

categories, noting, for example, that with larceny-theft there was no standard basis for assessing the value of stolen property. Thus, comparisons of theft statistics across jurisdictions could be misleading, if not completely meaningless.

The misuse and misrepresentation of UCR data has a long history. In an interesting example of this phenomenon, on January 1, 1962, *U.S. News and World Report* published an interview with J. Edgar Hoover, head of the FBI. Hoover asserted that crime was becoming more serious and frequent in the United States. In support, he noted that between 1950 and 1960, "serious crime" had increased by 98 percent, although the population of the United States grew by only 18 percent. More specifically, Hoover invoked the rather questionable practice of combining all crimes in the crime index to claim that "in 1960, more than 7,700 police departments of this country reported 1,861,300 murders, forcible rapes, robberies, aggravated assaults, burglaries, automobile thefts, and larcenies of \$50 or more." The bulk of this total was, however, comprised of more minor crimes such as larcenies. Hoover attributed these increases in crime to a decline in parental authority as well as more general moral standards in the United States. Presaging a view that is popular in explanations of crime today, Hoover also asserted that there was a relationship between exposure to violence in the media and crime: "The highly suggestive, and at times, offensive scenes, as well as the frequent portrayal of violence and brutality on television screens and in motion pictures, are bound to have an adverse effect on young people."

### The Dark Figure and Additional Problems with Crime Statistics

[A criminologist] studies the criminals convicted by the courts and is then confounded by the growing clamor that he is not studying the real criminal at all, but an insignificant proportion of non-representative and stupid unfortunates who happened to become enmeshed in technical legal difficulties. (Tappan 1947:96)

Although official data from the UCR and other sources were commonly used by journalists, criminologists, and other social scientists to comment on crime trends and the causes of crime, many commentators began to recognize their potential weaknesses. Beattie (1941) noted that police statistics were manipulable for political purposes and hence questionable in their validity. "Traditionally, police departments are anxious to make a good showing in their annual figures, and there is, therefore, a

natural tendency to record and report those facts which show a good administrative record on the part of the department" (p. 21). Vold (1935), commenting on an alleged crime wave in St. Paul, Minnesota, in the early 1930s, suggested that "it has been impossible for the present writer to determine whether this represents an actual increase in serious crime in this part of the country, or merely much needed improvement in police statistics" (p. 802). Sellin (1931) also argued that crime statistics primarily reflected the activity of law enforcement personnel and therefore could not be accepted as indicating particular trends in crime. Indeed, along these lines, Sellin is perhaps best known for his suggestion that "in general, it may be said that the value of a crime rate for index purposes is in reverse ratio to the procedural distance from the commission of crime and the recording of it as a statistical unit" (p. 346).

Early 20th-century criminologists also recognized that crime rates could be affected by the policies and practices of individual police departments. For example, in London, a change in recording practices was implemented in 1932, whereby citizens' reports to the police of thefts that had not previously been recorded in official data were now included. This change resulted in an increase in indictable (more serious) offenses from approximately 26,000 in 1931 to 83,000 in 1932 (Radzinowicz 1945)—an increase of more than 300 percent. Similarly, a change in police administration in New York City in 1950, which led to a modification of recording practices, resulted in increases of 400 percent in robberies, 700 percent in larcenies, and more than 1,300 percent in burglaries over a one-year period (Brantingham and Brantingham 1984).

One of the most influential studies that drew attention to the dark figures of crime and the possibility of criminal justice system biases in generating police and judicial statistics was Robison's (1936) study of delinquent youth in New York state. Robison challenged the common use of juvenile court statistics in examinations of delinquency and was particularly critical of the delinquency area technique adopted by the Chicago School sociologists Shaw and McKay (1931). Shaw and McKay, relying on juvenile court referral data in a number of U.S. cities, found that the highest rates of delinquency occurred in neighborhoods characterized by rapid population change, poor housing, poverty, tuberculosis, adult crime, and mental disorders. In addition, they found that delinquency rates were highest in inner-city or core areas of cities but declined with the distance from the center of the city. They emphasized the importance of social disorganization in explaining how youth in these areas became involved in delinquency. This disorganization was manifested in the alienation of children from their

parents and adult institutions, resulting in detachment from informal social controls that would normally produce conformity.

