

## MEASUREMENT PROBLEMS IN OFFICIAL AND SURVEY CRIME RATES<sup>1</sup>

WESLEY G. SKOGAN

Department of Political Science  
Northwestern University  
Evanston, Illinois 60201

### ABSTRACT

*This paper analyzes sources of error in the two major methods we use to measure crime in America—official police statistics and victimization surveys. The two produce quite different pictures of the volume and distribution of crime, but it is not clear that this is because victim-based statistics are “accurate.” Each measurement procedure has its characteristic errors, some of which it shares with the other. Comparisons of official and survey data on crime are helpful in revealing the dimensions of these error terms, and they point out the analyses which must be conducted if we are to specify their exact proportions.*

### MEASUREMENT PROBLEMS IN OFFICIAL AND SURVEY CRIME RATES

The development of sample surveys which measure the volume and distribution of crime in the United States will provide social scientists and public administrators with valuable new data. In particular, the National Crime Panel and Central City samples currently being monitored by the Bureau of the Census should produce a rich body of information on aspects of criminal and victim behavior which previously escaped systematic analysis. This data base may be used to confront a host of problems for which current statistics are unsuitable.

An immediate use of survey estimates of crime rates, however, has been to compare them to official statistics. Reports released by the Law Enforcement Assistance Administration have stirred public interest by their contrast with police figures on crime of the type summarized in the F.B.I.'s yearly *Uniform Crime Report*. Such comparisons inevitably reveal wide gaps between rates registered by the two sources. National or city-level survey

figures overshadow official police statistics by a substantial margin. This type of analysis has been encouraged by the government's decision to calculate U.C.R.-compatible figures from citizen surveys, although this is perhaps the least useful application of the data. The observation that there are varying discrepancies between official and survey crime estimates does not tell us where the error lies. Every statistic (and this includes survey as well as police figures) is shaped by the process which operationally defines it, the procedures which capture it, and the organization which processes and interprets it. Survey and police crime-measurement procedures produce different figures, but the reasons for this and their implications require analysis. A discussion of how survey and official crime statistics differ and why we obtain these discrepancies may clarify both their comparability and their interpretation, and may speak to their improvement in the future.

#### MEASUREMENT ERROR AND OFFICIAL CRIME STATISTICS

The presence of error of considerable magnitude is not unique to measures of crime, although a half-century of continuous criticism has focused more attention upon the errorful nature of crime measures than enjoyed by most social statistics. Measurement is the process of mapping an empirical system into a symbolic system. It involves the application of definitions to delineate aspects of the empirical system which are of interest, and a series of "If. . . Then. . ." rules matching selected attributes of those phenomena to symbols. The resulting symbols, usually numbers, always map the richness of the referent system simplistically and inexactly.

In measurement terms, all of these observed scores are composed of two elements: they are partially "true score" (reflecting what we wish to observe) and partially error. Even rapidly repeated, apparently identical measurements of the same phenomenon will produce different numerical readings. The degree to which they are similar—our ability to reproduce our findings—is the "reliability" of a measurement process. Reliability tests, for example, would gauge the extent to which various police patrol teams classify the same set of events in the same manner. While the ability to examine events twice and find the same thing is the *sine qua non* of good measurement, even reliable measures may not be useful. The procedures may not be measuring the actual object of interest, or the resulting figures may be artifacts of the measurement process. Police districts with ambitious commanders may consistently produce low crime totals. This is a validity problem. In order to obtain valid, non-artifactual measures we employ multiple and differing techniques, cross-checking our findings at every turn (Bohrnstedt, 1970).

Disciplines with well-developed measurement traditions have evolved routine procedures for coping with these problems. Economists have stressed reliability; they require measures which are stable and comparable across time (Morgenstern, 1963). Psychologists emphasize validity. The intangibility of the psychological domain heightens concern that its apparent orderliness may be an artifact of specific methods of investigation. Sophisticated psychological measurement combines the fruits of interviews, projective evaluations, and physical observations (Campbell and Fiske, 1959).

The measurement of crime is a substantive and methodological problem of interest to researchers in a variety of disciplines. Perhaps as a result, most of the effort expended upon measurement problems has been conducted outside of any coherent measurement model. Scattered validation studies of official statistics have been reported. Price (1966) compared state-level property-crime totals with insurance rates and uncovered only moderate cor-

relations. But such criterion validation requires a dependent measure which is relatively error-free, and in this case "crimes known to the police" are probably a better indicator of the underlying distribution of events than the independent validator. A better example of criterion validation is the California Criminal Statistics Bureau's comparison of police and American Bankers Association's figures on bank robbery. The latter appears to have been clearly defined and exhaustively enumerated, and it proved to be reflected quite accurately in official statistics (California Criminal Statistics Bureau, 1967).

Validity studies of official measures of more typical events, those which are less clear-cut and involve more discretion on the part of police officers and administrators, have been less hopeful. Comparisons between official records and self-reports of delinquency or informal police "contact" reports indicate that official figures greatly underestimate the volume of events which might be uncovered in other ways (Chambliss and Nagasawa, 1969; Quinney, 1970).

