Self-Report Surveys as Measures of Crime and Criminal Victimization

by David Cantor and James P. Lynch

ABSTRACT

Self-report surveys of victimization have become commonplace in discussions of crime and criminal justice policy. Changes in the rates at which residents of the country are victimized by crime have taken a place alongside the Federal Bureau of Investigation index of crimes known to the police as widely used indicators of the state of society and the efficacy of its governance. While a great deal has been learned about this method for producing data on crime and victimization, a number of fundamental issues concerning the method remain only partially explored. This paper outlines what we have learned about victimization surveys over the past 30 years and how this source of information has been used as a social indicator and a means of building criminological theories. It also identifies major methodological issues that remain unresolved and suggests some approaches to exploring them. The evolution of the National Crime Victimization Survey is used as a vehicle for this discussion, because the survey has been conducted continuously for 25 years and has been the subject of extensive methodological study.

David Cantor is an Associate Study Director with Westat in Rockville, Maryland. James P. Lynch is a Professor in the School of Public Affairs with American University in Washington, D.C.

A Review of Self-Report Surveys

Self-report measures of criminal victimization have become widely used social indicators and research tools in criminology and criminal justice. A great deal has been learned about the strengths and weakness of this methodology. Substantial improvements have been made in this methodology since its inception in the late 1960s, yet problems and limitations persist. This essay examines the evolution of the victimization survey methodology. It identifies (1) the contribution that these surveys have made to our understanding of crime and victimization, (2) what we have learned about the methodology, (3) what more we need to know, and (4) what additional research could be done to help us know it.

Assessing the self-report methodology in all its guises would require many more pages than we have been allotted.[1] Consequently, we will focus on household surveys of the general population for the purpose of continuing statistical series on the incidence and characteristics of criminal victimization. The National Crime Survey (NCS), conducted for the U.S. Department of Justice's (DOJ's) Bureau of Justice Statistics (BJS), will be at the center of our attention, along with its immediate precursors and their genesis.[2]

NCS has many design features that are not employed in other large-scale household surveys of victimization. Settling on many of these features involved conscious tradeoffs between data quality and the costs of administering the survey within the environment of the U.S. Census Bureau. Some of these decisions were supported by extensive testing to determine the effects of varying design features on the reporting of crime and the feasibility of fielding a survey with given features. Knowledge of this methodological work is important for understanding the current state of self-report victimization surveys. The uniqueness of NCS will be described briefly. A more detailed description of the history and evolution of the survey will follow.

Contributions of the survey method to our understanding of crime

Victim surveys substantially changed the definition of crime and the nature of the information available on crime events. Prior to the availability of victim surveys in the United States, much of our information on the volume and nature of crime came from the police, specifically the Uniform Crime Reports (UCR). Since 1929, UCR compiled statistics submitted by participating local police departments on offenses known to the police, persons arrested, and officers killed or assaulted. This information was collected on a subset of crimes that the International Association of Chiefs of Police (IACP) at the time considered prevalent, serious, and well reported to the police. Consistent with the technology available at the time, local police departments submitted aggregate counts of offenses known for seven categories of crime: homicide, rape, robbery, aggravated assault, burglary, larceny, and motor vehicle theft. These annual counts by jurisdiction of crimes known to the police were the principal product of UCR, and they were used by many to assess the level and change in level of crime in the United States.[3]

The events defined as crime and the information collected on these events were shaped by the needs of police organizations. The surveys had a different set of limitations related to the survey enterprise. Police record systems available at the time included only those events reported and recorded by the police, collected data on a selected subset of crime, and data presented as aggregate counts of crimes. The victim surveys included events that were reported to the police as well as those that were not. They included extensive information on victims and the social context of the crime and made those data available on an incident or victim basis. These surveys gathered data from victims and nonvictims. All of these differences increased the utility of available data on crime as a social indicator and for research.

In this section, the major contributions that victim surveys have made to criminological theory and policy are reviewed. Given space limitation, this review takes a broad-brush approach. For more detailed reviews, at least through the mid-1980s, the reader is referred to several other excellent reviews (Gottfredson 1986; Sparks 1981).

Implications for crime as a social indicator

One of the major functions of crime statistics is to provide a social indicator. Crime statistics serve this function by providing estimates of the level and change in one aspect of the well-being of a nation, state, or locality. Victim surveys substantially improved the information available on the volume of crime. The data from victim surveys included many crimes that were not reported to the police or other criminal justice agencies (Biderman and Reiss 1967). Victim surveys also provided more detailed information on crime events than did national data systems based on police records. These surveys would ask respondents to provide information on themselves, the offenders, the nature of the crime, and the context in which it occurred. While this type of information may have been available in local police files, it was not assembled nationally by agencies like the Federal Bureau of Investigation (FBI) in a form that allowed easy access. Moreover, the detail available in police files varied substantially, depending on the willingness of police officers to ask victims systematically for the specifics of crime events. Sample surveys take much greater pains to ensure that all respondents are asked to provide the same information on every crime. Finally, victim surveys provided event-level data, whereas UCR offered aggregate counts of eligible events in a given jurisdiction. As a result, victim survey data could be reported in various ways, while police counts could not. For example, the survey could present change estimates for lower class, black males in central cities or for females over 50, while UCR could present only a count of crimes by type nationally and for a jurisdiction.

These differences substantially improved our ability to estimate the volume and change in the volume of crime. At last, an estimate that included unreported crime could be made, and estimates could now be made for subgroups as well as for the population as a whole. This went a long way toward improving on police data as a social indicator. The volume of crime could be estimated for young males or whites or American Indians, for example, so that one could assess whether the volume of crime and the change in the volume of crime was the same for everybody. It became clear with the release of the survey data that this was not the case (U.S. Department of Justice, National Criminal Justice Information and Statistics Service 1976). This was a tremendous step forward for the use of crime data as a social indicator.

Finally, the surveys allow for the creation of new ways of classifying crimes other than the ubiquitous index crime classification, which had come to dominate and limit our understanding of the crime problem. The survey could produce estimates of "stranger to stranger" crime, crime "among intimates," "crime at work" or "vehicle-related crime" rather than staying with rape, robbery, aggravated assault, etc. These alternative crime classifications shed a whole new light on crime. Just as population-specific crime rates demonstrated that different groups of people had different crime problems, these alternative classifications showed that there were different problems as defined by the social context of the act that were not visible when events were classified by the criminal act alone.

The availability of victim surveys in tandem with data from administrative police records has raised the level of sophistication among consumers of crime information.

The benefits of the victim surveys as social indicator arose as much from the organization of the survey enterprise as from the enhancement in the information provided. Prior to the institutionalization of victim surveys, crime information was entirely under the control of the criminal justice system. This raised questions about the accuracy and scientific impartiality of the resulting data. Because the police have an immediate and specific interest in the crime problem, there is always the suspicion that they are "cooking the books." Victim surveys brought the "patina of science" to crime statistics. The Census Bureau and survey research agencies were not interested parties with respect to the crime problem, and there was greater trust that the resulting crime estimates were not purposely manipulated.

The availability of victim surveys in tandem with data from administrative police records has raised the level of sophistication among consumers of crime information. They use both of these indicators to try to understand the crime problem and how it might be changing. Initially, the two indicators were pitted against each other as the one true indicator of crime, but gradually this is giving way to the complementary use of the two series (McDowall and Loftin 1992). BJS, for example, is issuing reports that include data on relevant topics from both UCR and NCS (Zawitz 1988). Journalists make references to both indicators in their routine crime stories, and disparities between the police and survey data are taken as issues to be explained rather than used to impeach one or the other statistic (*New York Times* 1981; *Washington Post* 1981). These are all positive signs that the consumers of crime statistics are appreciating the complexity of describing the crime problem and are treating these social indicators with appropriate caution. The depth and breadth of this sophistication is difficult to assess in a period when the two series have tracked each other for a number of years. It will become more clear when the series diverge again. Nonetheless, this movement toward greater sophistication in the production and consumption of crime statistics would not have occurred without the routine availability of victim surveys.

Finally, victim surveys have played a role in the evolution of UCR. In 1984, the FBI undertook a study of UCR for the purpose of improving the system (Poggio et al. 1985). This redesign effort may have been prompted directly by the NCS redesign that had been under way for several years at that time. The NCS redesign was uncovering a great deal about the survey, and it gave BJS the ability to deflect any criticism of the survey by pointing to the redesign as evidence that something was being done about it. UCR had no such protection unless similar efforts to improve the series were undertaken. Moreover, the redesign recommended that UCR adopt a number of the features of NCS, specifically incident-based rather than aggregate reporting. Although this recommendation was affected by the increased use of computing and management information systems in local police departments, it was also an attempt by UCR to match the flexibility of NCS in reporting crime rates.

Implications for building theories of crime and its consequences

Victim surveys have also had a profound effect on theories of crime causation. The availability of highly disaggregated information on crime events, including events not known to the police, facilitated the development of a whole new way of looking at crime. Routine activity, opportunity theory, and even rational choice theories of crime flourished in large part because of the availability of victim survey data (Hindelang, Gottfredson, and Garofalo 1978; Gottfredson 1984; Kennedy and Forde 1999; Hough 1987; Cohen and Cantor 1981; Miethe and Meier 1994; Maxfield 1987; Lynch and Cantor 1992). The surveys also provided an opportunity to identify and investigate the consequences of crime. By documenting the durable and psychic harm that resulted from victimization, the surveys prompted researchers to investigate why the degree of harm differed across crimes and victims (Resnick et al. 1993). Self-report surveys also allowed responses to crime events to become the object of study. Much of the attention was focused on why victims call or do not call the police, but the mobilization of resources other than the police has also been investigated with these data (Skogan 1984).

Theories of crime causation

Self-report victim surveys have contributed to the building of criminological theory. The availability of these data encouraged the development and testing of victim-centered theories of crime (Hindelang, Gottfredson, and Garofalo 1978; Cohen and Felson 1979). These theories focused on the occurrence of crime events rather than criminal motivation. They emphasized the routine activities of victims as sources of opportunity for the motivated offenders. The social, structural, and spatial location of victims influenced their routine activities, which in turn affected their risk of criminal victimization.

Because victim surveys provided a wealth of disaggregated and detailed data on victims and the social context of victimization, they were ideally suited to the testing of routine activity theory. Attributes of persons and social contexts could be used to measure concepts within opportunity or activity theory. This type of detailed information on victims and events was not easily or reliably available from police data. The testing of these theories was facilitated further by the fact that the surveys collected the same information from samples of victims and nonvictims. Using the data from NCS and other surveys, researchers confirmed that the basic tenets of opportunity theory were consistent with the data. The higher the exposure and the lower the guardianship, the greater the probability of victimization.

While opportunity and routine activity theories are a major contribution of victim surveys to criminological theory, the recently revived interest in repeat victimization warrants mention (Pease 1998). Early on in the development of crime surveys, scholars observed that a small number of victims accounted for a relatively large portion of victimization (Sparks 1981; Nelson 1980). A flurry of activity followed, wherein several articles were published demonstrating that the distribution of repeat victimization was greater than would have been expected if the risk of victimization had been equal across persons. This led to speculation that repeat victimization was due to the fact that some people have different risks of victimization, so that persons with higher risk will become repeat victims in higher numbers than others with lower risk. This heterogeneity of risk would account for the distribution of repeat victimization observed throughout the historical development of victim surveys (Biderman et al. 1967; Sparks 1981; Nelson 1980). An alternative hypothesis was that the first victimization exposed the victim to subsequent victimizations, as in the case where a burglary makes the offender aware of other property, which motivates the offender to break in a second time. This was referred to as "state dependence."

Recent research using the British Crime Survey (BCS), as well as police records, has refocused attention on repeat victimization, reasoning that because repeat victims accounted for so much of the cross-sectional crime rate, it would be efficient to target resources to repeat victims and thereby lower the crime rate (Ellingsworth, Farrel, and Pease 1995). This research has found that prior victimization substantially increased the risk of subsequent victimization, and that the time interval between the victimizations was extremely short (Polvi et al. 1990). This work was used to develop police intervention programs that would take advantage of the newfound knowledge about repeat victimization (Forrester, Chatterton, and Pease 1988).

The fact that BCS is a cross-sectional survey limits the degree to which it can be used to investigate the sources of repeat victimization. The British work rekindled interest in repeat victimization in the United States, where the longitudinal data from NCS and other surveys may be more useful in disentangling the relative importance of heterogeneity in risk versus state dependence in explaining repeat victimization (Lauritsen and Davis Quinnet 1995). This research bears watching as an area where victim surveys can contribute to our understanding of why crime occurs (Pease 1998).

Victim surveys have also been useful for shedding light on the composition of the offender population. As part of many victim surveys, respondents are asked about characteristics of the offenders involved in "contact" crimes (i.e., those where the victim actually saw the offender). These data provide a profile of offenders that had not been caught by the police. Analysis has compared offender characteristics collected from victim reports with those provided in official records (Hindelang 1978, 1981). This research found considerable similarity in the characteristics of victims and offenders. That is, "people tend to victimize people like themselves" (Gottfredson 1986, 268).

This line of research has evolved by explicitly linking victimization to offending (Singer 1981, 1986; Lauritsen, Sampson, and Laub 1991; Gottfredson

1984). Analyses of surveys that contain reports of both offending and victimization have shown that reporting offending is linked to reporting victimization, especially for violent events. This relationship is an important theoretical jump that moves toward unifying the disparate discussions of offender motivation and victim risk into a general theory of criminal events.

