as police crime control, not only is it difficult to put a price tag on the outputs of the service, but it is also difficult to define the outputs even without attempting to price them.

59:68

As noted in Readings on Productivity in Policing, edited by Wolfle and Heaphy, there are at least four management requirements for assuring that productivity improves on a regular basis:

61:8-38

1. Commitment.
2. Analytical capability - Good analysis will not be hit or miss. It will always entail the systematic application of common sense.
3. Know-how.
4. Development of new ideas - Any gain is worth having.

Top administrators should fix responsibility for productivity improvement. Instead of relying on larger budgets garnered by emotional pleas before elected councils, the police manager is going to have to become more reliant on his productivity improvement efforts to meet the service demands he faces. There are too many opportunities for the application of our scarce public resources for us not to start holding police managers accountable for efficient and effective use of those resources. It is clearly the duty of elected officials to make a serious effort to determine if the money spent on policing is being wisely and efficiently used.

61:8-38

Targets that are too easy to meet will be of little use; targets that are unrealistic are likely to be frustrating, with respect to agencies. Target setting on productivity for individual police employees is likely to be a very controversial issue, and should probably not be undertaken without intensive participation by the individual employees.

61:120-121

In Readings on Productivity in Policing, James P. Morgan wrote that although the terms productivity measurement and productivity improvement may seem new to police work, many police departments have been practicing them for years. Productivity improvement in police work has been limited to a small range of problems, but all departments will have to adopt

productivity improvement as an overall, continuous goal. Productivity programs offer the police manager the opportunity to assure maximum utilization of tax dollars using an organized analytical approach to productivity, with the commitment, as well as endorsement, of the chief. 61:129

Productivity improvement is only half of the productivity process. The other half, and what must come first, is productivity measurement. Most city officials have continued to judge the quality of police service provided their citizens by comparing their own city's UCR figures with those in other cities. The police have helped perpetuate the use of this unreliable evaluative tool by not developing more realistic measurements of police work. 61:129

The introduction of a productivity improvement program will be challenged by obstacles, as would any innovation in a police department. The maximum involvement of employees throughout the various phases of a productivity program will greatly improve the chances of overcoming these obstacles. Any program which promotes innovation at the expense of cohesion and employee moral will probably be counterproductive. Obstacles to a productivity program include:

1. The political reality that programs must appear successful soon after their adoption.
2. The absence of technical and analytical in-house employees.
3. The customary idea that police effectiveness is shown by crime rates, when there are more relevant ways to measure police effectiveness.

4. The tendency to support traditional rather than innovative programs.
5. The belief that productivity programs threaten job security.

Getting more work done, or getting it done better, with the same amount of resources are not the only objectives of productivity programs. An equally important objective is to improve the morale and professionalism of each officer and the department as a whole. The police manager must lead his department to the accomplishment of its goals. By assigning officers to quality cases (both law enforcement and service types) the majority of the time, departments will motivate their officers to improve their skills and increase the satisfaction felt in handling cases competently. The goals of a police department, then, can be achieved by productivity programs without sacrificing human dignity. 61:146-149

Barnett, Larson, and Odoni, for evaluations of experimental programs and/or day-to-day operations that require quantitative performance measures and models, made the following recommendations: 4:4-18--4-19

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.
2. LEAA should select an already completed CJS experiment and test the developed methodology in an "after-the-fact" manner, so the methodology may be modified or changed, if necessary.

3. To facilitate evaluations of day-to-day operations, LEAA should select a small number of common operational problem areas that are conducive to improvement through the use of quantitative performance measures, methods, and models. 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.

Barnett, et al., also said 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.

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, including computer-assisted dispatch systems, automatic vehicle monitoring systems, computerized criminal history files, in-the-field inquiry systems, and computer-assisted court scheduling systems. 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. Barnett, et al., suggest that LEAA'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 by-product of the technology.