Our current system of gathering and publishing official statistics on crime was a response to such problems. The invalidity of local department's efforts at data collection and the limited reliability of the reported figures led to the development of the Uniform Crime Reporting system in the late 1920's. This system improved reliability, but sacrificed validity. Standardized definitions, data-collection forms, and data-gathering techniques produced city-level crime totals which are usually comparable from year to year, and inter-city comparisons undoubtedly are vastly improved by the U.C.R. system. But several important compromises were made in the formulation of this statistical system. The data are still gathered by local authorities, participation in the network is not mandatory, and the F.B.I.'s only option in the face of fraud is not to publish the reported figures (Pittman and Handy, 1965). As early as 1931 the Wickersham Commission called for the creation of a centralized data collection service and rigorous data-quality control (United States National Commission on Law Observance and Enforcement, 1931). The misreporting and under-reporting apparently endemic in current official statistics has led to their widespread devaluation.

## SURVEY MEASURES OF CRIME

Continuing dissatisfaction with official measures led to the development of alternative techniques to gauge the scope and distribution of crime. The most important of these is the population survey, a measuring device (with its own characteristic reliability and validity problems) which yields different pictures of crime.

The use of sample surveys to study crime reflects dissatisfaction both with the accuracy of official figures and the paucity of information they reveal. The yearly *Uniform Crime Report* does not speak to questions about the characteristics of victims of crime. Offender data is available only on arrestees, although victim testimony might shed some light on the characteristics of successful criminals. Finally, little data is reported on the physical and social circumstances under which most crimes occur, even though this has tremendous implications for their solution and deterrence.

It appeared to the President's Crime Commission that population surveys potentially could speak to all of these inadequacies, and in the mid-1960's the Commission funded several pilot projects and a national sample survey to test their utility.<sup>2</sup> Since then, the federal government has inaugurated a regular surveying program on a national scale and has funded several local and state-level investigations.<sup>3</sup>

It was inevitable that the victim-based data gathered by these large-scale surveys would be used to gauge police-reported crime statistics. Suspicion of official statistics has become widespread and appreciation of the errors in crime data particularly well-known, much more so than the caution of researchers who regularly employ attitude measures and self-reports of behavior. The latter deal skeptically with data and demand elaborately scaled, multiple-item indicators of concepts before they use them with any confidence. The questions on the Crime Panel surveys elicited a much larger volume of events than reported by police, so it is widely assumed that they are "more accurate" measures of the true volume of crime in society. But such gaps are inevitable. Despite the surface similarity of the resulting figures, the measurement operations and their errors differ greatly when we compare police and survey procedures for estimating crime rates. The social and organizational processes which stand between events occurring in the world and our survey or official maps of them produce quite different kinds of crime statistics.

### SOURCES OF MEASUREMENT ERROR

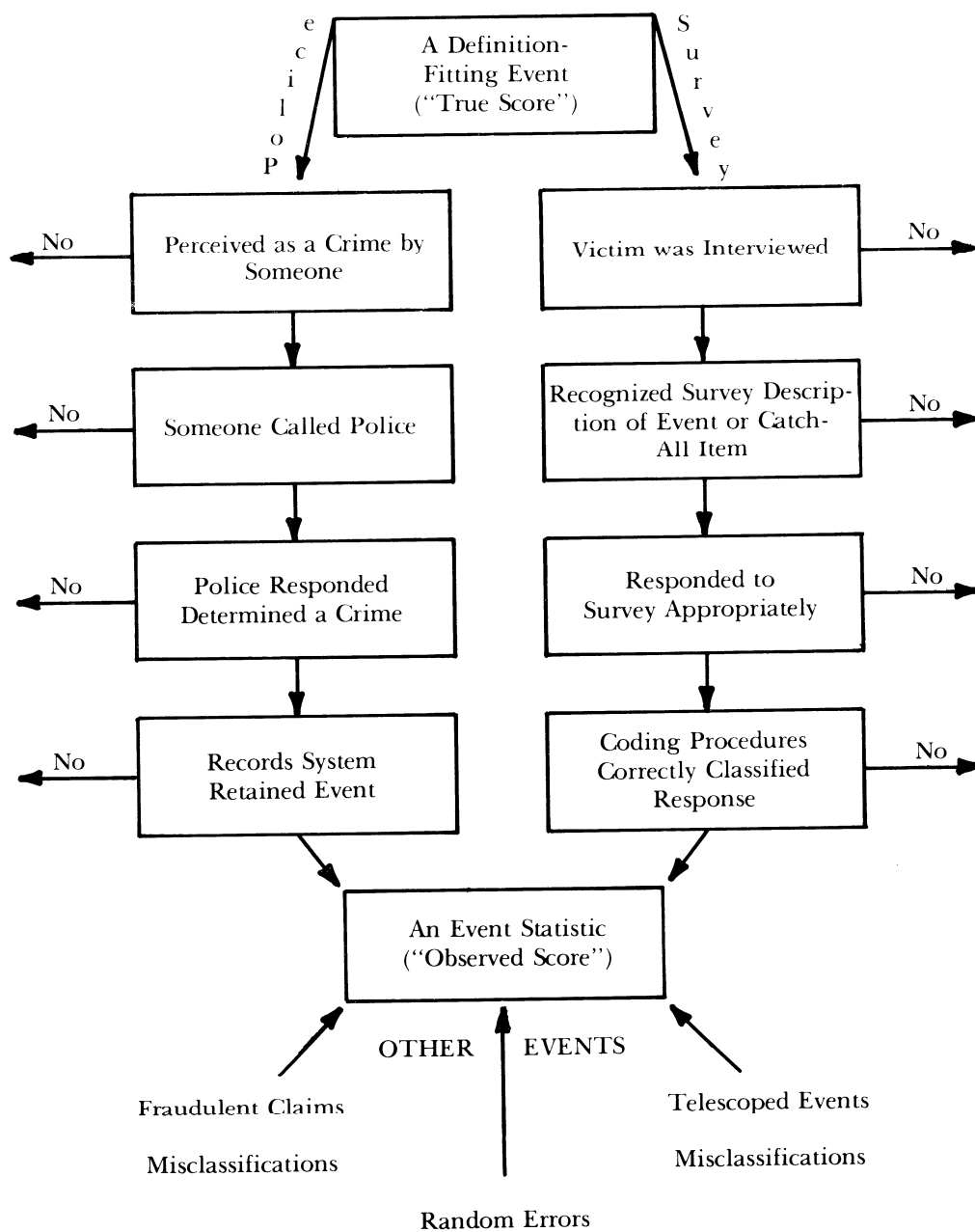
In the course of mapping crime events into a numerical system, both official and survey measurement procedures generate considerable error. If we think of error as the gap between a true score and an observed score for an event, Figure 1 may be a useful summary of what we know about its sources. At each step, an exit from the measurement process leads to error. On the survey side, measurement error has been investigated in a series of pilot studies which began in 1966. Our knowledge of error-generative processes on the police side is older, but it has been enhanced considerably by studies of victim behavior and systematic observations of police work during the past decade.