Responses to victimization

Self-report surveys also offer a unique opportunity for understanding which resources victims mobilize in response to victimization. The principal focus here has been on why the police are called in response to criminal victimization. The most prevalent answer to the question of why people call the police is the seriousness of the event in terms of loss or injury (Skogan 1974). This has been the finding across many different types of surveys in many different countries (Mayhew 1993; Kury 1993; Skogan 1984). This tradition of research using crime survey data suggests that citizens respond to the nature of the crime only so that other factors, such as the perceived legitimacy of the police, are not as important (Garofalo 1977). One of the more interesting findings from a victim survey in this area has been that the nature of prior service by the police affects subsequent willingness to call the police in response to victimization (Conway and Lohr 1994). This analysis was done with longitudinal data from NCS, and it raises anew the question of whether factors outside the crime event can influence the decision to call the police. Perhaps additional analysis of longitudinal data will reveal nuances not visible in cross-sectional data.

Consequences of victimization

Finally, self-report surveys of victimization have been essential to identifying and explaining the consequences of victimization. Here again, the fact that the victim surveys include crimes both reported and not reported to the police provides a more complete picture. The surveys provide a reasonably good picture of the immediate durable harm (i.e., injury and loss) resulting from crime (Harlow 1989). The cost (both to insurance companies and out-of-pocket) of a recent burglary, for example, is captured reasonably well in victim surveys. A number of surveys have assessed various forms of psychic harm that can result from criminal victimization, specifically, sexual assault (Gidycz and Koss 1991; Resnick et al. 1993; Norris, Kaniasty, and Thompson 1997; Finkelhor 1997). These studies have found that depression and posttraumatic stress syndrome are more prevalent among victims of crime than among victims of other traumatic events. Moreover, they have found that some categories of victims (e.g., rape victims) experience more enduring psychic harms than others.

Summary

A large number of the analyses referenced in this section were conducted with large-scale, ongoing household surveys such as NCS or BCS. This was the case in part because these data were easily accessible over a long period of time. These surveys were also extremely large compared with special purpose surveys in the social sciences, and these large samples were required for the study of rare events such as violent crime. NCS had the additional advantage of being the subject of a great deal of methodological study to inform design decisions made both at the inception and over the life of the survey. This methodological work outlined the error structure of the survey so that users of the data could use them appropriately. The error profile of NCS contributed to our understanding of the victim survey method more generally. The next section describes the unique features of the NCS design and subsequent sections review the methodological work done to inform decisions about specific design features of the survey.

NCS had the additional advantage of being the subject of a great deal of methodological study to inform design decisions made both at the inception and over the life of the survey.

Unique Features of the NCS Design

NCS employs a rotating panel design of addresses in which persons in sample households are interviewed at 6-month intervals over $3\frac{1}{2}$ years. All members of the household 12 years of age and older are asked about their victimization experience in the previous 6 months. In addition, one household member is asked to report on the theft of common property as well as on his or her own personal victimization. The survey includes a screening interview in which respondents are asked to recall and report potentially eligible crime events and to fill out an incident form that contains questions about the details of the event. This detailed information is used to determine if the events mentioned in response to the screening interview are within the scope of NCS and how the crime should be classified. The survey data are most commonly used to estimate both the level and change in the level of crime for the seven UCR Index crimes and simple assault.[4]

No other victim survey in the world has the same design as NCS. NCS employs a continuing rotating panel as opposed to a cross-sectional design. Occupants of each housing unit in the NCS sample who are 12 years of age or older at the date the household is contacted (household members, as they reach age 12, are added throughout the survey's $3\frac{1}{2}$-year duration) are interviewed seven times over $3^1/_2$ years. Other victim surveys interview sample units only once and rarely include both children as young as 12 and adults as respondents in the same survey.

NCS employs an address sample taken from the list of addresses compiled in the decennial census and updated throughout the decade by the U.S. Census Bureau. Many other victimization surveys in the United States use lists of telephone numbers or random-digit dialing to obtain their samples. In many other countries, excellent universal national lists simplify sampling. NCS attempts to interview everyone in the housing unit who is 12 years of age or older; most other surveys interview only one (often randomly selected) person in the household, and generally that person is an adult (i.e., 16 or older). The Census Bureau works assiduously to keep noncompletions to a minimum, and its completion rates are rarely matched by any other general population survey.

Respondents in NCS are asked to report all victimizations that occurred during the past 6 months or since the previous interview. In theory, victimization data are to be used only from respondents who had been interviewed 6 months previously. The respondent's recall and a record possessed by the interviewer of incidents reported in the prior "bounding" interview serve to exclude events that occurred during the prior reference period from the current one. Respondent mobility and noncompletions are so common, however, that intolerable data losses would incur were data to be used only from individuals who had been successfully interviewed 6 months earlier. In practice, bounding means merely that the unit was in the sample at the time of the prior interviews. Most other victimization surveys ask the respondent to report on an entire year or longer and do not employ a prior interview for temporal bounding.

The instrumentation employed in NCS is divided into a screening interview and an incident form. The interview presents cues to the respondents that are designed to help them recall and report possible criminal victimizations. Once a candidate event is mentioned, the respondents are asked detailed questions about the event to determine if it is a crime of interest to the survey and, if so, to provide information that can be used to classify the crime. All of the screening questions are administered before the incident form is administered. Some victimization surveys employ a screener/incident form logic, but many others do not. In those that do not, any positive response to a screening question would be considered a crime event. The type of crime event would be determined by the screening question that elicited the positive response. Moreover, the gathering of information about the incident occurs immediately after the respondent answers a screening question positively and before other screening questions are asked. The implications of these different approaches to screening will be discussed in detail later.

NCS asks only one respondent per household about thefts of certain kinds of things that are considered the common property of the household. All respondents are asked about thefts of their personal property. Specifically, these household respondents are asked screening questions about burglary, motor vehicle theft, and the theft of specific household property such as plants or lawn furniture. Because most other surveys have only one respondent per household, that person is asked about the theft of his or her personal property as well as the theft of common property of the household.

NCS uses a "series incident" procedure to accommodate victims who report a large number of victimizations and cannot report the details for each incident. Currently, if a respondent reports six or more incidents that are similar in kind and cannot provide the date and other details for each of the six events, then all of the events are treated as a "series incident." This means that the interviewer notes the number of events but completes an incident form only for the most recent one. Some victim surveys count all the events that a person mentions without concern for the ability of the respondent to recall the date of or other specifics about the event. Still other surveys record the number of events but administer the incident form only on a set number of events (e.g., a maximum of five). Limiting the number of incident forms is an attempt to reduce the burden on respondents and interviewers.

There are other ways in which NCS differs from other large-scale household surveys of victimization, but the features mentioned previously are among the most consequential from a cost and error standpoint. These particular design features were adopted because those implementing NCS believed that a survey designed in this manner would minimize the error in the estimates of the change in the level of crime. Some of the evidence relevant to these design decisions and the evolution of NCS to its current form are described in the next section.

The Evolution of NCS

NCS evolved into its current design in a series of stages. The first stage set the foundation for what followed.[5] In the early 1930s, the Wickersham Commission proposed a comprehensive national criminal justice statistics program under an independent central statistical agency. Although this plan did not achieve fruition, it led to making a cooperative national system for statistical reporting a Federal function under the FBI (National Commission on Law Observance and Enforcement 1931). Annual reports of a crime Index in UCR that the FBI compiled from these data became the most influential indicator for defining the seriousness of the Nation's crime problem.

Thirty years later, this achievement of criminology and statistics was increasingly being called into question by official and journalistic investigations of police offense statistics and by critical social science analyses. In the early 1960s, a few social scientists speculated about the possibility of adapting self-report national household surveys to produce an indicator of the nature of and changes in the crime problem that would be less vulnerable to the vagaries attributed to UCR.

In the second stage, these ideas received a receptive hearing from two new presidential commissions appointed in 1965 for the reform of law enforcement and the administration of justice (hereinafter referred to as the President's Commissions). During this period, the fundamental idea that citizen self-reports of crime could be used as the basis for crime statistics was formulated, proposed, and accepted by government officials and the public. In addition, many of the important methodological and logistical issues required to field a victim survey were addressed by several pilot studies.

In the third stage, the Census Bureau addressed many of these issues within the context of a large-scale household survey in preparation for fielding NCS on an ongoing basis. Some of the lessons from the earlier field tests were incorporated into the Census version, but others were not. A number of methodological tests were done during this period, shedding additional light on the effects of various design features on reporting in victim surveys.

The fourth stage began with the launching of the actual survey. After the National Criminal Justice Information and Statistics Service (the precursor to BJS) published the first of its annual reports, *Crime in the United States 1973* (U.S. Department of Justice, National Criminal Justice Information and Statistics Service 1976), the survey immediately achieved prominence as an indicator in the public media and in academic research and discussion. As a major social survey, it attracted the attention of leading experts and organizations in the social sciences and statistics as well as the U.S. Congress. Specific problems which the leading experts and the BJS census team identified led BJS to sponsor a reevaluation of the National Crime Survey by a committee of the National Academy of Sciences (NAS). Shortly after the publication of the results of the NAS study, congressional hearings were held on the possible suspension of NCS (U.S. House 1977).

The NAS evaluation resulted in a fifth stage, during which a 5-year program of research, instrument development, and redesign planning was conducted to deal with the issues raised by the report. During this period, a large number of methodological tests were conducted with particular emphasis on underreporting and screening issues.

The sixth stage of development began with the implementation of the changes in the design of NCS that were recommended as part of the redesign research. Again the Census Bureau engaged in extensive testing of various designs in preparation for implementing the new design.

In the remainder of this section, the last five of these six stages will be discussed in more detail by noting the advances made for designing and conducting victimization surveys.

Setting the stage in the 1960s

The confluence of several forces made the 1960s an auspicious time for the development of victim surveys. The brief period of détente in the Cold War moved defense-related issues off the front page. Demographic changes, both in terms of both the Baby Boom and the movement from rural to urban areas, moved crime to the forefront as a public issue. The waning of defense issues freed research and development professionals to seek other pursuits, and survey research enjoyed rapid growth. At the same time, UCR was coming under fire for not accurately reflecting how the crime problem was affecting society (Biderman 1966). These factors provided the skills, technology, and motivation for exploring the possibility of a victim survey.

The proposals for using interview surveys of samples of the general population, or polling methods, for measuring crime incidence rested on the belief that a vast reservoir of crime was not reflected in the statistics on offenses known to the police. It was recognized that many crimes were not reported to the police, and that officers at all levels of report processing could exercise great discretion in recording events. Reforms of several of the Nation's metropolitan police departments were accompanied by exposés of the previous practice of killing crime on the books. It was suspected that more reports would make their way through to published statistics when police departments believed crime was not being controlled properly because they were not allowed adequate resources or freedom of action.

For a sample survey to be practical and useful as a measure of levels and changes in rates of crime, two things had to be true (Biderman 1967):

1. The existing statistical indicators had to be found to be so inadequate and potentially misleading that it was worthwhile to develop and test an untried and expensive alternative.

2. The existing indicators had to be erring in the direction of massively understating crime rates. Were crime incidence not much higher than the official

statistics suggested, extravagantly large and expensive samples would be required to achieve sufficient numbers of incidents for statistically reliable results. In 1965, the total rate for all Index offenses combined for that year was a bit under 1,500 per 100,000 persons. If this was the true rate of crime, the expected number of robbery victims in a sample of 10,000 persons would be fewer than 10. The chances of encountering even one rape victim in such a sample would be quite remote. At that time, there was only one continuing national survey with a sample that large.

Research would be needed to demonstrate that both of these conditions were true. In the remainder of this section, we highlight the key milestones, both political and methodological, that led to the creation of NCS.

The President's Commissions' studies

The development of the crime victimization survey began in earnest with explorations for two commissions appointed in 1965 by President Lyndon B. Johnson: the President's Commission on Crime in the District of Columbia and the President's Commission on Law Enforcement and Administration of Justice. In cooperation with the President's Commissions, the first research grants by the Office of Law Enforcement Assistance, precursor of the National Institute of Justice, included the research and development of interview surveys to illuminate public experiences with crime and with justice agencies. A goal of both President's Commissions was to reduce the amount of crime that eluded the attention of the police. This was to be accomplished, in part, by increasing citizen cooperation with law enforcement (e.g., increase the amount of crime reported to police). The President's Commissions realized, however, that this goal had to be coupled with developing measures of the incidence and impact of crime that were independent of the efforts of the police. Otherwise, reliance on national or local statistics on offenses reported to the police or which otherwise became known to the police might be paradoxically affected by the President's Commissions' successes.

The victimization survey developed rapidly in its early stages. The idea of the survey was first broached in writing to the D.C. Crime Commission in September 1965 and to the President's Commission on Law Enforcement and Administration of Justice shortly thereafter. Independently of the President's Commissions, the National Opinion Research Center (NORC) incorporated victimization items in its ongoing omnibus amalgam survey in November 1965. The initial pilot survey, conducted by the Bureau of Social Science Research (BSSR) for the two President's Commissions in three Washington, D.C., precincts, began in January 1966. Field work for a supplementary BSSR precinct study that began in July 1966 was integrated with precinct studies by the University of

Michigan in Boston and Chicago using the same instrument and method. Interviewing for a national survey of 10,000 households by the NORC was conducted that same month. By the end of the year, all three organizations (NORC, BSSR, and the University of Michigan) had completed their exploratory studies. Their reports were published in three separate volumes by the President's Commission on Law Enforcement and Administration of Justice (Biderman et al. 1967; Ennis 1967; Reiss 1967b). The BSSR pilot survey is given prominence in the following discussion because it clearly shows a link between this early research and many of the methodological issues that have continued to emerge to the present day.