The Advisory Group on Productivity in Law Enforcement reported in 1973 the need for raising productivity as a national concern. Productivity analysis is in part a response to diversity. Attempts to develop productivity measurement tools for the public sector in general, and for police services in particular, are in their infancy. The Advisory Group decided to focus effort on limited areas of police work. Three areas were selected because they were believed to be both of great importance to most police departments, and subject to significant productivity improvements through existing techniques or knowledge. The areas treated are:

1. Patrol: direct services
2. Crime Prevention: specific programs
3. Human Resources: management of people

The Advisory Group's suggestion was for national support to assist State and local police agencies in developing their own capabilities for improving productivity.

Human resources account for 80 - 90 percent of present-day police costs. Perhaps the most important productivity ingredient is the attitude or motivation of a department's personnel. Two types of factors affect the productivity of human resources:

1. Personal factors, such as the values placed on work, family, and leisure time.
2. Organizational factors, such as the recruiting, selection, assignment, training, and organizational development processes.

Although it is important to be sensitive and responsive to personal factors, most are largely beyond the control of law enforcement agencies.

The need is to design work to simultaneously increase job satisfaction
54:47
and improve performance.

There is an importance of individual departments taking action themselves, but there is a need for national impetus to encourage dissemination, experimentation, and continuing development of findings. Productivity improvement clearly is not a "one shot" affair. Increasing productivity in police services is an especially complex undertaking, and success ultimately will depend upon a continuing and lively debate. Responsible and capable agencies or organizations willing to accept the responsibility for playing an active role in encouraging implementation and development of the programs and ideas contained in the Advisory Group's report and in general
54:71
to promote police productivity throughout the country, must be found.

Re-emphasizing what Deutsch has said, "optimal allocation of tax dollars among competing programs or agencies requires a knowledge of their performance"... "Implicit is this optimal allocation is the determination
20:10
of an optimal level of effort (expenditure) for each program."

If the reader has endured this report in hope of finding a happy ending and a magic productivity measurement formula and wonders what became of that page, there is but disappointment. There is no such happy ending, no such page. And there won't be - until sufficient high level interest and support are brought to bear and overwhelm the agency resistance to change and self-interest problems.

It is believed that a feasible, valid evaluation method can be generated using the two concepts of a total CJS model and a seriousness or weighting

index to establish relative importance or priorities. Certainly, agencies exist within the CJS: law enforcement, prosecution, courts, and corrections, which have intense interest in performance measurement. These agencies would genuinely like to have objective, valid techniques, accepted by society and the system, to either support their current direction or to delineate a superior alternative. Certainly there are agencies both within the CJS and without such as LEAA, HEW's Mental Health Study Center, the National Science Foundation, and others who are able to fund a project aimed at valid, acceptable CJS (and component agency) productivity measurement and improvement. Then too, there are those agencies such as GAO and OMB who are charged with evaluating CJS agencies, who should be delighted with an objective method to evaluate agencies and programs.

What is needed, then, is high level interest, cooperation, and support to conceive, finance, implement, and evaluate productivity measurement for the various criminal justice agencies as well as for the system as a whole.

The alternative is continuing to use scattered, unreliable, and sometimes nonexistent data which suffers from definitional and other problems measuring, if anything, efficiency in nonuniform terms, using nonuniform data, data collection, and processing methods.