The first stages in the official measurement process lie in the hands of civilians: the victims of crime, their relatives, neighbors, and bystanders. The first public filter through which events must pass is perceptual: someone must know that a specific incident has taken place. This is in part an information problem. For example, a great deal of larceny from commercial establishments (shoplifting and employee theft) is discovered only in the form of inventory shrinkage (Dodge and Turner, 1971). In this case we know that crime is taking place, but *events* remain unknown and uncountable. The general difficulty is that discrete events may escape detection, while continuous indicators of their occurrence—like dollar losses per quarter or shortages at audit—cannot be enumerated under our current system of social accounts. The problem is also conceptual: people must define an event as falling into the domain of events about which "the police must do something." This appears to inhibit the reporting of consumer fraud, and it is the difference between crime and "ripping-off." Attitudinal studies of the legitimacy of theft or fraud upon large private and governmental bureaucracies indicate that there is far from universal agreement about the labeling of some behaviors in our society (Smigel and Ross, 1970). The problem of who does the perceiving is also of interest. Pilot surveys in Dayton and San Jose revealed that 25% of all personal crime and 20% of all property crime is reported by someone other than the victim (*Crimes and Victims*, 1974). The motives which lead non-victims to send for the police are simply unknown.

The decision to call the police has been the focus of considerable research, for it is probably the most important factor shaping official statistics on crime. In the Dayton-San Jose pilot surveys conducted in 1972, victims recalled that about 60% of all robbery, 56% of all larceny, and 40% of all household burglaries were *not* reported to the police. Their reasons

FIGURE 1

## SOME SOURCES OF MEASUREMENT ERROR



for failing to do so were numerous: the largest categories chosen were "not serious enough" (25-30%), "nothing can be done" (25%), or that the harm or loss was slight (10%) (*Crimes and Victims*, 1974). Other analyses of the reporting problem have focused upon the race, class, or even personality characteristics of victims rather than their manifest responses, although the utility of this approach is not particularly clear. It appears that the characteristics of the *event* are controlling: who did it (relative or stranger); why it was done (economics or passion); what was the damage to person, property, or propriety; and what were the participants' estimates of the burdens and benefits of invoking the police? Only a portion of the latter calculation—that involving the victim's fear of the police—would appear to be a straightforward race-and-class problem. Despite much discussion of this factor, neither Ennis' (1967) national survey nor the Dayton and San Jose studies revealed more than 2% giving that response (*Crimes and Victims*, 1974).

Observational studies of police behavior indicate that even after the police are called the outcome of the crime-measurement process remains problematic. Crime recording becomes an organizational activity. Black's (1970) and Black's and Reiss' (1970) descriptions of police-citizen encounters in Chicago, Boston, and Washington, D.C. indicate that extra-legal factors greatly influence a policeman's decision to write a formal report. They are loathe to file a report when the relational distance between the participants in a dispute is small, in part because they know that it is very unlikely that the case will be pursued in the courts. They tend to defer to the dispositional preferences of the complainant, who often mobilizes the police only to warn or threaten another party. Both complainants who are deferential to the police and higher-status victims are more likely to be successful in persuading the police to file a report. The police also act upon their own assessment of the complainant's culpability. Often responsibility for personal crimes or their outcomes may be apportioned among the parties, and police respond to the division of blame (Curtis, 1974). Finally, in cases where juveniles are parties to a dispute the police tend to defer to the dispositional preferences of adults at the scene.

These observations suggest another reason why official statistics on crime should be lower than survey estimates. Unlike survey enumerations, where the victim's claim ultimately must be recorded on his terms, police "measurement" takes place within the context of the event. Complainants are surrounded by witnesses and bystanders who contribute their interpretations of events. Surprisingly often the suspects themselves are present to offer countercharges and alternative explanations. The decision to file a formal report is "judicial" in the sense that an officer weighs claims and counter-claims before making a disposition in a case. Patrol officers quickly learn to be suspicious of the motives of complainants, for their authority is often invoked for private purposes; claims of victimization are not taken at face value (Rubinstein, 1973). As the *Uniform Crime Report* does not present predisposition case totals, but only "founded" complaints for each city, we have no idea of the dimensions of this process. Scattered reports of large departments on hand indicate that the effect of "unfounding" is considerable: approximately 25% of rapes, 13% of robberies, and 19% of gun assaults reported to the police were discounted in these cities. They probably would generate self-reports of victimization, but they did not become official statistics.