Robison (1936) took issue with Shaw and McKay's studies. She argued that the method used by these sociologists not only was invalid for measuring the extent and nature of juvenile delinquency but also was ineffective for providing theoretical explanations of delinquency and policies for delinquency prevention. The Chicago School studies and others like them had implicitly assumed that the extent of delinquency could be measured accurately through the identification of apprehended delinquents—a questionable assumption, at best. The studies also failed to sufficiently differentiate between different types of delinquency, by generally treating truancy, stealing, and malicious mischief as similar behaviors.

Robison (1936) challenged the commonly held notion that there was a fundamental association between poverty, race/ethnicity, and delinquency, claiming instead that these relationships were due to the differential treatment of individuals of different socioeconomic status and racial/ethnic backgrounds by criminal justice system officials. She argued that the customs of diverse nationality and cultural groups had an impact on which youth would be labeled officially delinquent; therefore, variation in the behavior of parents and authorities confronted with "troublesome children" was potentially more important in determining official rates of delinquency than real differences in the proportion of delinquents in these groups. Robison further noted that "is it not rather a human tendency to regard less critically the behavior of the children who live on the right side of the tracks than that of the urchins who, surprised in suspicious activity, react with an almost reflex furtiveness? The policeman is more prone to suspect the poor man's child of theft and the rich man's child of a prank" (p. 30).

Robison's study combined administrative records of delinquent behavior from a cross section of public schools, family agencies, and agencies that cared for neglected children in New York City. The inclusion of these additional data resulted in an increase in the total number of delinquents from 7,090 (the figure derived from juvenile court records) to 15,898 children who were under the care for the commission of delinquent behavior in 1930. In her analysis of these and other data, she pointed to racial and ethnic differences in the identification and processing of delinquency cases. Referring specifically to official juvenile court data, which indicated that there were seven Catholic youth for each Protestant youth processed and three Catholic to every Jewish, Robison noted that "the same behavior by a boy in a Jewish family, no matter where he lives, has evidently not the same chance of being labeled delinquent and referred to the court as that of a boy

in an Italian family. Apparently a misbehaving boy in a Protestant family has even less chance of being referred as delinquent, either to the court or an unofficial agency, than the Italian or Jewish boy, and with Negro boys the chances are still different" (pp. 195-6). In short, Robison's research suggested that reliance on official data to study juvenile delinquency would result in misleading theories about its causes and misguided policy solutions to the juvenile delinquency problem.

Another prominent critic of official crime data was Edwin Sutherland (1947), who made a number of important points regarding the weaknesses in these data. Sutherland was one of the first to note that, for the purposes of comparison, crime statistics needed to be calculated in proportion to the population or some relevant base. He suggested that the rates should be corrected for variations in the age, sex, racial, and urban-rural composition of populations. Underlining the importance of using relevant bases in calculating rates, Sutherland used the example of an increase in convictions for violations of motor vehicle legislation in Michigan from 1,566 in 1912-13 to 27,794 in 1931-32. He noted that the number of cars in Michigan had increased from 54,366 to 1,230,980; using the number of motor vehicles instead of population figures for calculating the conviction rate for these offenses indicated that there had in fact been a significant decline in such convictions over the 20-year period.

Sutherland (1947) also noted that the way the crime index was constructed could result in misleading interpretations of changes in crime over time. Focusing on homicide, he noted that the principal increase in such offenses was primarily due to homicides by negligence, especially in the form of killings related to automobile use. He asserted that the number of homicides by negligence was approximately equal to the number of other criminal homicides. Because the total number of homicides had decreased while the proportion of homicides due to negligence had increased from the 1930s to the 1940s, it followed that the crime of murder had decreased significantly.