BIBLIOGRAPHY

1. ALT, Francis B. and DEUTSCH, Stuart Jay, "The Effect of Massachusetts' Gun Control Law on Gun-Related Crimes in the City of Boston," Evaluation Quarterly, Sage Publishing Co., Beverly Hills, Ca., 11-77.
2. ALTEMOSE, John R. and WALDRON, Ronald J., "Determining and Defending Personnel Needs in Criminal Justice Organizations," Public Administration Review, July/August 1979.
3. AVI-ITZHAK, Benjamin and SHINNAR, Reuel, "Quantitative Models in Crime Control," Journal of Criminal Justice, Vol. I, Pergamon Press, 1973.
4. BARNETT, A., LARSON, R., ODONI, A., "Performance Measures for Evaluation of LEAA and CJS Programs," Performance Measurement and the Criminal Justice System: Four Conceptual Approaches, October, 1976.
5. BECK, James L. and HOFFMAN, Peter B., "Time Served and Release Performance: A Research Note," Journal of Research in Crime and Delinquency, July 1976.
6. BECK, James L. and HOFFMAN Peter B., "Parole Decision-Making: A Salient Factor Score," Journal of Criminal Justice, Vol. 2, p. 195-206, 1974.
7. BELKIN, J., BLUMSTEIN, A., CASSIDY, G. COHEN, J., "Performance Measurement and the Criminal Justice System", Performance Measurement and the Criminal Justice System: Four Conceptual Approaches, LEAA, October 1976.
8. BIDERMAN, Albert D., "Surveys of Population Samples for Estimating Crime Incidence," The Annals of the American Academy of Political and Social Science, Vol. 374, November 1967.
9. BLUMSTEIN, Alfred and LARSON, Richard, "Models of a Total Criminal Justice System," Institute for Defense Analyses, Science and Technology Division, November 1970.
10. BLUMSTEIN, Alfred and LARSON, Richard, "Models of a Total Criminal Justice System," 17 Operations Research 199, 1969.
11. BOLAND, Barbara and WILSON, James Q., "The Effect of the Police on Crime," Law and Society Review, Vol. 12, No. 3 Spring 1978.
12. BOTTOMS, Albert H. and NILSSON, Ernest, "Operations Research," Police Chief Magazine, May 1970.

13. CARR, T. and WYCKOFF, L.B., "Beyond the Routine Uses of OBSCIS," Proceedings from the OBSCIS Seminar, 1978.
14. CHAPMAN, Jeffrey I., "An Economic Model of Crime and Police: Some Empirical Results," Journal of Research in Crime and Delinquency, January 1976.
15. CHAIKEN, J., CRABILL, T., HOLLIDAY, L., JACQUETTE, D., LAWLERS, M., QUADI, E., "Criminal Justice Models: An Overview", Santa Monica, Ca. Rand, 1975.
16. CLARK, Ramsey, Crime in America, New York: Simon and Schuster, 1970.
17. CLINARD, Marshall B. and QUINNEY, Richard, Criminal Behavior Systems - A Typology, New York: Holt Rinehart and Winston, Inc., 1973.
18. Criminal Careers of Habitual Felons, Rand Corp., July 1978.
19. "Curbing the Repeat Offender; A Strategy for Prosecution," Institute for Law and Social Research, Washington, D.C., Publ. No. 3
20. DEUTSCH, Stuart Jay, "A Conceptual Basis for Effectiveness Measurement of Law Enforcement Activities," Performance Evaluation and the Criminal Justice System: Four Conceptual Approaches, October 1976.
21. DEUTSCH, Stuart Jay, Deterrence Effectiveness Measurement, Atlanta, Georgia Institute of Technology, 1978.
22. DEUTSCH, Stuart Jay and RICHARDS, Jerry E. "An Evaluation of CJS Models: Simulation Approaches and Recommendations," Industrial and Systems Engineering Report, Series No. J-78-14, Atlanta, Georgia Institute of Technology, May 1978.
23. DEUTSCH, Stuart Jay and MALMBORG, Charles J., A Dynamic Model for Discrimination Between Incapacitation and Deterrence, Atlanta, Georgia Institute of Technology, 1978.
24. DEUTSCH, Stuart Jay and PFEIFER, Phillip E., "A Three-State Iterative Procedure for Space-Time Modeling," Atlanta, Georgia Institute of Technology.
25. DEUTSCH, Stuart Jay, "Stochastic Modeling and Analysis of Crime." National Institute of Law Enforcement and Criminal Justice Quarterly Report, 4 (April - June 1976).
26. "Evaluation in Criminal Justice Programs," LEAA, June 1973.
27. FIGLIO, R., SELLIN, T., and WOLFGANG, M., Delinquency in a Birth Cohort, University of Chicago Press, 1972.