Technical considerations, including difficulties with the classification scheme employed in gathering official statistics, may introduce measurement errors on the police side as well. The Uniform Crime Reporting System imposes a set of definitions which do not match the legal pigeon-holes into which the police must sort events. The translation from local to national terminology appears to vary from jurisdiction to jurisdiction,

enhanced by local differences in training and data quality control (Federal Bureau of Investigation, 1973). Errors of this sort will shift over time within cities as well. I would interpret the tremendous variation and apparently random distribution of "manslaughter by negligence" totals reported in the *Uniform Crime Report*, for example, to be a function largely of variations in local practice.<sup>4</sup> Survey studies of crime, on the other hand, utilize standardized measurement operations which may vary among interviewers, but should not vary considerably across cities. Because these error terms differ, further "gaps" will appear between figures from the two sources.

The final source of error on the police side is organizational and political. The ability of official records systems to retain information once it has been entered is problematic. In 1966, a department audit of stationhouses in New York City revealed 20-90% underreporting of events in their files (Wolfgang, 1968). These and other discoveries suggest that crime is an organizational problem in police departments. Especially in cities where commanders are evaluated on their ability to reduce crime, we observe a consistent tendency toward underreporting or the down-grading of offenses by police departments (Seidman and Couzens, 1974). Events also disappear individually in response to political influence or bribes, but this is less likely to skew the totals in common types of crime.

The dramatic impact of variations in police record-keeping procedures upon crime statistics is illustrated by "before-and-after" studies of cities which have overhauled their systems. Many of these were noted by researchers for the Crime Commission in their discussion of crime statistics (President's Commission, 1967). New York City's 1950 reorganization, for example, boosted that department's robbery totals by 400%, larceny 700%, and assault with a weapon 200% (Wolfgang, 1968). The Commission correctly perceived such overhauls as part of a more general phenomenon: the increasing professionalism of big-city police departments. A working hypothesis would be that as departments centralize their administration, automate their information systems, and encourage more legalistic behavior on the part of beat patrolmen, error in the official measurement of crime should be reduced significantly.

## SURVEY MEASUREMENT

The sources of measurement error on the survey side have been investigated in a series of national and city-level studies. In some, alternative techniques are employed in different random samples of a population and the results are compared. In others, police records are sampled to locate respondents who are known to have been victimized. They are then interviewed and their recall patterns analyzed. Each method gives us a different check of the reliability and validity of survey measures of crime.

These investigations suggest that the first question we must ask is, "Will the victim be interviewed?" This raises both data collection and sampling problems. In early studies, a randomly selected adult often was used as an informant for an entire household. Interviewers quizzed this respondent about the victimization experiences of each family member. In the Dayton-San Jose surveys, half of the sample households in each city were completely enumerated; interviewers questioned every household member over the age of 13 to elicit self-reports of victimization. Apparently, informant fatigue or lack of information about other household members is a substantial problem, for individual questioning elicited significantly more events. The differences were so marked that future federal surveys will employ complete household enumerations despite their increased cost.

Sampling deficiencies, on the other hand, have not been remedied. In the city-level studies conducted by the Bureau of the Census household sampling procedures are employed, and the sampling frame is bounded by the territorial limits of the central city. But an average of 13% of the daytime populations of the nation's core cities are commuters (Kasarda, 1972). In Chicago, for example, over 400,000 workers leave the city at sundown. Tourists and other transients account for another fraction. Although they may be victimized and can report their experiences to the police, they are currently not eligible for interviewing.

Even if they enter the sample, victims of crime may not recall the event. As Albert Biderman et al. (1967) have noted, one striking finding of the victimization pretests was the relatively low salience of many crime events. In practice, most respondents seem to find it difficult to remember incidents of victimization other than recent cases. The problem of memory fade has been investigated in two ways. First, known victims have been selected from police reports and interviewed. Their recall rates have climbed from 62% (Washington) to 74% (San Jose), reflecting successive improvements in the Census Bureau's questionnaire. Second, respondents have been required to recall known events within time frames ranging from three months to one year. These tests reveal a sharply decreasing recall rate for temporally distant events. The same phenomenon may be observed by plotting the date of occurrence of each event recalled by randomly selected respondents. Monthly crime rates estimated from survey responses drop sharply as an inverse function of time (Ennis, 1967). Accurate survey measurements require brief recall periods. This means that very large samples are required to provide yearly crime estimates. The current compromise for the National Crime Panel is six months; respondents in the city studies are asked to recall events for an entire year. Police estimates, on the other hand, are subject to few of these difficulties.

Reverse checks of police records also indicate that recall rates in an interview setting are sensitive to variations among the events themselves. They suggest that responses may not be forthcoming even if an event is recalled. Victims appear to be unwilling to report clashes with friends or relatives, for example. In San Jose, those whom the police noted had been victimized by strangers recalled the event during an interview 75% of the time; only 22% of the cases where the police recorded that the offender was a relative were recalled, and 58% of those cases involving an acquaintance. Rapes were revealed cautiously; in the San Jose pilot survey all recalled rapes were described as "attempted." It should be noted that these variations are similar to those which appear to affect the willingness of victims to relate their experiences to the police as well. Disputes within families and rapes are both highly under-reported. And, as it was noted above, the police appear to be less willing to file formal reports when disputants are acquainted. In this case, survey and official procedures both systematically undercount the same classes of events. This is a serious measurement problem.