The BSSR pilot studies

The initial set of BSSR pilot studies identified many issues that persisted throughout the development of NCS. Many of these issues, especially with regard to screening and scope, remain controversial among researchers to this day. Much of this work was the result of stating the case for why and how such surveys might be done. Other valuable information came from collection and analysis of the data.

Lessons from conceptualization and planning. From the start of the planning process, two contrasting aims had to be reconciled. On one hand, there was the need to present results that could be compared directly with those from police statistics. This restricted much of the planning to the conceptual structure, definitions, and perspectives of police statistics, and to the UCR Index offense rates, in particular. On the other hand, it was important for the survey planners to incorporate in their instruments provisions for information on incidents and their victims that had not been collected before. In some cases, these two goals were in direct conflict.

The provisions of the surveys for comparing police and survey statistics had to permit adjusting survey *victimization* rates of individual persons (which UCR once claimed to be, but in many key respects was not) and infer from them *offense* rates for specific jurisdictions. This was no easy task because it required the designers of these surveys to stretch the methodology in a variety of ways, including:

- **Place.** Police statistics provide rates of occurrence within a jurisdiction, not for residents of that jurisdiction. What befell residents from the suburb while in the city had to be discounted for survey comparisons. Using a national survey for comparisons with national UCR rates is simpler, but not where comparisons are made for subnational places or types of places.

- **Residential mobility.** There was recognition that some means would be needed to deal with persons moving into and out of the areas under study. Although necessity might dictate the assumption that the premove victimizations of in-movers balanced out the postmove experience of out-movers, this assumption was tenuous.

- **Multiple and collective victims.** For offense classes where police statistics count only one incident even when there may be multiple victims, the survey needed provisions to identify events that someone else eligible for the survey might also give, if sampled.

- **Offenses against organizations.** As a sample of households and persons, the survey was an inappropriate vehicle for collecting information on crimes against businesses and other organizations. Separate exploratory surveys were undertaken of samples of businesses that are not under discussion here (Reiss 1967a, 1969; Aldrich and Reiss 1970). However, UCR does not consistently distinguish between residential and commercial crime. This made the comparison of UCR with the BSSR pilot data difficult. The decision was made to include robberies and other offenses against the person carried out against a respondent at a business or who is performing an organizational role, with the harm done to the individual distinguished from that done to the organization.

Although the requirement of making victimization estimates from the survey comparable to police offense rates constrained those designing the BSSR pilot survey, this survey was different from UCR in several important respects. First, the BSSR pilot survey would not attempt to validate crime reports in ways similar to police records. Theoretically, police reports are backed by the officer who fills out an incident report. The information from persons claiming to be witnesses or victims is subject to evaluation, and the report, in turn, is subject to evaluation at higher organizational levels and may be labeled unfounded on many grounds. The survey method, by contrast, places its ultimate trust in the unsupported testimony of the individual citizen respondent. It is assumed that the pledge of anonymity and the absence of material consequences, positive or negative, for the information given, should leave respondents with scant motive for deceit, invention, embroidery, reticence, or other departures from disinterested performance (Biderman and Reiss 1967). Rather, the survey exercises quality control by trying to identify miscomprehension or incorrect execution of the procedures.

A second difference lay in the scope of offenses covered. Provision was made for the interview to cover victimizations by a far more extensive range of offenses than the set making up the UCR Index. It included any acts of which the respondent was a victim and that the respondent thought was a crime in that it could be punished by imprisonment or fine. It included cues to a variety of frauds, forgeries, swindles, extortions, defamations, false accusations, and official misconduct as well as arson and vandalism. Proponents of direct UCR–NCS comparisons viewed this expansion to be counter to the goals of the survey. As will be noted in a later section, this feature of the BSSR pilot survey was greatly curtailed in later implementation by the Census Bureau and DOJ.

A third area expanded by the BSSR pilot survey was measuring the impact of crimes. The offense classifications used by UCR were highly constrained by its need to provide the least able cooperating departments with a set of categories and instructions for sorting and hand-tallying offenses in each category. The survey was not restricted by these categories. Planning for the survey could envision more refined discriminations within the traditional common-law categories and categorizations along other dimensions, as well. The survey instruments could explore the significance of victimizations from the victim's perspective, and they could cover many variates of relevant social values or policy issues.

Lessons from fielding and analysis. Once the interviews for the BSSR pilot survey began, a number of fundamental conclusions, both substantive and methodological, emerged. The first related to the salience of victimization events. The BSSR pilot survey found that most victimizations were not readily recalled by respondents, including victimizations that are classifiable as Index offenses or have high scores on the Wolfgang-Sellin seriousness scale (Biderman et al. 1967). Increments in the specificity of questions, prompts, and pauses for reflection brought forth large increments in the number of victimizations recalled. The first BSSR pilot survey questionnaire employed 70 discrete probes for victimizing incidents. Although these facts figure in the literature primarily for their methodological significance, their substantive significance for criminology is also important. That crimes are not highly salient events in memory implies that they do not rank high relative to many other life events in their importance for individuals. If we reflect upon how crowded lives can be with trials and tribulations of everyday life, even the most serious crimes are, apparently, paltry. The earliest report gave other reasons that so many victimizing incidents were apparently forgotten:

> Forgetting these events also stems from the unpleasant and embarrassing aspects of the experience. . . . Further, few of the incidents led to a path of action that might serve to reinforce the ability to recall the event. The large majority of the . . . [events] are happenings that would have been difficult to avoid—measures to prevent repetition . . . would usually involve greater cost than . . . the risk deserves. In very few of them is the victim known; hence there is no individual target on whom the victim can fix

> whatever affect the event may arouse. In most instances, there is nothing to do to gain either material or emotional indemnification for the loss. (Biderman 1966, 12)

The final report of the BSSR pilot study went on to explain why the low recall salience of crime incidents does not mean that they were unimportant events for the victims. Their importance, it was argued, resides in their being indicative of the fragility of the social order, and these experiences are assimilated to and may be outweighed by other signs of disorder.

A second important finding was that the incidence of victimization was far more frequent than existing statistics suggested. The feasibility of a national survey was asserted in a progress report 3 months after the pilot project began (Biderman 1966; also reproduced as appendix G in Biderman et al. 1967). The report was based on the high percentage of respondents giving victimization reports in pretest interviews and in the earliest interviews of the survey proper (only 183 interviews in all).

A third important substantive finding related to the great excess of the survey rates over those reported to the police. To compare police and BSSR pilot survey data, a procedure was applied for reconciling survey offense rates with those of police reports for the same precincts. Even after eliminating from the calculations those incidents that respondents said had not been reported or were not otherwise known to the police, the survey rates were far higher. The conclusion was that nonreporting *by* the police may account for more of the dark figure than nonreporting *to* the police.

These initial trials identified problems of interviewing for victimization that have continued to receive methodological attention to the present day:

1. A recency bias in recall so pronounced that a reference of period of no more than 6 months was recommended for future surveys.

2. The need for singularly focused incident recall tasks.

3. Far greater victimization reporting by self-respondents than by household members acting as proxy respondents for other members.

4. The problem of crimes against the household and of multiple-victim incidents.

Integrating results across surveys

The foregoing results of the BSSR pilot studies were both consistent with and contrary to the NORC pilots. By comparing and contrasting across these surveys, several key findings emerged. The first was the revelation that the incidence of crime generated by victim surveys was sensitive to survey procedures. The Washington, D.C., study (Biderman et al. 1967) and the NORC survey (Ennis 1967) used different methodologies. The Washington, D.C., field test was organized around principles that would facilitate recall and reporting of crime events. In practice, this meant minimizing cognitive burdens that occur when the interview imposes official rules, terms, and definitions that hamper straightforward internal and conversational discourse. The procedures avoided complicating the respondents' memory work with filtering, composition, and decomposition tasks to make their thoughts and answers fit official categorical molds.

This contrasted with the tack taken by the NORC questionnaires. Those questionnaires used a battery of screening questions, each one devoted to a specific Uniform Crime Reports crime class and containing all the elements needed to define a victimizing event as belonging to that class. Screening questions were worded to exclude experiences that did not fit the official definition of the crime class to which the item was devoted. They included wording that sought to ensure that the item encompassed all the experience fitting the criteria for the class. By having the respondent answer positively to only one screening question for any incident, analysts could make its preliminary victimization counts by crime class simply by tallying "yes" answers to screening questions. The NORC survey then followed the next step of police statistics: further incident interviewing to inform an "unfounding" procedure for eliminating questionable reports. This included providing interviewer ratings of the veracity of the respondent's testimony and then a review by experts of a subsample of incidents, including police and lawyer raters of the incident report for inconsistencies and appropriate classification.

The NORC approach made for long screening questions, as illustrated by this one for robbery:

> Within the last twelve months, did anyone actually take or try to take by force or threat of force from you personally or anyone in the household any money or property? This would include bicycles forcibly taken away from children, or a violent purse snatching. (Ennis 1967, appendix A, 3)

As previously noted, the BSSR Washington, D.C., pilot survey, by contrast, proceeded by orienting the respondent to the crime victimization recall task and then presented the respondent with a long list of short cues, largely of between one and five words, giving the respondent time to think between each one. The screening questions were not to be used as data (other than for methodological analyses), but simply triggered the execution of an incident form. The detailed incident questioning had the burden of getting the information needed to

The pilot surveys answered the basic questions about the need for and the feasibility of a survey-based indicator of crime. Moreover, they identified (and informed) many of the basic design issues in creating such surveys.

determine what offense(s), if any, had occurred in the incident(s) the respondent recounted, who the victim(s) was (were), and additional information about the incident and its aftermath.

The BSSR pilot survey procedure yielded far higher annual victimization rates (0.80/respondent) than that yielded for central cities by NORC (0.08). After taking account of what was learned in the initial Washington, D.C., pilot work, the BSSR instrument was modified in collaboration with the University of Michigan's Metropolitan Areas surveys for the President's Commission on Law Enforcement and Administration of Justice. The revised BSSR interview procedure yielded approximately 2.0 incidents per respondent (Biderman et al. 1967, 50). The greater productivity of the BSSR/Michigan survey suggested that one's approach to screening will dramatically affect the resulting incidence estimates. More specifically, it suggested that organizing surveys in a manner consistent with the principle of facilitating the recall and reporting tasks was preferable to emphasizing legal principles which complicated the respondent's task.

Another indication of the dependence of the rates yielded by the survey on method was the positive correlation between education level and victimization by the types of crimes where it might not be expected. This suggested better performance as an interview subject of the better educated. Biderman (1967) wrote more generally of "class-linked under-reporting" in the survey.

A second key finding found across all the pilots was a severe recency bias in the data. This was observed by increased reports of victimizations at the earliest and most recent ends of the reference periods. Increased reports at the beginning of the period were thought to reflect incidents occurring outside the period being brought forward in time into the period. The increase at the end of the period was seen as a mix of telescoping and greater recall of events that are closer to the interview. These phenomena, identified earlier in a Census Bureau experimental housing survey (1965), were regarded as applicable here and figured in much of the future design research on NCS.

Summary of the pilot studies

The pilot surveys answered the basic questions about the need for and the feasibility of a survey-based indicator of crime. Moreover, they identified (and informed) many of the basic design issues in creating such surveys. These studies confirmed suspicions that police data substantially underestimated the level of crime because both citizens failed to report and the police did not record eligible events. They also showed that there was a vast reservoir of crime that could be estimated with a household survey. Grappling with actually fielding such a survey identified the design issues that needed to be addressed. Principal among the lessons learned here was the inherent tension between the logic of police record systems, particularly UCR, and the logic of surveys. The constraints of the former would prohibit realizing the full potential of the latter. Asking about crimes using the legalistic framework of police record systems would probably inhibit complete reporting of events in the survey. Moreover, constraining the scope of crimes and the information collected about crime events to that which is customarily included in police record systems would fail to exploit the potential of these surveys.

Other valuable lessons were learned. First, respondents had trouble recalling and reporting events, so steps should be taken to facilitate the task. The recall task should be simplified and many cues should be provided to jog memories. Second, temporal placement of events within and outside of the reference period by respondents was problematic, so some attention should be given to making this easier. Third, self-respondents were preferable to proxy respondents. Fourth, some attention needed to be given to the problem of reporting the theft of collective property. Asking everyone in the household about these items would result in some duplicate reporting, but asking less than everyone would result in underreporting.

Implementation of the survey by BJS and the Census Bureau

It was clear from the pilot studies that large samples would be required to obtain reliable estimates of victimization for crime classes of intense interest (e.g., rape).[6] The Census Bureau was the only organization that could field such a large survey and was chosen to conduct the ongoing NCS. In preparation for implementing the survey within the Census environment, some of the lessons from the field surveys were taken into account although others were ignored. In addition, the Census underwent an extensive program of pretests, trial surveys, and record check experiments beginning in 1970 (Lehnen and Skogan 1981).

Several important features of the current NCS design resulted from this work. One experiment in Dayton and San Jose (Kalish 1974) assessed the effectiveness of proxy reporting for the survey. This continued the line of work reported by the President's Commissions that proxy respondents were far less productive than self-respondents. The conclusion from the pilot studies, not surprisingly, was that when a single respondent reported for the entire household, far fewer crimes were reported when compared with interviewing all members of the household as self-respondents. As a result, the idea of interviewing all members of the household was eventually adopted for NCS.