28. FISHER, K.D., NAGY, T., SATHER, G.A., WHINERY, L.H., Predictive Sentencing, Lexington Books, 1976.
29. FONTAINE, E.M., "A General Theory on Crime, Recidivism and Sentencing," International Criminal Police Review, No. 318, May 1978.
30. FORST, Brian E., Participation in Illegitimate Activities: Further Empirical Findings, 1976.
31. GIBBONS, Don C. Changing the Lawbreaker, Englewood Cliffs, NJ, Prentice-Hall, Inc. 1965.
32. GLASER, Daniel, Handbook of Criminology, Chicago: Rand McNally College Publishing Co., 1974.
33. GOTTFREDSON, D.M. and BALLARD, K.B. Jr., "The Validity of Two Parole Prediction Scales," ISCD, Research Service Center, California, December 1965.
34. GREENBERG, David F., "The Incapacitative Effect of Imprisonment - Some Estimates," Law and Society Review, Vo.9, No.4, Summer 1975.
35. HARRIS, Carl M. and MOITRA, Soumyo D., "On The Transfer of Some OR/MS Technology to Criminal Justice," Interfaces, Vol. 9, No. 1, November 1978.
36. HARRIS, Carl M. and MOITRA, Soumyo D., "Improved Statistical Techniques for the Measurement of Recidivism," Journal of Research in Crime and Delinquency, July 1978.
37. HARRIS, Carl M. and STOLLMACK, Stephen, "Failure Rate Analysis Applied to Recidivism," 1973.
38. HOOD, Roger and SPARKS, Richard, "Key Issues in Criminology," McGraw-Hill Book Co., New York, 1970.
39. International Criminal Police Review, Interpol, No. 317, April 1978.
40. JUNG, Clarence, Phone Interview on June 19, 1979, University of Richmond - School of Business - Economics Department.
41. KASSENBAUM G., WARD, D. and WILNER, D., Prison Treatment and Parole Survival: An Empirical Assessment, New York, John Wiley & Sons, 1971.
42. KITCHENER, H., SCHMIDT, A., and GLASER, Daniel, "How Persistent is Post-Prison Success," U.S. Bureau of Prisons, Unpublished Article, December 1975.

43. KROHN, Marvin, WALDO, Gordon P. and CHIRICOS, Theodore G., "Self-Reported Delinquency: A Comparison of Structured Interviews and Self-Administered Checklist," The Journal of Criminal Law and Criminology, Vol.65, No.4, December 1974.
44. LEVY, Girard W. and MYERS, Louis B., "Description and Prediction of the Intractable Inmate," Journal of Research in Crime and Delinquency, July 1978.
45. LIND, Robert C., "A Program of Research on Performance Measurement and Evaluation for the Criminal Justice System," Performance Measurement and the Criminal Justice System: Four Conceptual Approaches, October 1976.
46. MALTZ, Michael D., Evaluation of Crime Control Programs, Washington, D.C., U.S. Government Printing Office, 1972.
47. MCCLINTOCK, "The Dark Figure," Institute of Criminology, Cambridge, United Kingdom. 1970.
48. MCGUIRE, William J., Benefit Flows and Costs Characterizing Federal Correctional Institutions, James Madison University, Harrisonburg, Virginia, 1978.
49. MOVAHEDI, Siamak and OGLES, Richard H., "Prediction and Inference in Criminology," Criminology, Vol.14, No.2, 1976.
50. NAGEL, Stuart S. with NEEF, Marion, Operations Research Methods: As Applied to Political Science and the Legal Process, Beverly Hills, University of Illinois, Sage Publications, 1976.
51. NIJMEGEN, Van de Vegt and Dekker, and ZWANEBURG, M.A., "Prediction in Criminology," 1977, Deller & Van De Vegt, Netherlands.
52. NISHIMURA, H., "Individual Prediction Tables for Further Delinquency Used in the Japanese Juvenile Police," International Criminal Police Review, Vol. 33, No. 317, April 1978.
53. Operations Research in Law Enforcement, Justice, and Societal Security, edited by, BROUNSTEIN, Sidney H. and KAMRASS, Murray, Lexington Books, 1976.
54. "Opportunities for Improving Productivity in Police Services", Report of the Advisory Group on Productivity in Law Enforcement, Washington, D.C., National Commission on Productivity, 1973.
55. "Predicting Subsequent Conviction for Individual Male Prison Inmates," Schumacher, M., Australian and New Zealand Journal of Criminology, Southdown Press, Australia 3/74.