As noted in Figure 1, the final step in the survey measurement of crime involves the coding and classification of reported victimizations. It is difficult to judge how successfully this process reflects the event. In his report to the Crime Commission, Ennis (1967) related a modest test of the inter-coder reliability of his classification scheme. Teams of lawyers and detectives were successful in classifying citizen-reported victimizations in the same U.C.R. categories as his research staff about 65% of the time. In a validity test of the more advanced San Jose Survey instrument, Census personnel classified 259 of 292 recalled victimizations into the same categories as the local police who initially recorded them. Since we have no confidence that police and the interviewer were told exactly the same story, this is a remarkable correspondence. Coupled with the face validity of the current survey



instrument—the items are drawn to tap the dimensions which define Part I offenses in the *Uniform Crime Report*—this suggests that the classification stage of the process is probably less troublesome than most.

A final and potentially important source of error in both survey and official measures is the intrusion of other events into the observed score for a city or household. On the police side, fraudulent claims may be registered. People may misuse the police in personal vendettas, they may invent stories to disguise their own culpability, or they may attempt to register excessive insurance claims. In addition, actual events which lie outside the domain of interest may be misclassified as falling within it. The most serious problem on the survey side is “forward telescoping.” Method checks of all kinds indicate that the tendency of respondents to recall events which occurred outside of the reference period of the survey and to claim that they occurred within the specified interval is quite strong. Experiments with the Census’ Quarterly Household Survey panels indicate that “bounded” interviews may avoid distortions of this kind. Respondents who are asked to recall events which have occurred since an interviewer’s *last visit* report as few as one-half the number of victimizations recalled by those who are quizzed about the same period but who previously have not been questioned (Turner, 1972). Given the low salience of most crime events and their steep forgetting curve, victims require signposts to guide their recall.

#### ESTIMATING ERROR MAGNITUDE

Like all measures, estimates of crime rates contain error. Given the magnitude of the sources of error discussed here, it is remarkable that official and survey measures of crime covary as closely as they do. The existence of these multiple measures may help us estimate in very rough fashion the magnitude of the error in each. Additional methods tests and analyses of existing data may contribute further to our understanding of the dimensions of error.

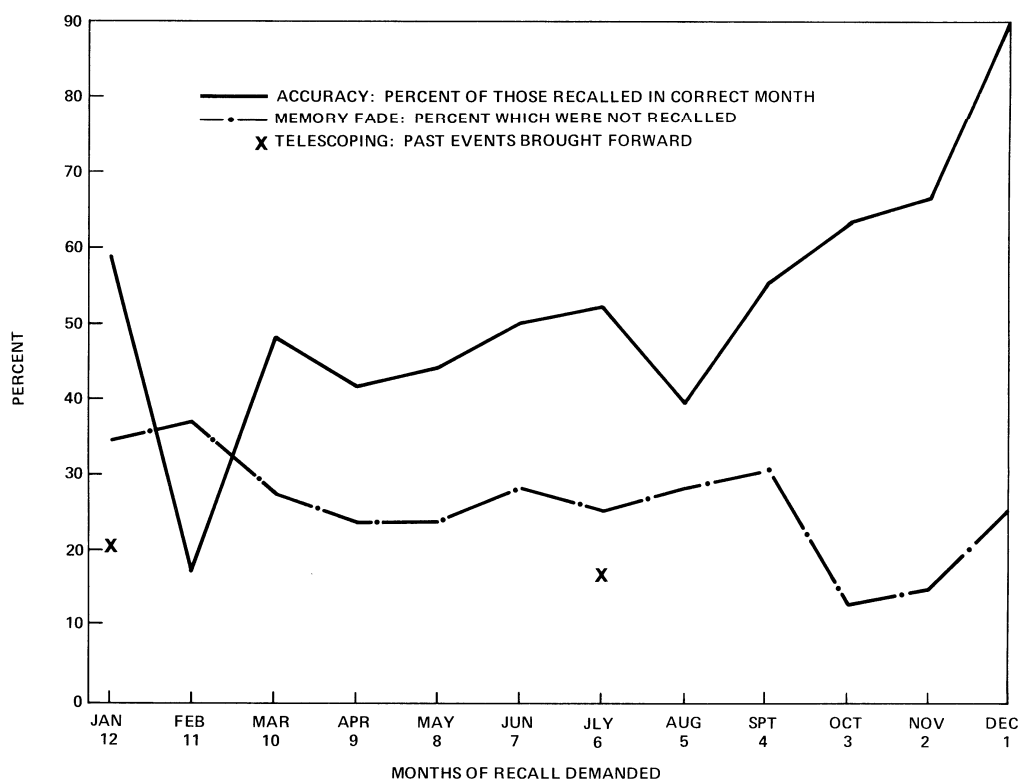
Crosschecks of recall errors in the survey measurement process indicate that the rate at which interviews “recover” events is fairly high. In the San Jose pilot survey of 1971, of the 394 known victims who were located for questioning, 292 recalled the event in some form (San Jose Methods Test, 1972). Table 1-A presents the recall rate for various subcategories of events. Note that rates for frequent crimes, larceny and burglary, were higher than those for less frequent events. Table 1-A also presents the total number of personal and household victimizations recalled by the residents of San Jose proper in the standard population survey phase of the pilot study. These are then projected into “corrected” totals which roughly take into account patterns of non-recall. As the column totals indicate, the San Jose survey may have recovered approximately 75% of the five classes of events of interest. This is a *very* rough indicator of the recovery power of the victimization survey instrument, one that requires further refinement.