A series of reverse record check studies was also conducted in different cities. These surveys drew a sample of known victims from police records and then interviewed these individuals to see if the incident was reported on the survey. Theoretically, this type of study provides an external criterion to judge the accuracy of reports on the survey. These studies were used primarily to determine the accuracy of the recall of incidents by respondents. The Census researchers drew two main conclusions about the optimal length of the recall period:

> If the objective is to determine whether a crime occurred, as opposed to placing it in a more accurate timeframe, then a 12-month reference period is as good as one of 6 months. . . . To the extent that it is desirable to place an incident in a specific timeframe, greater accuracy is obtained from a shorter reference period. Thus, a 6-month reference period is better than 12, and a 3-month period is better than 6. (Dodge and Turner 1981, 3)

These conclusions were used as a basis for a 12-month reference period for surveys done across cities (Hindelang 1976). However, this basic result was not accepted by a number of researchers (e.g., Biderman 1981a; Biderman and Lynch 1981, 31), partly due to the problems associated with a reverse record check design (see following text for problems). A 6-month reference period was eventually adopted for NCS.[7]

A second basic result from the reverse record check studies served as a precursor of issues that still haunt victim surveys. This result was the conclusion that

> [recall] was very high for crimes involving theft of property (80 to 85 percent). With respect to personal crimes, robbery was well reported (75 percent and above), but rape and assault were less so ($66^2/_3$ percent and 50 percent, respectively). An important factor in the recall rates for cases of personal victimization is the relationship of the offender and victim. Recall rates vary directly with the nature of that relationship; that is, when victim and offender are strangers, recall rates are high. . . . Acquaintance, and even more kinship, results in lower reporting rates. (Dodge and Turner 1981, 3)

As will be noted, one of the primary faults found with victim surveys has been their inability to illuminate violence among persons that know one another.

A third feature of NCS was also adopted from these experiments: the use of a household respondent to report about crimes against household property. This was based on the conclusion that a single household member could report on crimes such as burglary, auto theft, and larcenies against household property (e.g., lawn furniture, plants). This resulted in arranging NCS screening so that a single person (the household respondent) is administered a set of screen items that specifically ask about these types of crimes. Once this part of the screen is complete, all household members are administered a set of questions that are meant to apply to personal crimes.

It should be noted that the final design of NCS was, in several ways, contrary to the recommendations initially made by the President's Commission on Law Enforcement and Administration of Justice based on the results of the field surveys. Specifically, the principle of facilitating recall and reporting was compromised somewhat in favor of some of the legal principles and the desire to classify crimes neatly (Dodge and Turner 1981, 4). The major impetus behind this was the attempt to mimic UCR. The Census survey restricted its screening to Part I crimes in UCR, such that questions were asked with the intent of eliciting mentions of these crimes and only these crimes. Although the Census instrumentation separated the screening task from the provision of detailed information for classification, there was a one-to-one correspondence between the screen questions and the UCR crimes. Related to this was that the NCS questionnaire departed from the "short cue" approach adopted in the BSSR/Michigan pilot studies in favor of a more rigid approach that attempted to direct attention to legal categories. Evidence from the pilot studies, as well as evidence that has been cumulated since (see the following discussion of the NCS redesign), suggest that all of these departures reduced the rate of reporting in NCS.

Other design features of NCS were occasioned by the need to fit into the organization of the Census Bureau and the Current Population Survey (CPS). CPS is the largest intercensal survey conducted in the world and, at the time, NCS was to be the second largest of these surveys. Sharing interviewers between the two surveys would mean great efficiencies for the organization. CPS employed a rotating panel design. This was viewed as an advantage to NCS for a number of reasons. One was the ability to use prior interviews to "bound" subsequent interviews (Neter and Waksberg 1964). A second was that the rotating panel design substantially increased the precision of the year-to-year change estimates. The panel design feature produces a natural positive correlation across annual estimates. This, in turn, substantially reduces the standard error on change estimates.

In addition to these decisions regarding the design of NCS, the Census Bureau also instituted a survey of commercial establishments and a set of cross-sectional surveys conducted in a number of the largest U.S. cities. The commercial surveys were developed because exploratory studies of small business showed that these establishments had victimization rates several times that of households (Reiss 1967a, 1969; Aldrich and Reiss 1970). Moreover, the household survey was not a good vehicle for measuring this component of the crime problem. The city surveys were fielded in an effort to evaluate the impact of crime prevention and crime reduction programs implemented with DOJ funding in the largest cities. These surveys were intended to assess the change in the level and distribution of crime in these cities as a result of the programs.

The National Academy of Sciences report

When NCS began to produce information on crime, various groups began to question the quality and usefulness of these data. Groups supportive of police-based crime statistics were already suspicious of this new data collection system. Academics began to raise questions about a multimillion-dollar data collection with few variables that could be used in testing theories of crime and that could not produce estimates for local jurisdictions. They also worried that this new data collection would take funds away from criminological research. The sponsoring agency also began to wonder about its new creation when the first years of data began to show the same large increases in crime as UCR (Parkinson, Paez, and Howard 1977). While much of the concern was focused on the commercial and city surveys, not NCS, all aspects of this new data collection came under scrutiny.

In response to these concerns, the Law Enforcement Assistance Administration asked the Committee on Social Statistics of the National Academy of Sciences–National Research Council (NRC) to evaluate the surveys. The committee selected a panel that represented a variety of disciplines and recruited staff to carry out the investigations necessary to perform the work. The study took place between January 1974 and June 1976. The panel examined every aspect of NCS, from the goals of the survey to its staffing and management to the publications produced with NCS data. The panel's recommendations and deliberations were published in *Surveying Crime* (Penick and Owens 1976).

Many of the panel's recommendations were pertinent to the management of the survey within the Census Bureau and DOJ, and others sought the elimination of the commercial and city surveys. Among the recommendations that addressed survey design, procedures, and instrumentation were some familiar calls for improvements in screening procedures. The panel suggested that:

> [1] The function of screen questions should be to facilitate . . . recall and reporting of happenings that fall within the scope of the survey. The

> usefulness as data of screening responses themselves is restricted to their use in the analysis of the effectiveness of screen questions. The operating rule is that data for final tabulations come from the detailed exploration of the pertinent events by the use of subsidiary incident forms. (Penick and Owens 1976, 82)
>
> [2] A set of well defined screen questions must recognize that there are not just seven crimes in which one is interested. . . . The screener can be a cue to any. . . . The screen questions can also be a cue to some element of place or some other circumstance of the victimization that may help bring about mention of an incident. . . . (Penick and Owens 1976, 84)
>
> [3] Screening procedures should reflect less worry about redundancy and about eliminating ineligible events than about unnecessarily cluttering up discrete questions and the respondent's thinking. (Penick and Owens 1976, 87)
>
> [4] The household informant should be limited to questions on breaking and entering and to household property items. . . . Alternatively for research purposes, everyone who is interviewed within the household should be asked about household as well as personal, crimes. Interviewers . . . should assume the burden of eliminating separate mentions . . . of identical incidents, including the theft of jointly owned property such as automobiles. (Penick and Owens 1976, 87)

The panel also recommended that the "screen questions take account of the large volume of incidents now classed as series" (Penick and Owens 1976, 88). The urgency of this issue became apparent with the availability of the first years of data from the survey, in which "an estimated 20–30 percent of reported personal victimizations were treated as a series and excluded from the personal victimization count." The panel took exception to the fact that the series incident procedure (1) excluded a large number of relevant events simply because they did not conform easily to the incident logic of the survey, (2) required the respondent to make difficult judgments about combining a set of incidents into a series and to estimate for each series the number of events involved, and (3) allowed the determination of whether the series procedure should be invoked to be made by the interviewer in the field rather than by data analysts.

In addition to these recommendations pertaining to screening, the panel called for research and development work on the best combination of reference period, frequency of interview, retention time of an address in sample, and bounding rules that would address the following questions:

> [1] Should the reference period be 3 months, 6 months, or 12 months?

> [2] Should a household address be interviewed once, twice, three times, seven times, or some other number?
>
> [3] Is the bounding interview worth its cost and does it introduce a new significant bias into the results?
>
> [4] What are the shapes of the reporting decay interviewer overload and telescoping functions? (Penick and Owens 1976, 68)

Surveying Crime and the work of the NAS–NRC panel raised questions about some of the design decisions made during the implementation of NCS within the Census Bureau. These questions would soon become the agenda for a program of research and development that would shed further light on the relative desirability of different approaches to surveying victims of crime.

The NCS redesign and other improvements

Since the publication of the NAS report, a number of studies have been completed that have informed the design of NCS as well as the conduct of victimization surveys more generally. In response to the NAS recommendations, BJS sponsored a long-term redesign of NCS by convening a consortium of Government, private, and academic experts in various fields relevant to the design of NCS (e.g., statisticians, criminologists, survey researchers). The result was to implement significant changes to the survey, the most drastic coming in 1992.[8] Research related to the design of victimization surveys more generally was undertaken by other researchers interested in improving the method. In this section, we briefly review the results of this research.

Reference period

As noted in the NAS report, further work was needed to assess the optimum reference period for NCS. The reverse record check studies completed in the early 1970s were a first step in this process, but were not viewed as definitive. A reference period experiment (RPE) was conducted by the U.S. Census Bureau and sponsored by BJS in the 1978–80 period. In the experiment, portions of the NCS sample were randomly assigned to interviews with 3-month, 6-month, or 12-month periods. Analysis of these data found that aggregate level estimates increased substantially as the reference period was shortened (Bushery 1981). The 3-month period produced significantly higher rates than the 6-month period, which produced significantly higher rates than the 12-month period. This finding runs counter to the conclusions of Dodge and Turner (1981) that the production of incident reports did not vary by length of the reference period.

RPE also indicated that the 3-month period displayed significantly different relationships of victimization with key sociodemographic variables when compared with the 6-month period. In particular, the effect of age was found to be stronger for all personal crimes in the shorter period, and the relationship to race was stronger for serious assaults and robbery (Kobilarcik et al. 1983; Cantor 1985).

Czaja and Blair (1990) conducted an experiment that compared reference periods of 6 and 9 months for three different types of crimes (burglary, robbery, and assault). They found significant underreporting of all three types of crimes, with burglary (16 percent) and robbery (28 percent) having significantly lower underreporting than assault (71 percent). The authors attribute the significantly higher rate of underreporting for assault to conceptual issues related to whether the victim defined the event as a crime. They also found that underreporting varied by race. Nonwhites were significantly more likely not to report the crime than whites.

Czaja and Blair (1994) did not find reference period length to have a significant effect. Respondents seemed about equally able to report crimes across the two different reference periods. This is consistent with initial analysis of the early reverse record check studies previously discussed (Dodge and Turner 1981), but it is inconsistent with RPE. It seems likely that the differences between RPE and the reverse record check studies is due, at least in part, to the types of crimes that were investigated across the two studies. RPE asked about all types of crimes, while the reverse record check studies examined events that were reported to the police. The latter are most likely to be remembered by respondents, so the difference of 6 to 9 or 12 months in the reference period may not be critical for reporting these types of crimes. It may be more important for crimes that are never reported to the police.

Using the results of RPE, several individuals developed formal statistical models that quantified the error properties of designs with different length reference periods (Lepkowski 1981; Bushery 1981). These analyses are based on the assumption that any increase in the rate of reporting victimization is better. Under this assumption, these analyses concluded that using a 3-month reference period was the best alternative among the three tested in the experiment. However, the assumption that more is better has been questioned in a number of contexts (e.g., Skogan 1981, 12). Increased reporting rates may occur, for example, if respondents telescope more crimes into the reference period.[9] Ultimately, NCS did not shorten the reference period, in part because of a fear that such a change would have a serious impact on the statistical power of key comparisons.[10] Nonetheless, this line of research has led designers of victimization surveys to be cautious when trying to extend reference periods beyond 12 months and to prefer shorter reference periods whenever possible.

A second change considered by the redesign was to simplify the reference period for NCS respondents. NCS imposed both an early and recent boundary on the period. Respondents were asked to report for the 6-month period ending at the end of the previous month and beginning 6 months prior. For example, if interviews were conducted in August, respondents were asked to report for the time period between the end of July and the beginning of January. As previously described, this ran counter to the initial pilot designs, which specifically deemphasized all of the reference period boundaries. The idea was to emphasize the recall of any eligible incidents without placing many filters on the cognitive task of the respondent. The use of any specific boundaries ran counter to this basic premise. The use of two boundaries, as in NCS, further complicated the recall process.

The redesign of NCS found that instituting the most recent bound resulted in substantial telescoping of crimes from the month of interview to the last month of the reference period (Biderman et al. 1986, 80). For this reason, starting in 1992, NCS asked respondents to report victimizations up to the day of the interview.

Improvements in screening

As previously noted, the early pilot studies indicated that respondents need cues and examples to help them recall incidents of victimization (Biderman et al. 1967). Specific screen cues and questions serve both to orient respondents to the types of events covered by the survey and to jog their memories for incidents that do not immediately come to mind as instances of crime (Biderman et al. 1986, 88–103).

In addition to jogging their memories, it is desirable to reduce any inhibition respondents might have to report crimes that might be sensitive. Victims may be reluctant to report incidents that are a source of pain, fear, shame, or embarrassment. One way of coping with a painful experience, in fact, is to try to forget it. Reporting the incident in a survey forces the victim to reexperience it and, perhaps, disclose information that could become known to other household members. Of particular concern is the gross underreporting of domestic violence on household victimization surveys.