56. "Police Effectiveness & Productivity Measurement: A Package of Concepts, Tools & Guidelines for Building & Using A Measurement System," Vol 1-4. American Justice Institute, 1978.
57. "Police Research: An Overview", Research Briefs, Law Enforcement Assistance Administration Newsletter, March 1977.
58. "Problems with the Use of Regression Analysis in Prediction Studies", Journal of Research in Crime and Delinquency, January, 1976.
59. "Proceedings of the Criminal Justice Symposium Focusing on Police Productivity", edited by RICCIO, Lucius J., Bethlehem, Lehigh University, 1974.
60. "Productivity - Improving Police Productivity - More for Your Law Enforcement Dollar - A Brief for Elected Officials," The National Commission on Productivity.
61. "Readings on Productivity in Policing," edited by WOLFLE, Joan L. and HEAPHY, John F., Police Foundation, 1975.
62. "Report to the Congress", Comptroller General of the United States, February, 1978.
63. "Research Highlights," Office of Research Programs, National Institute of Law Enforcement and Criminal Justice, LEAA, 1976.
64. RICCIO, Lucius J., A Model for the Study of Crime Control Administration, May 1971.
65. ROBISON, J.O., SMITH, G.W., WENK, E.A., "Can Violence Be Predicted?", Crime and Delinquency, Berkeley, University of California, 1972.
66. SCHUMACHER, Mary, "Predicting Subsequent Conviction for Individual Male Prison Inmates," Australian and New Zealand Journal of Criminology, Vol.7, No.1, March 1974.
67. SELLIN, Thorsten and WOLFGANG, Marvin E., "The Measurement of Delinquency," Patterson Smith, Montclair, NJ, 1978 (1st ed 1964).
68. SHINNAR, Reuel and Shlomo, "The Effects of the Criminal Justice System on The Control of Crime: A Quantitative Approach," New York, City College of New York, 1975.
69. SIMON, S. and COCKERHAM, W., "State's Prisons Fail to Deter or Help Most Criminals," The Hartford Courant, Vol.CXXXVII, No. 55, February, 1974.

70. STAATS, Elmer B., Comptroller General of the United States, Statment Made February 14, 1978.
71. STEWART, Douglas K., Evaluation for Criminal Justice Agencies: Problem-Oriented Discussion, Washington, D.C., U.S. Government Printing Office, 1978.
72. STOLLMACK, Stephen, "Comments on the Mathematics of Behavioral Change," Evaluation Quarterly, Sage Publications, Beverly Hills, Ca. 2/79.
73. STOLLMAC, Stephen, "Predicting Inmate Population from Arrest, Court Disposition, and Recidivism Rates."
74. "Survey of Federal Evaluation Activities", Office of Management and Budget, Evaluation and Program Implementation Division, September 1975.
75. "Uniform Crime Reports for the United States," Federal Bureau of Investigation, Washington, D.C., 1975.
76. U.S. Bureau of Prisons, 1970.
77. van ALSTYNE, David J. and GOTTFREDSON, Michael R., "A Mutidimensional Contingency Table Analysis of Parole Outcome," Journal of Research in Crime and Delinquency, July 1978.
78. von HIRSCH, Andrew, "Prediction of Criminal Conduct and Preventive Confinement of Convicted Persons," 21 Buffalo Law Review 717, Spring 1972.
79. Wickersham Commission Reports, No.3, "Report on Criminal Statistics," Washington, D.C., U.S. Government Printing Office 1931.
80. WILKINS, Leslie T., "Inefficient Statistics," Parole, New York, Federal Legal Publications Inc, 1975.
81. WILLIAMS, Kristen M., Predicting Recidivism With Promis Data, 1975 The Institute for Law and Social Research.
82. WOLFGANG, Dr. Marvin E., Phone Interview on July 13, 1979, University of Pennsylvania - Center for Studies in Criminology and Criminal Law.

END