Table 1-B examines the respondent’s contributions to errors in survey measures of crime. The memory curve plotted in Table 1-B indicates that recall periods exceeding three months may lead to the substantial undercounting of offenses in the population. A test of the ability of those recalling events to place them in the proper month—an essential check of the ability of surveys to provide time-series estimates of the type anticipated—indicates that recall accuracy degrades sharply after about three months as well (Turner, 1972; Ennis, 1967). These curves, which were computed from data in the report of the San Jose pilot study, suggest that the six-month recall period used in the National Crime Panel and the 12-

TABLE 1  
ESTIMATES OF SURVEY MEASUREMENT ERROR

1-A: Interview Recall by Type of Household and Personal Victimizations: City of San Jose Only			
Type of Crime	Total Recalled <sup>1</sup> Victimizations	Recall <sup>2</sup> Rate	Projected Victimizations
Rape	100	67	149
Robbery	2840	76	3737
Assault	8980	48	18708
Burglary	17610	90	19567
Larceny	45900	81	56667
Total	75430		98828
Event Recall = 75430/98828 = 76%			

1-B: RECOVERY RATE AND ACCURACY OF INTERVIEW RECALLS BY TIME:  
SAN JOSE AREA AND WASHINGTON, D.C.<sup>3</sup>



1-C: *Sampling Frame Loss Estimate:  
Motor Vehicle Theft in Chicago, Illinois, 1972*

Vehicle Registration	1,260,000 <sup>4</sup>
Survey Theft Estimate	38,700 <sup>5</sup>
Loss Probability	.03
Total Commuting Autos Daily Entering City	206,000 <sup>6</sup>
Commuter Vehicle Loss at One-Half City Rate	3,090

Source: <sup>1</sup>*Crimes and Victims, 1974*: Table 12 and 39. The San Jose Robbery total presented in Table 12 of that source is clearly incorrect—this is my estimate from the robbery sub-totals. Rape totals are not reported in raw form; 100 is estimated from percentages elsewhere in the report.

<sup>2</sup>*San Jose Methods Test, 1972*: Table C.

<sup>3</sup>Accuracy and memory fade figures were calculated from: *San Jose Methods Test, 1972*: Table 4; telescoping data were reported in: "Victim Recall Pre test," 1970: Table G.

<sup>4</sup>Illinois Secretary of State, 1972.

<sup>5</sup>*Crime in the Nation's Five Largest Cities, 1974*: Table 1.

<sup>6</sup>*Journey to Work, 1972*.

month period bounding the city-level samples may contribute significantly to the error components of those measures. Because these curves were calculated from the same data used to estimate survey recovery rates in Table 1-A, it is impossible to untangle here the distinct contributions of the salience of events *and* their temporal distribution, however. The estimates of the magnitude of forward telescoping error presented in Table 1-B are based upon the Washington, D.C. pilot survey. There, 17% of the victimizations recalled by selected respondents occurred before the indicated cut-off point when a six-month limit was specified, and 21% telescoped in events which occurred before a 12-month limit. As I noted before, telescoping effects—which lead to an overcounting of events—can be controlled by "bounding" the recall period with a salient event. The National Crime Panel utilizes the previous visit of an interviewer, while city-level interviews must rely upon verbal instruction. The latter surveys are much more likely to overestimate crime rates due to telescoping errors.

Error introduced at stages preliminary to the interview are more difficult to estimate. Sampling errors for individual cities are introduced by the systematic elimination of commuters, conventioners, and tourists from the sampling frame. The effect of this loss upon one crime statistic, motor vehicle theft, is very roughly estimated in Table 1-C, where motor vehicle statistics for the city of Chicago are presented. The victimization survey of Chicago projected that residents there suffered about 38,700 vehicle thefts in 1972, or a loss probability of .03 per motor vehicle. Projecting commuter vehicle losses at only one-half of the rate for city residents, it appears that excluding commuters probably undercounts victimization by about 8% for this offense. Other crimes dealt with in the victimization surveys cannot be so easily projected. Commuters are susceptible to personal larceny,

robbery, assault, and rape, probably in that order; conventioners and tourists may be more at risk than commuters in all of these categories.

The problem of systematically eliminating potential victims of crime from the sample is compounded when we consider the distribution of known victimizations. The surveys indicate that many offenses, most notably assault and robbery, disproportionately affect young males; they are also a demographic group which is difficult to enumerate in a household survey. A reasonable estimate would be that we undercount by 5% or more due to sampling limitations. Together, the sources of error on the survey side of Figure 1 probably accumulate to undercount events by 30%.

Reversing the analysis enables us to probe the magnitude of error terms on the police side. The most satisfactory test would reverse the record-check procedures utilized in the pilot surveys: follow-up studies of offenses which were apparently reported to the police would be conducted to determine which could be found in police records. None of the police departments which have extended their cooperation to researchers has granted access to its records on this scale, and most police record-keeping systems are not organized in ways that make this a simple task. An analysis of the marginal frequencies of reported and officially recorded events in some of the sample cities suggests that the gap between the two sources would be considerable.

Many events which occur and are said to have been turned over to the police do not appear to survive police processing. Official and survey estimates of city crime rates consistently differ. The ratio of robberies recalled in interviews to robberies known to the police in Portland—one of the smallest cities analyzed here—was greater than 3-to-1 in 1971-72; in New York City, more than 2-1/2 robberies were recalled to interviewers for every event recorded by the police. Table 2 presents official and survey robbery estimates for five cities surveyed in 1972 as part of the Bureau of the Census' Large City study.