Procedures for conducting NCS were not set up to promote disclosure of incidents among household members. For example, the Census Bureau does not treat the guarantee of confidentiality as applying to other people who are present during an interview; as a result, many NCS interviews are not conducted in private. There is evidence that this does, in fact, inhibit reporting of violence. In the case of domestic violence, the offender may actually be present (Coker and Stasny 1996).[11]

Experimental tests conducted as part of the redesign tested screening strategies using an enhanced cuing approach. It was found that this approach significantly increased the number of reports of all types of crimes relative to the screener used on NCS (Biderman et al. 1986, 104–165). Increases were thought to be due to widening the concepts respondents have about eligible events as well as facilitating retrieval from memory. In the late 1980s, the Census Bureau went on to test a revised version of this screening strategy in several field experiments. The tests uniformly showed increased reporting of all types of crimes except robbery and motor vehicle theft (Hubble and Wilder 1988; Kindermann, Lynch, and Cantor 1997). The strategy was eventually adopted by NCS. Particularly large increases occurred for crimes that were thought to be the most underreported, such as sexual assault (especially among nonstrangers) and simple assaults (many of which are attempts without completion).

The introduction of a computer into the survey process changed not only the way interviews were administered but also the way survey organizations were managed.

Computer-assisted interviews

The introduction of a computer into the survey process changed not only the way interviews were administered but also the way survey organizations were managed. These changes have had a dramatic effect on the quality of the information that is collected on surveys in general (Couper et al. 1998) and NCS in particular.

The past 15 years have seen the universal adoption of computer-assisted telephone interviews (CATI). This has generally been seen to have had a positive effect on data quality for three reasons. First, it allows for programming more complex skip patterns. This takes the burden of navigation off of the interviewer and allows her to concentrate on the respondent (at least theoretically). Second, the computer forces the interviewer to at least see and review all questions for all respondents. When administered by paper and pencil, interviewers have more control over what questions they will and will not administer to the respondent. If they view certain questions as burdensome or feel a respondent may not react well to them, they can easily skip over them. When a computer is used, they are at least forced to view the screen before passing through. This could be especially important in conjunction with the detailed screening strategies described previously. With the increased cuing, the screener is longer and could be viewed as especially burdensome on the respondent. There might be more of a tendency to skip parts of the screener if the interviewer does not believe the questions are worth asking.

A third advantage to CATI is that administration is centralized. Rather than interviewers working out of their homes, they work in a central facility. This allows much tighter quality control over their work. Interviews are routinely monitored by supervisors. This makes it much more difficult for interviewers to deviate from accepted protocols, not to mention fabricating data.

As part of the redesign of NCS, the Census Bureau conducted a series of split-ballot experiments investigating the use of CATI to conduct interviews. Cases were randomly assigned to be interviewed either by CATI or by an interviewer out of his or her home by telephone (the traditional method). The results indicated a substantial increase in the reporting of all types of crimes in the CATI condition (Hubble and Wilder 1988). It is not clear whether the increase was due to computerization of the instrument, the centralized monitoring, or both. Nonetheless, as the new methods of NCS were implemented in 1992, a significant proportion of the jump in the reported victimization rate was attributed to this aspect of the redesign.

Within the past 3 to 5 years, survey researchers have developed methods for the respondent to complete a survey using a computerized self-administered procedure. The primary motivation behind this has been to reduce response inhibition and distortion. If respondents do not have to report sensitive information to an interviewer, they are more likely to report socially sensitive incidents. Computer-assisted self interviews (CASI) were first developed for use in self-report drug surveys and have been applied to a wide range of sensitive behaviors (e.g., same-sex sexual activity, abortion). An enhancement of this method has been to add an audio component, audio computer-assisted self interviews (ACASI), which reads the questions to the respondent. Respondents wear a set of headphones while following the questions on the computer screen. The audio component assists in overcoming possible literacy problems as well as enhances the privacy of the interview (e.g., respondents are free to blank out the screen and use only the audio).

Experimental research has found that ACASI leads to more reports of sensitive information when compared with an interviewer-administered instrument (Tourangeau and Smith 1996; Turner, Ku, and Sonenstein 1996). Respondents seem to feel more comfortable reporting sensitive behaviors when interviewers or other observers are not involved.[12]

Computerizing a self-administered instrument is particularly convenient for victimization surveys, given the relatively complex skip patterns and questionnaire structures. The need to first administer a screener and then follow up each incident mentioned with detailed questions (e.g., What happened? When did it happen? Who did it? Where did it happen?) makes it extremely difficult, if not impossible, to use a paper and pencil self-administered form. The skip instructions are simply too difficult to communicate and implement. A computer takes care of this problem without complicating the respondent's task.

ACASI has not been widely implemented on victimization surveys. The exception is the British Crime Survey (Mayhew 1995; Percy and Mayhew 1997), which uses it to administer questions on sexual assault and domestic violence. Results, although not experimental, indicate a large increase in the reporting of these incidents. Application of this new methodology is likely to spread as it becomes more available.

Revising the series incident procedure

Dodge and Balog (1987) examined series incidents in NCS largely for the purpose of determining if classification of events as series incidents was due to interviewers' unwillingness to collect data on a large number of incidents from a given respondent. To do this, they first identified respondents who had initially reported series incidents and then reinterviewed these individuals using two separate surveys modified from NCS. Dodge and Balog found that in most series incidents with five or fewer events, the respondent could give details of the event if the interviewer asked for those details. This was not the case for the majority of the series incidents with six or more events. On the basis of this study, the Census Bureau changed the requirements for invoking the series incident procedure from situations in which three or more events were reported to those in which six or more events were reported. Moreover, the other information necessary for using the series procedure was explicitly built into the interview. Interviewers were required to ask or verify if (1) the events were similar and (2) the respondent could not report the details of each event. These changes followed closely the recommendations of the NAS panel and substantially reduced the number of events treated as a series incident in NCS. They also help reduce the effects of interviewer discretion on the identification of high-volume victimization. We can be more confident that events treated as series incidents are different from those events that are more discreet and distinguishable for the respondent. This moves us closer to being able to treat series incidents as a distinct form of victimization to be investigated rather than as measurement error.

The Census Bureau also introduced questions about the interrelationship of events in a series incident. They asked if the events involved the same offender or different offenders, if they occurred in the same place or in different places, and if the victimization had stopped or was continuing at the time of the interview. These few questions add a great deal of information about series incidents. It provides an idea of whether the victimization is a repeated

The magnitude of the effects of differing design features on victimization rates is so striking as to raise serious questions about the implications procedural variations may have for uses made of the survey data.

encounter with the same individuals or a much less particular event. This, in turn, provides valuable insight into the genesis of this continuous event (Lynch, Berbaum, and Planty 1998).

These changes in the series procedure are noteworthy because they mark a break with the point-in-time assumptions of NCS regarding crime events. In some cases, it may be more appropriate to consider crime as part of an ongoing event or condition such that one event can be a precipitant of another. This was not known until the survey collected information on the relationship between crime events. This is a small break with the emphasis on incident rate estimation and the assumption that crimes are best viewed as point-in-time events rather than events of continuous duration.

Limitations and Future Research

The institutionalization of victim surveys has encouraged their use in many debates of controversial policies. For example, victim survey data figured prominently in the debates about the Campus Crime Act, the Violence Against Women Act (Gilbert 1992; McPhail 1995; Murray, n.d.), and gun use (Kleck 1991, 1996; Cook 1985). This intense use of victim surveys has identified a number of longstanding methodological issues as well as raised new ones (see Fisher and Cullen in this volume). In addition, a number of longstanding issues were not resolved in the research and development work of the past two decades; they, too, should be addressed.

Controversies with the design and analysis of victimization surveys

Researchers' opinions differ on the importance survey design has for interpretation of victimization surveys. The problem of understanding the implications of survey methodology for analysis is not unique to victimization surveys. However, the magnitude of the effects of differing design features on victimization rates is so striking as to raise serious questions about the implications procedural variations may have for uses made of the survey data.

The differences between the BSSR and NORC surveys, the differences in NCS before and after 1992, and more recent surveys on violence show that self-reports of victimization vary by a factor of 2 or greater, depending on the design features of the survey. Not only is this variation extremely large, but some of it is related to characteristics of the respondent or the event itself. An important controversy, and area of research, centers on the implications this variation has for conducting research with victimization surveys.

Several examples illustrate this point more clearly:

- NCS incurs great expense to use a bounded 6-month reference period because of extensive evidence of improvements in data quality that bounded, brief periods provide superior data. The original city surveys (Hindelang 1976; Hindelang, Gottfredson, and Garofalo 1978), the British Crime Survey, and most of the surveys on violence against women use an unbounded 12-month period. What are the implications of this for interpretation of analysis of these surveys? For example, many of the analyses of multiple victimization (see our earlier discussion) were based on unbounded reference periods. Does this increase estimates of multiple victimization artificially?

- As noted previously, when NCS changed methods in 1992, the level of crime jumped by 50 percent to 200 percent, depending on the type of crime. This is generally attributed to the change in screening methods and the use of CATI (Persley 1995). If one assumes the postredesign data are better, does this argue against ever using preredesign data?

- Screening, the context of the questions, and automation (CATI and ACASI) have been shown to increase reports of sexual assault by factors of at least 2, depending on the domain of interest. Does this invalidate surveys that use methods where underreporting is the greatest?

The use of different methods both within and across surveys is unavoidable, given the costs associated with data collection. In many cases, the design feature that is considered better is also more expensive to implement than alternatives. This includes, for example, using shorter recall periods, using bounded recall periods and self-reports rather than proxy reports. Because of the expense of these "best" design features and the need to survey fairly large samples of people to yield statistically reliable analyses, only a survey of the magnitude of NCS can hope to institute many of these procedures. Even NCS, however, treats respondents nonuniformly (see the following discussion). Consequently, it is of both scientific and practical interest to understand what sacrifices survey planners and users may make in adopting particular designs or their products.

When judging alternative designs, it is important to keep in mind two basic analytic goals. One is to estimate the actual level of particular crimes. This, for example, has been the main controversy surrounding the surveys focusing on violence against women. A second goal is to look at relationships among different variables to evaluate particular policies or examine year-to-year change.

When estimating the level of crime, it is readily evident that variations in methods provide vastly different answers to questions. Which method or sets of methods provides the best estimate of the particular concept of interest? We discuss in more detail in the following sections issues of data validity along with the research needed to clarify key issues related to understanding these wide variations for purposes of estimating the level of crime.

For more elaborate analyses (e.g., analysis of change over time, relationships among variables, and evaluation of policies), judging validity is more complicated. The critical question is not only which methods increase the validity of level estimates, but also whether different methods produce different substantive conclusions. A particular data collection method may be better at estimating the true level of crime (at least as defined by a particular study), but if measurement error is uncorrelated with the domain of interest, substantive conclusions may be unaffected. For example, one might argue that the use of longer recall periods is legitimate if measurement error is not correlated with critical relationships that may be of interest. If the primary relationship of interest, say, is between the victim's race and rate of victimization and the underreporting associated with longer reference periods does not vary by race, then analysis may be unaffected by the use of a longer period.[13]

We know that measurement error is correlated with a number of important characteristics related to victimization reporting. Studies have shown, for example, that race is related to underreporting (Czaja and Blair 1990), differential error by race is associated with the length of the recall period (Kobilarcik et al. 1983), and blacks underreport simple assaults (e.g., Skogan 1981, 30–31). Respondent event dating and definitional problems are correlated with the saliency of the event, at least as indicated by reverse record check studies.

Little research has been done on how these types of relationships vary by design feature. Are the underreporting patterns by race and/or education different when the screener is modified to encourage more complete reporting? When the reference period is shorter? When the instrument is self-administered? Answers to questions like these would enhance both the design and interpretation of results across surveys.

The use of differential methods also exist within particular surveys. It may be more convenient, for example, to conduct interviews by telephone and conduct in-person interviews for those persons who do not have telephones. Similarly, it may be more convenient to interview particular household members using a proxy interview than to use self-reports. These treatment nonuniformities may be correlated with measurement error and the domain of interest. In this regard, NCS is, perhaps, the most egregious culprit. For example:

- The panel design of NCS does not allow for an initial bounding interview for persons who move into a sampled unit after its panel's first time in the sample (i.e., at waves 2–7). Thus, data from persons who move into a housing unit are combined with data from those who had been living in the house in the previous interview. As noted earlier, unbounded data produce more reports of victimization and those most likely to move have higher victimization rates (Biderman and Cantor 1984). It follows that NCS will overestimate the relationship between mobility (and its correlates) with victimization. In addition, as the percentage of unbounded households changes from year to year, there is potential that the yearly change estimates may also be affected (Biderman and Lynch 1991).

- A single member of the household is administered screening questions devoted to crimes against the household. This screener, however, reveals more crimes against individuals as well (Biderman, Cantor, and Reiss 1985). The selection of household informants is negatively correlated with victim risk (the household member who tends to stay home is most likely to be selected as the household informant). This depresses relationships associated with risk.

- The use of CATI on NCS is restricted to those who have telephones and who are willing to participate using this mode. Since CATI increases reporting of victimizations (Hubble and Wilder 1988) and its use is negatively correlated with risk, relationships examining risk factors are depressed.

Future research might further explore differences across methods as they impact substantive relationships. Research along these lines can be done in several ways. The most elaborate, and expensive, is through experiments, much like those described in the development of NCS (e.g., Kobilarcik et al. 1983) and by the work of Czaja and colleagues (1990, 1994). Treatments can be randomly assigned across respondents and results compared across treatments.

A second line of research would be to conduct identical analyses across data sets that vary systematically by design features. There has been little detailed comparison across datasets for key analyses (e.g., NCS versus BCS). Such research might illuminate how the different designs affect key relationships. Surveys, unfortunately, do not typically differ by only one feature. Consequently, these comparisons could not assign definitively the reason for any observed difference. They might, however, suggest the magnitude of the effects of particular combinations of design features.

The suggestion of more methodological research is complicated by the absence of good criteria for assessing the validity of the resulting data.

Validation

The suggestion of more methodological research is complicated by the absence of good criteria for assessing the validity of the resulting data. Reverse record check studies using police reports have been shown to be flawed conceptually (Biderman and Lynch 1981). One concern is that police records cover only events that are, by definition, not in the "gray area" the survey is meant to cover (i.e., crimes that do not come to the attention of the criminal justice system). This problem is illustrated when comparing the reverse record check studies (Kalish 1974; Czaja and Blair 1990) and the reference period experiment (Bushery 1981). They came to different conclusions partly because the former covers police events, whereas the latter does not. A second criticism of the reverse record check methodology is the difficulty of matching across the two mediums. Information in police records about the event may not be reported by the victim. Consequently, determining whether reports by victims match a police report are difficult to determine. In one study, for example, Miller and Groves (1985) demonstrated that the conclusion is influenced by what matching rules are applied.

Comparative studies of different survey procedures are useful alternatives to external data for validation purposes, but require assumptions in order to say something about validity. The most common assumption has been that there is more underreporting than overreporting and that, as a result, more is better. As retrospective surveys appropriately cast broader nets in search of eligible events and screening procedures become more sophisticated (as in the case of NCS), this assumption becomes less tenable.

The use of both reverse record check and comparative studies should still play some role in the development of new procedures. These provide external validation measures that normally cannot be obtained in survey research.[14] However, as more aggressive and "broad net" screening techniques are employed in victim surveys, much more pressure must be put on incident forms to filter out ineligible events and to classify events deemed eligible for inclusion. These methods increase the number of events that fall into the gray area; that is, events for which questions arise about *content validity.* In the case of rape, for example, where the question of consent is extremely important, victims may indicate a lack of consent when the circumstances of the event indicate ambiguity in that area (e.g., prior intimacy, absence of force). As Biderman (1981c) noted, critical to understanding interpersonal harm is an accurate assessment of what actually happened. This includes not only the circumstances of the event (who, what, when, where, how), but also something about the sequence of the events, and, possibly, the motivations of each of the actors.

The best way to address this issue is to collect the appropriate attributes related to the event and to use those attributes to construct a crime classification that reflects what happened. For example, if the purpose of the survey is to measure the number of crimes that occur in a population, an appropriate classification scheme can be developed similar to what is currently on NCS. Violent events, such as rape, can be classified by their typical components (e.g., forcible sexual penetration by a stranger). Those possessing all of these components can be put in one class, while events that have only two of the attributes would be put in another. They could all be classified as rape, or not, depending on the goal of the study. Furthermore, other attributes could be used to distinguish between degrees of certainty.

This is particularly important for events about which there is intense interest in the prevalence of the event but little consensus about definitions. The more complex classifications possible with attribute-based classification can prevent citing statistics for a large and heterogeneous class of events while claiming that all of the events in that class have the attributes of a much smaller and much more serious subset of these events. Loftin, Logan, and Addington (1999) are working on this type of classification scheme for hate crimes, and more work of this type should be done for other types of crime.

NCS currently has an extensive set of characteristics used to classify events. This is one of its strengths. However, these characteristics are geared primarily to the purpose of classifying events into official classes of crime. More research is needed in the development of incident forms to reflect both broader screening strategies and other uses of the data.

This can be done in several different ways. One way would be to conduct more qualitative analyses through collection of verbatim incident descriptions, focused respondent debriefings, or more intensive cognitive interviewing methods. More quantitative approaches would involve reinterview studies to examine test-retest reliability. The focus of these studies would be twofold. One would be to examine the attributes of events reported by different screening items. This would begin to provide evidence of content validity for reports using different screening strategies. A second focus would be to match the qualitative descriptions with the picture presented by the attribute-based classifications. The latter would provide some sense of the accuracy of the incident form in characterizing the event (e.g., motivations, interactional sequences, intent of victim). These would also provide the survey designer with a sense of the response processes that are used to formulate reports.

Continuations in the development of screening procedures

Parallel to the developments in screening for NCS, researchers interested in violence against women have developed screening methods based on many of the same principles. These methods have raised questions about the content of the survey instrument and methods that should be used to understand the types of events captured by a victimization survey. (For a more detailed description of these studies, especially as they compare with NCS, see Fisher and Cullen in this volume.)

The approach used by a number of violence researchers (e.g., the Conflict Tactics Scale) is to rely on extensive, and quite explicit, cues that narrowly focus on violent events, including sexual violence (Strauss et al. 1995). In addition to its narrow focus and explicit cues, this approach sets a different context than NCS. This is done by the use of a different type of introductory statement. Family conflict studies (Strauss 1998) set up the survey as one concerned with family or marital problems. As noted previously, this contrasts with the design of NCS, which frames the survey around concerns with crime. A second variant of this approach is surveys on personal safety (Tjaden and Thoennes 1998). This strategy directly addresses problems of failures of concept by not asking the respondent to make a value judgment about whether the incident is a criminal event.

Studies using this strategy have also adopted a quite different approach from NCS when sampling and interviewing household members to minimize problems with response inhibition and distortion. These studies (Tjaden and Thoennes 1998; Koss 1996; Kilpatrick et al. 1987) typically interview one person in a household. This is done, in part, to prevent others in the household from knowing what is actually on the questionnaire (e.g., if an abuser knows what is on the questionnaire, it may endanger the victim). Before administering the questions, the interviewer makes sure the respondent is in a private room where no one can overhear the conversation. It is made clear to the respondent that if someone walks in during the conversation, the interview will be continued at another time. This might include, for example, abruptly ending a telephone interview if the respondent feels it is necessary to do so.

The family conflict and personal safety surveys find extremely high rates of violence, especially rates of violence against women by nonstrangers. Strauss (1998), for example, presents a table that shows family conflict studies finding an average rate of violence among family members to be 16 percent for family conflict studies, 2 percent for a personal safety study (Tjaden and Thoennes 1998), and 0.9 percent for NCS.

The differences in the estimates appear to be functions of both the context of the items and the cuing used. As previously noted, NCS explicitly sets the context of the interview as one concerned with crime, whereas the conflict and safety studies explicitly avoid the use of any legal connotations (Strauss 1998, 3). Whether the event is considered a crime by the respondent is irrelevant. One example of this is a question used from the Sexual Experiences Survey (SES) (Koss 1993), which asks: "Have you had sexual intercourse when you didn't want to because a man gave you alcohol or drugs?"

This contrasts to the question on NCS that asks about sexual assault by a nonstranger:

> People don't often think of incidents committed by someone they know. Did you have something stolen from you OR were you attacked or threatened by: (a) someone at work or school, (b) a neighbor or friend, (c) a relative or family member, (d) any other person you've met or know?

SES asks about a situation that ignores any criminal intent, it simply asks about unwanted sex that was preceded by using alcohol or drugs. NCS asks about "incidents committed" within the context of being a victim of a crime.

Perhaps just as important as the context of the questions, the conflict and safety studies focus cues exclusively on violence, especially among nonstrangers. NCS screens on all types of crimes and only has one or two questions (with multiple cues) that specifically target (domestic) violence. The higher density of more specific cues will lead to reports of more events, as shown by the studies referenced earlier.

The discrepancies between these two approaches pose both conceptual and methodological challenges. Biderman (1981b, 1981c) makes the distinction between an indicator of crime and that of interpersonal harm (also see the discussion by Skogan 1981, 9–10). The former implicitly relies on the judgments of the respondent to report details about the culpability of the offender in the event. The latter does not, at least when initially asking the respondent to report the event. As Biderman (1981c, 49) notes, by restricting attention to events that are crimes, the survey may be leaving out events that are critically important for understanding the causes and consequences of interpersonal harm:

> Victims apply their own conceptions of whether the act indeed was "criminal," whether it should be made a matter for official attention, and whether the official system would be likely to act sufficiently in accordance with the victim's view and desires were a complaint made. . . . These grounds for excluding events from the criminal justice process include all of the

> classes of judgement that are the central objectives of victimology. . . . Victimological research that is based exclusively on officially recorded offenses thereby may be excluding most of the social phenomena with which it is particularly concerned.

As a measure of crime, however, the use of an indicator of interpersonal harm leaves out the formal criteria that separate events as criminal from others, such as accidents, legitimate retaliation, or other explanations that are not criminal.

The approaches taken by studies on violence against women have elaborated on the harm approach illustrated in the early work described by Biderman (1981b). They ask about actions and consequences without any reference to criteria related to criminal events. Many of these studies have not, however, taken the additional step of then asking details about the event to "establish sound actuarial knowledge of the magnitude of hazards various types of social situations present. [To do this] the data employed should be phenomenologically comprehensive and phenomenologically analyzable" (Biderman 1981c, 51). Further research needs to begin to move in this direction in order to begin conceptually relating harm to what society (and victims of harm) conventionally view as crimes.

While the context of the survey is important in defining the scope of eligible events for the respondent, the cues presented also serve this function, as well as influence the process of locating specific events in memory. Intensive cuing of particular types of events should yield more reports of these events, as illustrated by the NCS experiments, as well as the violence studies. This implies that reporting events will reflect, in part, the distribution of cues in addition to the distribution of crime events. Further evidence of this can be seen in a comparison of NCS and the National Violence Against Women Study. The latter cued extensively for rape and for crimes among intimates. This resulted in rates of physical assault that were generally higher in NCS, with the exception of rape and assaults by intimates (Bachman 1998). We know little about how cues interact with the survey context (e.g., harm versus crime). This requires more research into the effects of cues and how they should be allocated given the purpose of a particular survey instrument.

Sample design, coverage, and nonresponse

A number of issues related to the sample design, coverage, and nonresponse continue to be problematic for victimization surveys. These include developing efficient sample designs, improving coverage, and nonresponse imputation for groups at high risk of victimization.

Developing efficient sample designs

Conducting a victimization survey is an expensive undertaking. Because a relatively small percentage of people will report an event for a fixed time period, large sample sizes are needed to generate reliable population estimates. Some reduction in reliability can be compensated for by lengthening the reference period and improving the screening methods. However, both of these have their limitations and costs (as noted earlier). The former increases memory error, and the latter complicates the design and detail required for the instrumentation (and time needed to design and administer the instrument).

One of the major innovations over the past 10 to 15 years has been to increase the number of surveys that are done by telephone using samples generated by random digit dialing methods. By dispensing with expensive area-based, in-person designs, the project can increase the number of interviews per dollar spent. Many of the surveys referenced that have examined violence and the enduring effects of victimization have been conducted using this method. The disadvantage of this method is that it typically yields relatively low response rates (in the 60- to 70-percent range) and misses the population that does not have a telephone.

Yet to be fully exploited are less traditional methods, especially those using networks of victims. Network designs are based on the idea of using respondents as informants on other persons in the network. The respondent is asked to provide information on whether other persons that the respondent knows have been victims of a crime. If the information is accurate, and one can precisely enumerate the counting rules involved, then it is possible to develop estimates of victimization. Czaja and Blair (1990) conducted an evaluation of this method and did not find the network methods they employed to be better than a traditional approach using a mean square error criterion. However, they noted a number of problems with their design and recommended further research into this type of sampling process.

Other methods for reducing the expense involved in these surveys have been suggested, including the use of other types of multiplicity estimators. One variant of this logic is to use telephone prefixes or area-locations to find victims. Oversampling in areas yielding a large number of reports of victimization, using this logic, may produce more efficient sample designs. Of course the overall efficiency and utility of any such strategies depend on the goals of the survey (e.g., estimating population rates versus comparison across different subpopulations). Nonetheless, using some type of stratification or double sampling process needs to be explored when trying to reduce the costs of victimization surveys. If this were done, the utility of this method of evaluation and analysis would be greatly increased.

Nonresponse and coverage

A persistent observation in victimization surveys is relatively equal simple assault rates between whites and other minority groups. One possible explanation for this is differential error related to response problems such as comprehension and recall (Skogan 1981). Another explanation is differential coverage and nonresponse. There is evidence, in fact, that both nonresponse and coverage must be taken into account for victimization surveys (Reiss 1977; Biderman and Cantor 1984; Griffin Saphire 1984; Stasny 1991). These issues may be particularly problematic for telephone surveys, which are increasingly being used because of the economies they offer. With respect to coverage, research with NCS has shown that persons who do not own telephones are the most likely to be victimized (Woltman, Turner, and Bushery 1980). With respect to nonresponse, telephone surveys generally achieve response rates that are 15 to 20 percent lower than in-person surveys and that may be particularly vulnerable to issues of bias, especially for certain subgroups.

Even for in-person surveys, however, there is evidence that coverage and nonresponse biases are problematic. Particular problems have been found for certain minority groups, especially Hispanics and young black males. The hypothesis is that the surveys simply miss those who are most likely to be subject to crime. Using longitudinal data from NCS, Reiss (1977) found that persons who have high residential mobility have much higher victimization rates than those who are not mobile (see also Biderman and Cantor 1984). Further elaboration taking advantage of this correlation has found mobility to be an important covariate when imputing data (Griffin Saphire 1984; Stasny 1991). For cross-sectional analysis, it may be possible to adapt information from reports of mobility for respondents (e.g., how many times a respondent moved in the last year). Future research should elaborate on this correlation, as well as developing other indicators for uses in imputation.