The victimization surveys asked each victim of a crime whether the event was reported to the police. This reporting rate can be used to "correct" survey estimates of the crime rate in each of the five cities for citizen-induced errors in police figures. While it is socially desirable to respond under questioning that one reported a crime, which will inflate this figure somewhat, it is clear from Table 2 that substantial distance remains between citizen "reported" and police recorded crime. The gap was smallest in Detroit, where police accounts of robbery added up to 73% of "reported" (by recall) by city residents. Philadelphia's extremely low robbery count, which amounted to only 37% of what her residents claim to have reported to the police, may be related to numerous charges that police there cheat on their statistics (Seidman and Couzens, 1974).

The figures presented in Table 2 suggest that "crimes known to the police" are probably not very accurate indicators of the true volume of crime in a community. In the case of robbery, official totals accounted for an average of only 38% of the victimizations recalled by citizens of these five communities. That figure varied considerably across cities: official robbery totals added up to 48% of the survey figure in Detroit, and only 22% in Philadelphia. If the interview recovery rate for individual and commercial robbery in these cities approaches the individual rate in San Jose (76%—See Table 1), official robbery totals might amount to an average of only 28% of the true total, and that figure might drop to only 17% in Philadelphia. The correction for error induced by citizen reporting practices improves this picture somewhat—official figures may reach 63% of "reported" robbery—but it appears that organizational processes contribute considerably to error in police-recorded crime statistics.

TABLE 2  
ESTIMATES OF OFFICIAL MEASUREMENT ERROR

<i>City</i>	<i>Total<sup>1</sup> Official Robbery</i>	<i>Total<sup>2</sup> Survey Robbery</i>	<i>Survey<sup>3</sup> Measure of Reporting Rate</i>	<i>Estimated "Reported" Robbery</i>	<i>Official As Percentage of Survey "Reported"</i>
Chicago	23531	64100	57.5	36881	64
Detroit	17170	36100	65.4	23638	73
Los Angeles	14241	36400	55.1	20064	71
New York City	78202	191400	59.5	113863	69
Philadelphia	9710	44000	58.9	25914	37
Average	28571	74400	59.3	44072	63

Source: <sup>1</sup>Federal Bureau of Investigation, 1973.

<sup>2</sup>*Crime in the Nation's Five Largest Cities, 1974*: Table 1. This includes both individual and commercial offenses.

<sup>3</sup>*Crime in the Nation's Five Largest Cities, 1974*: Table 8 (recomputed). This rate includes both individual and commercial reports.

## SUMMARY

This paper has attempted to draw together a number of lines of research and apply them to a measurement problem common to many researchers and administrators—the errorful nature of crime statistics. While it has long been known that police measures of crime are shaped by social and organizational processes, the persistence of similar kinds of error in new, survey-based crime indicators is less widely understood. Police statistics are shaped by citizens, who must interpret and report events, by patrol officers who must classify and record them, and by bureaucracies which must store and retrieve them and live with their political consequences. In order to measure crime through survey techniques, researchers must locate their victims, elicit thorough and accurate responses to questionnaire items which adequately define the concepts, and classify them in ways which resemble the official methods.

In the process, both measurement techniques seem to undercount systematically certain types of crimes—those which occur within families or among acquaintances, or which are difficult to identify as discrete events. Others are likely to be under-enumerated or over-enumerated in different ways: surveys will miss commuter victimization while inflating events in which the informant was in fact culpable; the police will tend to discount the latter, but their on-the-spot coverage of events will avoid many of the problems induced by memory fade.

The existence of alternative techniques for assessing the volume and distribution of crime enables us to identify and more fully understand the magnitude of these difficulties. Comparisons of victim self-reports with police records allows us to pick optimal recall periods for accurate survey measures, and to estimate the magnitude of patterned non-recall and its social origins. Comparisons of survey-generated and official crime rates will enable us to gauge the impact of bureaucratic processes upon the recording and releasing of crime data, especially as the number of cities which are surveyed increases. In addition, the identification of the limitations of the sampling frame indicates that the extension of the city samples to include the suburban fringe would further decrease discrepancies between official and survey estimates of the volume of events occurring in the central city. The existence of multiple measures of the same underlying distribution of events has numerous methodological payoffs.

## FOOTNOTES

- <sup>1</sup> A revised version of a paper presented at the meetings of the American Statistical Association in St. Louis, August 26, 1974. This revision was completed while the author was a Visiting Fellow at the Law Enforcement Assistance Administration. That agency bears no responsibility for its content.
- <sup>2</sup> Crime Commission Surveys included those of Washington, D.C. (Bideman et al., 1967), high crime areas of Chicago, Washington, and Boston. (Reiss, 1967a and 1967b), and a national sample (Ennis, 1967).
- <sup>3</sup> These included pilot surveys (*Victim Recall Pretest*, 1970; *Household Survey of Victims of Crime*, 1970, *San Jose Methods Test*, (1972) and a full-scale shake-down test (*Crimes and Victims*, 1974). Since then there have been two reports of multi-city studies (*Crime in the Nations' Five Largest Cities*, 1974); *Crime in Eight American Cities*, 1974). Local studies have been conducted in North Carolina (Richardson et al., 1972) and Seattle (Hawkins, 1970).
- <sup>4</sup> New York City, for example, has twice the population of Chicago and reports 3-1 2 times as much crime, but Chicago in 1972 recorded 260 negligent manslaughter cases while New York recorded 66. In general, manslaughter rates do not correlate with anything.