The correlation with mobility should also be viewed as a proxy for coverage problems. Persons who are most likely to be missed are, in part, those persons with unstable living situations who either may not have a residence at any particular point in time or may not be considered part of the residence when the interviewer conducts the initial household enumeration (Martin 1996). Indirect evidence of coverage problems on NCS was found by Cook (1985), who compared estimates of gunshot victims with external records available from hospital emergency rooms. These data seemed to indicate a gross undercount of such injuries on NCS. One leading explanation is that NCS misses those individuals who are most likely to be a victim of this type of crime.

Going beyond the assumption of crime as a point-in-time event

One of the weaknesses of the victim survey method is its emphasis on crime as incidents occurring at a point in time. This approach to crime stems from the carryover from the attempts to have the survey mimic UCR. Many of the types of crimes that have the greatest social (as opposed to individual) import are more readily approached as conditions that endure rather than incidents that begin and end at a given point in time (Biderman 1975). Among the kinds of victimizations that could be conceived and measured in prevalence rather than incident terms are various forms of continuing terrorization and extortion. This might include, for example, a spouse or sexual partner in continuous fear of violence or school children who must routinely give up their lunch money to gangs of fellow students. Here the victim is in virtually a continuous state of threat and victimization, but the survey requires that this condition be divided into its component parts, which minimizes the disruption of life and social relations.

To some degree, victimization surveys provide information about these kinds of situations through tabulations of series victimizations. Historically in NCS these were defined as three or more similar incidents of victimization mentioned by a respondent, but which, because of their frequency or similarity, the respondent cannot individually date or differentiate from one another. The terrorized spouse then could be identified in NCS through repeated incidence of spouse beating or the terrorized school child by repeated robberies. It is not necessary for the specific acts defining victimization to exist for there to be continuous victimization. To make a threat credible to a victim and to continue the state of terrorization, the offender need not continually repeat his threat or actually inflict violence.

Even the series victimization is captive to the point-in-time logic in that most of the questioning regarding series incidents is done for the purpose of counting incidents and not for the purpose of establishing duration or patterns of events. It may be more useful for understanding conditions of continuous victimization to have respondents explain the interrelationship of events in a series or to talk about the factors that are contributing to persistence. One logical way to do this is with a longitudinal design. Directly asking respondents to draw linkages between events, if there are any, would certainly be useful in identifying who is in the condition of continuous victimization.

Longitudinal surveys are expensive and time consuming to complete. Linking events within the same reference period would provide a significant advance and would not be as expensive to implement. Recent changes to the series victimization procedures in NCS have moved in this direction by asking respondents if all of the events in the series involved the same offender, if they occurred in the same place, and if the victimization is continuing. Unfortunately, these questions are asked only of respondents who satisfy the conditions of the series victimization procedure. It would be better if some provision were made for asking about the interrelationship of events in all instances of repeated victimization within a given interview period.

Another facet of victimization that might be better suited to a prevalence rather than an incidence approach to measurement is the durable and psychic consequences of point-in-time crime events. Here questioning that elicited the initiation and termination of conditions resulting from a crime event would be useful. Also, in surveys like NCS that involve more than one interview, asking about the conditions across interviews would be helpful in establishing the persistence of these conditions.

Conclusions

Self-report surveys of victimization have become an established feature of crime statistics in the United States and throughout the world. They are used routinely as social indicators and as tools for building criminological theory. Over the past 25 years, we have learned a great deal about asking persons to recount their victimization experiences. Much of that knowledge has come from the National Crime Survey and its antecedents. The process of selling, planning, and fielding the first surveys framed and informed many of the issues that needed to be resolved to conduct a household survey of victims. During the initial development there was a tension between a legalistic emphasis and one oriented to more traditional survey design concerns. Survey methodologists found that it was better to organize the survey to facilitate the recall and reporting task. Those interested in comparisons with police data were concerned with developing a social indicator comparable to police-reported measures of crime. Elements of both approaches to designing victimization surveys were retained by the Census Bureau and DOJ when they began fielding the ongoing NCS. However, a heavy emphasis was put on comparisons to the police record systems.

The subsequent methodological work moved NCS closer to an approach based on facilitating recall. Moreover, changes in the NCS design and the appearance of other victimization surveys have reiterated the lesson of the pilot studies: Data from victim surveys are heavily influenced by their design. The appearance of these alternative designs and the very large differences in reporting that resulted give us both the motivation and the ability to learn much more about the method. In the future, it will be critical to compare variations in design with differences in reporting to better understand the implications of the methodology (Lynch 1996). Unless this additional research and development work is done, the substantial effect of design on the resulting data will raise suspicions about whether results are less a reflection of the crime problem than they are of the design of the survey. If this occurs, then the widespread acceptance of this method may decline.

This paper and the authors have benefited immeasurably from discussions with and the writings of Albert D. Biderman.

Notes

1. There are a number of excellent reviews of various aspects of the design and contribution of victimization surveys. For more detailed discussion of specific topics, see publications by Gottfredson (1986), Sparks (1981), Hindelang (1976), and Skogan (1981).

2. From 1973 to 1991, the survey was called the National Crime Survey. Since 1991, it has been referred to as the National Crime Victimization Survey. Because we often refer to the survey throughout its existence, we use National Crime Survey (NCS) throughout the text.

3. For a more complete discussion of UCR's organization and procedures, see *The Uniform Crime Reporting Handbook* (U.S. Department of Justice, Federal Bureau of Investigation 1984). For a discussion of the implications of these aspects of UCR for the quality of the resulting data, see Biderman and Lynch (1991).

4. For a more complete description of the design, see Rand and Taylor (1995).

5. For a detailed review of the early development of NCS, see Hindelang (1976, 21–76).

6. The size of the sample required will be affected by the productivity of the screening interview. The less productive the screening interview, the greater the projected sample size for the same level of precision. Although the more evocative screening procedures used in the BSSR pilot study may have reduced the sample size required, even in this case, extremely large samples would be required to estimate rare crimes like robbery and rape with any precision.

7. Several additional factors led to adopting a 6-month period. One was related to the observation that shorter time periods led to more accurately dated events. Because NCS relies on this dating to determine which incidents are within a reference year, this type of error had to be minimized. The second related to the timing of data releases. With a 12-month period, an additional 6 months would have to elapse (compared with a 6-month period) because of the need to interview all persons who could possibly report a crime within the appropriate calendar year. For example, for a 6-month period, interviews needed to generate an estimate for year *t* would have to wait until interviews in June of year *t+1* are finished. For a 12-month period, one would have to wait until November interviews in year *t+1* are completed.

8. At the time these changes were made, the National Crime Survey was renamed the National Crime Victimization Survey.

9. RPE was conducted with bounded reference periods. This should minimize this type of telescoping. Nonetheless, the assumption that more is better as a measure of improved quality may overestimate the true gains in quality achieved by shortening the reference period.

10. A shorter reference period would result in covering less of the calendar period in each interview (3 months rather than 6 months). This cuts the sample size by a significant proportion.

11. It is less clear how other types of events may be affected. For example, the presence of other household members may actually encourage reporting if the other household members actually know about the event.

12. It should be noted that the evidence related to self-administration does not link improvements in reporting to external validation criterions (e.g., biological tests, arrest records).

13. This perspective is somewhat simplistic. Methods have effects on not only the direction of relationships (bias), but also reliability and statistical power. For example, making concepts clearer or using a self- rather than a proxy-respondent may reduce sampling error by eliminating variation due to misunderstandings or faulty knowledge.

14. Other external criteria that should be considered, especially when using more broad-net approaches, could be records that capture more general sets of injuries or incidents. This might include, for example, records from emergency rooms, hospital records, and insurance claims.

References

Aldrich, Howard, and Albert J. Reiss. 1970. The effects of civil disorder on small business in the inner city. *Journal of Social Issues* 26:187–206.

Bachman, Ronet. 1998. Comparing estimates of violence against women from the National Survey of Violence Against Women and the National Crime Victimization Survey. Paper presented at the National Conference on Violence Against Women, 18 October, Arlington, Virginia.

Biderman, Albert D. 1981a. Notes on the methodological development of the National Crime Victimization Survey. In *The National Crime Survey: Working papers, volume 1: Current and historical perspectives,* edited by Robert G. Lehnen and Wesley G. Skogan. NCJ 75374. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

———. 1981b. A social indicator of interpersonal harm. In *The National Crime Survey: Working papers, volume 1: Current and historical perspectives,* edited by Robert G. Lehnen and Wesley G. Skogan. NCJ 75374. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

———. 1981c. When does interpersonal violence become crime?—Theory and practice. In *The National Crime Survey: Working papers, volume 1: Current and historical perspectives,* edited by Robert G. Lehnen and Wesley G. Skogan. NCJ 75374. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

———. 1975. *Notes on the significance of measurements of events and of conditions by criminal victimization surveys.* BSSR 0003–58. Washington, D.C.: Bureau of Social Science Research.

———. 1967. Surveys of population samples for estimating crime incidence. *Annals of the American Academy of Social and Political Science* 374 (November): 16–34.

———. 1966. *Report on design of a national study.* Washington, D.C.: Bureau of Social Science Research, Inc.

Biderman, Albert D., and David Cantor. 1984. A longitudinal analysis of bounding respondent conditioning and mobility as sources of panel bias in the National Crime Survey. In *American Statistical Association 1984 proceedings of the Section on Survey Research Methods.* Washington, D.C.: American Statistical Association.

Biderman, Albert D., David Cantor, James P. Lynch, and Elizabeth Martin. 1986. *Final report of the National Crime Survey redesign.* Washington, D.C.: Bureau of Social Science Research.

Biderman, Albert D., David Cantor, and Albert J. Reiss, Jr. 1985. A quasi-experimental analysis of personal victimization by household respondents in the NCS. Paper presented at the Annual Meetings of the American Statistical Association, Philadelphia.

Biderman, A.D., L.A. Johnson, J. McIntyre, and A.W. Weir. 1967. *Report on a pilot study in the District of Columbia on victimization and attitudes toward law enforcement.* President's Commission on Law Enforcement and Administration of Justice, Field Surveys no. 1. Washington, D.C.: U.S. Government Printing Office.

Biderman, Albert D., and James P. Lynch. 1981. Recency bias in data on self-reported victimization. *American Statistical Association 1981 proceedings of the Social Statistics Section.* Washington, D.C.: American Statistical Association.

Biderman, Albert D., and James P. Lynch, with James Peterson. 1991. *Understanding crime incidence statistics: Why the UCR diverges from the NCS.* New York: Springer-Verlag.

Biderman, Albert D., and Albert J. Reiss, Jr. 1967. On exploring the "dark figure" of crime. *Annals of the American Academy of Political and Social Science* 374 (November): 1–15.

Bushery, J. 1981. Recall bias for different reference periods in the National Crime Survey. Paper presented at the Annual Meetings of the American Statistical Association, Boston.

Cantor, David. 1985. Operational and substantive differences in changing the NCS reference period. In *American Statistical Association 1985 proceedings of the Social Statistics Section.* Washington, D.C.: American Statistical Association.

Cohen, Lawrence, and David Cantor. 1981. Residential burglary in the United States: Lifestyle and demographic factors associated with the probability of victimization. *Journal of Research in Crime and Delinquency* 18:113–27.

Cohen, Lawrence, and Marcus Felson. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review* 44:588–608.

Coker, Ann, and Elizabeth Stasny. 1996. *Adjusting the National Crime Victimization Survey's estimates of rape and domestic violence for "gag" factors.* NCJ 173061. Washington, D.C.: U.S. Department of Justice, National Institute of Justice.

Conway, Mark, and Sharon Lohr. 1994. A longitudinal analysis of factors associated with the reporting of violent crime to the police. *Journal of Quantitative Criminology* 10:23–39.

Cook, Philip. 1985. The case of the missing victim: Gunshot wounding in the National Crime Survey. *Journal of Quantitative Criminology* 1:91–102.

Couper, M.P., R.P. Baker, J. Bethlehem, C.Z.F. Clark, W.L. Nicholls, and J. O'Reilly, eds. 1998. *Computer assisted survey information collection.* New York: John Wiley & Sons.

Czaja, R., and J. Blair. 1990. Using network sampling in crime victimization surveys. *Journal of Quantitative Criminology* 6:185–206.

Czaja, R., J. Blair, B. Bickart, and E. Eastman. 1994. Respondent strategies for recall of crime victimization incidents. *Journal of Official Statistics* 10:257–276.

Dodge, R.W., and F.D. Balog. 1987. *Series crimes: Report of a field test.* Technical Report, NCJ 104615. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Dodge, R.W., and A.G. Turner. 1981. Methodological foundations for establishing a national survey of victimization. In *The National Crime Survey: Working papers, volume 1: Current and historical perspectives,* edited by Robert G. Lehnen and Wesley G. Skogan. NCJ 75374. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Ellingsworth, Daniel, Graham Farrel, and Kenneth Pease. 1995. A victim is a victim is a victim: Chronic victimization in four sweeps of the British Crime Survey. *British Journal of Criminology* 35:360–365.

Ennis, Philip H. 1967. *Criminal victimization in the United States: A report of a national survey.* President's Commission on Law Enforcement and Administration of Justice, Field Surveys no. 2. Washington, D.C.: U.S. Government Printing Office.

Finkelhor, David. 1997. The victimization of children and youth: Developmental victimology. In *Victims of crime,* edited by Robert Davis, Arthur Lurigio, and Wesley Skogan. 2d ed. Thousand Oaks, California: Sage Publications.

Forrester, D., M. Chatterton, and K. Pease. 1988. *The Kirkholt burglary prevention project.* Rochdale Crime Prevention Unit Paper no. 13. London: Home Office.

Garofalo, James. 1977. *The police and public opinion: An analysis of victimization and attitude data in 13 American cities.* NCJ 42018. Washington, D.C.: U.S. Department of Justice, National Criminal Justice Information and Statistics Service.

Gidycz, Christine, and Mary P. Koss. 1991. Predictors of long-term sexual assault trauma among a national sample of victimized college women. *Violence and Victims* 6:175–90.

Gilbert, Neil. 1992. Realities and mythologies of rape. *Society* 29 (4): 4–10.

Gottfredson, Michael. 1986. Victimization surveys. In *Crime and justice: An annual review of research,* edited by Michael Tonry and Norval Morris. Vol. 7. Chicago: University of Chicago Press.

———. 1984. *Victims of crime: The dimensions of risk.* Home Office Research and Planning Unit Report no. 81. London: Her Majesty's Stationery Office.

Griffin Saphire, Diane. 1984. *Estimation of victimization prevalence using data from the National Crime Survey.* New York: Springer-Verlag.

Harlow, C.W. 1989. *Injuries from crime.* Special Report, NCJ 116811. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Hindelang, Michael. 1981. Variations in sex-race-age-specific incidence rates of offending. *American Sociological Review* 46:461–474.

———. 1978. Race and involvement in common law personal crimes. *American Sociological Review* 43:93–109.

———. 1976. *Criminal victimization in eight American cities: An analysis of common theft and assault.* Cambridge, Massachusetts: Ballinger.

Hindelang, Michael, Michael Gottfredson, and James Garofalo. 1978. *Victims of personal crime: An empirical foundation for a theory of personal victimization.* Cambridge, Massachusetts: Ballinger.

Hough, Michael. 1987. Offender's choice of target: Findings from victim surveys. *Journal of Quantitative Criminology* 3 (4): 275–281.

Hubble, David, and B.E. Wilder. 1988. Preliminary results from the National Crime Survey CATI experiment. In *American Statistical Association 1988 proceedings of the Section on Survey Research Methods.* Washington, D.C.: American Statistical Association.

Kalish, Carol. 1974. *Crimes and victims: A report on the Dayton-San Jose pilot survey of victimization.* Research Report, NCJ 13314. Washington, D.C.: U.S. Department of Justice, National Criminal Justice Information and Statistics Service.

Kennedy, Leslie, and David Forde. 1999. *When push comes to shove: A routine conflict approach to violence.* Albany: State University of New York Press.

Kilpatrick, D.G., B.E. Saunders, L.J. Veronen, C.L. Best, and J.M. Von. 1987. Criminal victimization: Lifetime prevalence, reporting to police, and psychological impact. *Crime & Delinquency* 33:479–489.

Kindermann, Charles, James P. Lynch, and David Cantor. 1997. *The effects of the redesign on victimization estimates.* NCJ 164381. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Kleck, Gary. 1996. Crime, culture conflict, and the sources of support for gun control: A multilevel application of the general social surveys. *American Behavioral Scientist* 39 (4): 387–404.

———. 1991. *Point blank: Guns and violence in America.* New York: Aldine de Gruyter.

Kobilarcik, E.L., C.A. Alexander, R.P. Singh, and G.M. Shapiro. 1983. Alternative reference periods for the National Crime Survey. In *American Statistical Association 1983 proceedings of the Section on Survey Research Methods.* Washington, D.C.: American Statistical Association.

Koss, Mary P. 1996. The measurement of rape victimization in crime surveys. *Criminal Justice and Behavior* 23:55–69.

———. 1993. Detecting the scope of rape: A review of prevalence research methods. *Journal of Interpersonal Violence* 8 (2): 198–222.

Kury, Helmut. 1993. Crime in East and West Germany: Results from the First Intra-German Victims' Survey. In *Fear of crime and criminal victimization,* edited by Wolfgang Bilsky, Christian Pfeiffer, and Peter Wetzels. Stuttgart: Enke Verlag.

Lauritsen, Janet L., and K.F. Davis Quinnet. 1995. Repeat victimization among adolescents and young adults. *Journal of Quantitative Criminology* 11:143–166.

Lauritsen, J.L., R.J. Sampson, and J. Laub. 1991. The link between offending and victimization among adolescents. *Criminology* 29:265–292.

Lehnen, Robert G., and Wesley G. Skogan. 1981. *The National Crime Survey: Working papers, volume 1: Current and historical perspectives.* NCJ 75374. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

Lepkowski, James. 1981. *Sample design issues from the National Crime Survey.* Ann Arbor: University of Michigan, Survey Research Center.

Loftin, Colin, Wayne A. Logan, and Lynn Addington. 1999. A proposed binary-based evaluative framework for the categorization of hate crimes. Paper presented at the Conference on Hate Crime Measurement, 23 April, School of Criminal Justice, University at Albany, State University of New York.

Lynch, James P. 1996. Clarifying divergent estimates of rape from two national surveys. *Public Opinion Quarterly* 60 (3): 410–29.

Lynch, James P., Michael Berbaum, and Mike Planty. 1998. *Investigating repeated victimization with the NCVS.* Final report, NIJ grant no. 97–IJ–CX–0027. Washington, D.C.: U.S. Department of Justice, National Institute of Justice.

Lynch, James P., and David Cantor. 1992. Ecological and behavioral influences on property victimization at home: Implications for opportunity theory. *Journal of Research in Crime and Delinquency* 29 (August): 335–362.

Martin, Elizabeth. 1996. Household attachment and survey coverage. Paper presented at the Joint Statistical Meetings of the American Statistical Association, 4–8 August, Chicago.

Maxfield, Michael. 1987. Household composition, routine activity, and victimization: A comparative analysis. *Journal of Quantitative Criminology* 3:301–320.

Mayhew, Pat. 1995. Some methodological issues in crime victimization surveys. In *Crime victimization surveys in Australia—Conference proceedings.* Brisbane: Criminal Justice Commission.

———. 1993. Reporting crimes to the police: The contributions of victim surveys. In *Fear of crime and criminal victimization,* edited by Wolfgang Bilsky, Christian Pfeiffer, and Peter Wetzels. Stuttgart: Enke Verlag.

McDowall, D., and C. Loftin. 1992. Comparing the UCR and NCS over time. *Criminology* 30:125–132.

McPhail, Beverly. 1995. The term is “femicide” and society must do something. *Houston Chronicle,* 1 October.

Miethe, T.D., and R.F. Meier. 1994. *Crime and its social context: Toward an integrated theory of offenders, victims, and situations.* Albany: State University of New York Press.

Miller, Peter V., and Robert M. Groves. 1985. Matching survey responses to official records: Validity in victimization reporting. *Public Opinion Quarterly* 49:366–380.

Murray, David. N.d. *America’s phony rape epidemic.* Washington, D.C.: Statistical Assessment Service.

National Commission on Law Observance and Enforcement. 1931. *Police.* Washington, D.C.: U.S. Government Printing Office.

Nelson, James F. 1980. Multiple victimization in American cities: A statistical analysis of rare events. *American Journal of Sociology* 85:871–891.

Neter, John, and Joseph Waksberg. 1964. Conditioning effects from repeated household interviews. *Journal of Marketing* 29:51–56.

New York Times. 1981. Editorial, The conflict of UCR–NCS statistics, 19 April.

Norris, Fran H., Krzysztof Kaniasty, and Martie P. Thompson. 1997. The psychological consequences of crime: Findings from a longitudinal population-based study. In *Victims of crime.* 2d ed. Edited by Robert Davis, Arthur Lurigio, and Wesley Skogan. Thousand Oaks, California: Sage Publications.

Parkinson, R.P., A.L. Paez, and N.F. Howard. 1977. *Criminal victimization in the United States, 1974: A National Crime Survey report.* NCJ 39467. Washington, D.C.: U.S. Department of Justice, National Criminal Justice Information and Statistics Service.

Pease, Ken. 1998. *Repeat victimization: Taking stock.* Crime Detection and Prevention Series. Paper 90. London: Home Office, Police Research Group.

Penick, B.K.E., and M. Owens. 1976. *Surveying crime.* Washington, D.C.: National Academy Press.

Percy, A., and Pat Mayhew. 1997. Estimating sexual victimization in a national crime survey: A new approach. *Studies on Crime & Crime Prevention* 6 (2): 125–50.

Persley, Carolyn. 1995. The National Crime Victimization Survey redesign: Measuring the impact of new methods. In *American Statistical Association 1995 proceedings of the Section on Survey Research Methods.* Washington, D.C.: American Statistical Association.

Poggio, E.C., S.D. Kennedy, J.M. Chaiken, and K.E. Carlson. 1985. *Blueprint for the future of the Uniform Crime Reporting program: Final report of the UCR study.* NCJ 98348. Washington, D.C.: U.S. Department of Justice, Federal Bureau of Investigation and Bureau of Justice Statistics.

Polvi, N., Terah Looman, Charles Humphries, and Ken Pease. 1990. Repeat break-and-enter victimization: Time course and crime prevention opportunity. *Journal of Police Science and Administration* 17:8–11.

Rand, Michael, and Bruce Taylor. 1995. The National Crime Victimization Survey redesign: New understandings of victimization dynamics and measurement. Paper presented at the Joint Statistical Meetings of the American Statistical Association, 13–17 August, Orlando, Florida.

Reiss, Albert J. 1977. *Victim proneness by type of crime in repeat victimization.* New Haven, Connecticut: Institute for Social Policy Studies.

———. 1969. Impact of crime on small business. U.S. Senate. Select Committee on Small Business. *Hearings before the Select Committee on Small Business.* 91st Cong., 1st sess. 21–23 May and 22 July.

———. 1967a. Impact of crime on small business. U.S. Senate. Select Committee on Small Business. *Hearings before the Select Committee on Small Business.* 90th Cong., 1st sess. 24–26 April.

———. 1967b. *Measurement of the nature and the amount of crime: Studies in crime and law enforcement in major metropolitan areas.* President's Commission on Law Enforcement and Administration of Justice, vol. 1 of Field Surveys no. 3. Washington, D.C.: U.S. Government Printing Office.

Resnick, Heidi S., D. Kilpatrick, B. Dansky, B.E. Saunders, and Connie Best. 1993. Prevalence of civilian trauma and post traumatic stress disorder in a representative sample of women. *Journal of Consulting and Clinical Psychology* 61 (6): 984–999.

Singer, S.I. 1986. Victims of serious violence and their criminal behavior: Subcultural theory and beyond. *Violence and Victims* 1:61–70.

———. 1981. Homogeneous victim-offender populations: A review and some research implications. *Journal of Criminal Law and Criminology* 72 (2): 779–788.

Skogan, Wesley G. 1984. Reporting crimes to the police: The status of world research. *Journal of Research in Crime and Delinquency* 21:113–137.

———. 1981. *Issues in the measurement of victimization.* NCJ 74682. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.

———. 1974. The validity of official crime statistics. *Social Science Quarterly* 55:35–48.

Sparks, Richard F. 1981. Multiple victimization: Evidence theory and future research. *Journal of Criminal Law and Criminology* 72:762–778.

Stasny, Elizabeth A. 1991. Hierarchical models for survey and nonresponse probabilities. *Journal of the American Statistical Association* 86:296–303.

Strauss, M.A. 1998. The controversy over domestic violence by women: A methodological, theoretical, and sociology of science analysis. In *Violence in intimate relationships,* edited by X.B. Arriaga and S. Oskamp. Thousand Oaks, California: Sage Publications.

Strauss, M.A., S.L. Hamby, S. Boney-McCoy, and D.B. Sugarman. 1995. *The Revised Conflict Tactics Scales (CTS2–form A).* Durham, New Hampshire: University of New Hampshire, Family Research Laboratory.

Tjaden, P., and N. Thoennes. 1998. *Stalking in America: Findings from the National Violence Against Women Survey.* Research in Brief, NCJ 169592. Washington, D.C.: U.S. Department of Justice, National Institute of Justice, and U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Tourangeau, R., and T. Smith. 1996. Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly* 60:275–304.

Turner, C.F., L. Ku, and F.L. Sonenstein. 1996. Impact of ACASI on reporting male-male sexual contacts: Preliminary results from the 1995 Survey of Adolescent Males. In *Health survey research methods,* edited by R. Warnecke. Hyattsville, Maryland: U.S. Department of Health and Human Services, National Center for Health Statistics.

U.S. Department of Justice, Federal Bureau of Investigation. 1984. *Uniform Crime Reporting handbook.* Washington, D.C.

U.S. Department of Justice, National Criminal Justice Information and Statistics Service. 1976. *Criminal victimization in the United States 1973.* NCJ 34732. Washington, D.C.

U.S. House. 1977. Committee on the Judiciary. *Suspension of the National Crime Survey: Hearing before the Subcommittee on Crime.* 95th Cong., 1st sess., 13 October.

Washington Post. 1981. Editorial, Two reports on violent crime in the U.S., 5 May.

Woltman, Henry F., Anthony Turner, and John M. Bushery. 1980. A comparison of three mixed mode interviewing procedures in the National Crime Survey. *Journal of the American Statistical Association* 75 (371): 534–543.

Zawitz, Marianne W., ed. 1988. *Report to the Nation on crime and justice.* 2d ed. NCJ 111096. Washington, D.C.: U.S. Department of Justice, Bureau of Justice Statistics.