## REFERENCES

- Bideman, A.D., et al. (1967). *Report on a Pilot Study in the District of Columbia on Victimization and Attitudes Toward Law Enforcement*. United States Government Printing Office, Washington, D.C. 20402.
- Black, D.J. (1970). "The Production of Crime Rates." *Amer. Soc. Rev.*, 35:733-748.
- Black, D.J. and Reiss, A. (1970). "Police Control of Juveniles." *Amer. Soc. Rev.*, 35:63-77.
- Bohrnstedt, G.W. (1970). "Reliability and Validity Assessment in Attitude Measurement." In: G.F. Summers (Ed.) *Attitude Measurement*, pp. 80-99. Rand-McNally, Box 7600, Chicago, IL. 60680.
- California Criminal Statistics Bureau (1967). *Bank Robbery in California: A 35-Year Comparison of California with the Rest of the United States and an Intensive Study of 1965 Offenses*. California Criminal Statistics Bureau, Sacramento, CA. 95801.
- Campbell, D.T. and Fiske, D.W. (1959). "Convergent and Discriminant Validation in the Multitrait-Multimethod Matrix." *Psych. Bull.*, 56:81-105.
- Chambliss, W. and Nagasawa, R.H. (1969). "On the Validity of Official Statistics." *J. of Res. on Crime and Delinq.*, 6:71-77.
- Crime in Eight American Cities*. (1974). National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, Washington, D.C. 20530.
- Crime in the Nation's Five Largest Cities*. (1974). National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, Washington, D.C. 20530.
- Crimes and Victims: A Report on the Dayton-San Jose Pilot Survey of Victimization*. (1974). Law Enforcement Assistance Administration, Department of Justice, Washington, D.C. 20530.

- Curtis, L.A. (1974). "Victim Precipitation and Violent Crime." *Soc. Probs.*, 21:594-605.
- Dodge, R.W. and Turner, A. (1971). "Methodological Foundations for Establishing a National Survey of Victimization." Unpublished paper presented at the 1971 meetings of the American Statistical Association.
- Ennis, P.H. (1967). *Criminal Victimization in the United States: A Report of a National Survey*. United States Government Printing Office, Washington, D.C. 20402.
- Federal Bureau of Investigation. (1973). *Uniform Crime Report 1972*. United States Government Printing Office, Washington, D.C. 20402.
- Hawkins, R.Q. (1970). "Determinants of Sanctioning Institutions for Criminal Victimization." Unpublished Ph.D. dissertation, Department of Sociology, University of Washington.
- Household Survey of Victims of Crime*. (1970). Demographic Surveys Division, Bureau of the Census, Washington, D.C. 20233.
- Illinois Secretary of State. (1972). *Annual Report*. Springfield: Illinois Secretary of State. 62701.
- Journey to Work*. (1972). United States Bureau of the Census, Washington, D.C. 20233.
- Kasarda, J.D. (1972). "The Impact of Suburban Population Growth on Central City Service Functions." *Amer. J. of Soc.*, 77:1111-1124.
- Morgenstern, O. (1963). *On the Accuracy of Economic Observations*. Princeton University Press, Princeton, N.J. 08540.
- Pittman, D.J. and Handy, W.F., (1965). "Uniform Crime Reporting: Suggested Improvements." In: A. Gouldner and S.M. Miller (Eds.), *Applied Sociology*, pp. 180-188. Free Press, 866 Third Avenue, New York, NY 10022.
- President's Commission on Law Enforcement and Administration of Justice. (1967). *Task Force Report: Crime and Its Impact—An Assessment*. United States Government Printing Office, Washington, D.C. 20402.
- Price, J.E. (1966). "A Test of the Accuracy of Crime Statistics." *Soc. Probs.*, 14:214-221.
- Quinney, R. (1970). *The Social Reality of Crime*. Little Brown, 34 Beacon Street, Boston, MA 02106.
- Reiss, A. (1967a). *Measurement of the Nature and Amount of Crime*. United States Government Printing Office, Washington, D.C. 20402.
- (1967b). *Public Perceptions and Recollections About Crime, Law Enforcement, and Criminal Justice*. United States Government Printing Office, Washington, D.C. 20402.
- Richardson, R. et al. (1972). *Perspectives on the Legal System: Public Attitudes and Criminal Victimization*. Institute for Research in Social Science, University of North Carolina, Chapel Hill, N.C.
- Rubinstein, J. (1973). *City Police*. Farrar, Strauss, Giroux, New York.
- San Jose Methods Test of Known Crime Victims. (1972). *Statistics Technical Report No. 1*. National Institute of Law Enforcement and Criminal Justice Statistics Division, Law Enforcement Assistance Administration, Washington, D.C. 20530.
- Seidman, D. and Couzens, M. (1974). "Getting the Crime Rate Down: Political Pressure and Crime Reporting." *Law and Soc. Rev.*, 8:457-493.
- Snigel, E.Q. and Ross, H.L. (1970). *Crimes Against Bureaucracy*. Van Nostrand, 450 West 33rd Street, New York, NY 10001.
- Turner, A. (1972). "Methodological Issues in the Development of the National Crime Survey Panel: Partial Findings." National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, Washington, D.C. 20530.
- United States National Commission on Law Observance and Enforcement. (1931). *Report on Criminal Statistics*. United States Government Printing Office, Washington, D.C. 20402.
- Victim Recall Pretest*. (1970). Demographic Surveys Division, Bureau of the Census, Washington, D.C. 20233.
- Wollgang, M.E. (1968). "Urban Crime." In: J.Q. Wilson (Ed.), *The Metropolitan Enigma*, pp. 245-282. Harvard University Press, 79 Garden Street, Cambridge, MA 02138.