Sutherland (1940) is best known for his assertion that explanations of crime were invalid because the official statistics they were based on did not include "white-collar criminals" (p. 4). Defining white-collar crime as "a crime committed by a person of respectability and high social status in the course of his occupation" (p. 4), Sutherland included in this category members of the medical profession—who, he alleged, illegally sold narcotics, provided fraudulent reports and testimony in accident cases, and split fees—and disreputable business and professional men who were "quacks, ambulance-chasers, bucket-shop operators, dead-beats, and fly-by-night swindlers" (p. 4). Although Sutherland's identification of white-collar

crime did not lead to changes in how crime was officially counted in the United States, his assertion that crime was committed by individuals in all social classes inspired the work of labeling theorists such as Lemert (1951) and indirectly contributed to the development of self-report measures as alternatives to official measures of crime.

Some commentators also suggest that crimes committed by females, as well as by middle- and upper-class individuals, were underrepresented in official statistics. As early as 1932, a conference sponsored by the White House recognized that official juvenile court data underestimated the extent of female delinquency. One contributor to this debate was psychologist Otto Pollack. As Coleman and Monynihan (1996) note, "While Sutherland had seen the dark figure wearing a collar and tie, Otto Pollack claimed he had seen the dark figure wearing a dress" (p. 11). Pollack's (1951) study of female criminality, while presenting a controversial and rather misguided biopsychological theory of female crime, made an important contribution to the debate on the validity of crime statistics by focusing on the underrepresentation of females in official data. Pollack argued that female criminality was concealed by the underreporting of offenses committed by women, the lower detection rates of female offenders compared to male offenders, and by the greater leniency shown to women by officials in the criminal justice system, including police, prosecutors, and judges.

The recognition that there was a significant amount of crime that was not recorded in official statistics, whether committed by middle- and upper-class individuals or females, led to the development of the first important alternative measure of crime and delinquency, that is, self-report studies.

## Early Self-Report Studies

Discussions of the development of self-report methodology in criminology typically have focused on pioneering studies of researchers such as Wallerstein and Wyle, Porterfield, and Short and Nye. Important precursors to self-report studies, however, were more general developments in the science of survey research and specific refinements in questioning people about their "deviant behavior." Early forerunners of self-report studies of crime and deviance were, in fact, studies examining the sexual behaviors of adults and college youth, many of which focused on sexually deviant behaviors. These studies demonstrated that people would answer questions about private and potentially embarrassing behaviors. For example, in 1897, Havelock Ellis published his then controversial book

*Sexual Inversion*,<sup>1</sup> which was based on interviews with people in London and Paris regarding their sexual attitudes and behaviors and revealed "the vast, tangled jungle of sex activities which flourish in human bodies" (*Time* 1936).

One of the first studies of sexual behavior in the United States was conducted at the University of Missouri in the late 1920s as part of a course on the family. Included in the survey administered to students were questions about their view of sexual relations, both premarital and extramarital, whether fear, religious convictions, pride, or other forces inhibited their sexual desires; and opinions on divorce and alimony. Although the results of this study were apparently never published, the reactions to it are notable. Two of those who conducted the survey were dismissed from the University of Missouri, and another was suspended for one year. In justifying these actions, a member of the executive board of the university stated, "The time has come for a crusade against such discussions and such literature; for ridding our schools of those who cannot distinguish between legitimate research and cesspool delving and for barring from local libraries volumes which reek with sex appeal" (*Literary Digest* 1929).

As attitudes toward sexual issues became somewhat more liberal in the United States following World War II, several other studies of sexual behavior emerged. In a study of 613 students in Texas, Porterfield and Salley (1946) queried their subjects about their views and behaviors with respect to premarital sex, among other things. Focusing on gender differences in these behaviors, Porterfield and Salley noted that 58.5 percent of the "pre-college" men and 59 percent of the college men in their sample reported involvement in premarital sex, compared to only 1 in 137 (less than 1 percent) of the women.

The best known and most widely publicized of these surveys of sexual behavior and attitudes were those published by Alfred Kinsey and the Institute for Sex Research in the late 1940s and early 1950s. While teaching a course on marriage and the family at the University of Indiana, Kinsey, originally trained as a biologist at Harvard University, began to question his students about topics such as their age at first premarital intercourse, their frequency of sexual activity, and the number of sexual partners they had. He was eventually given funding to conduct more detailed studies of sexual behavior by the Committee for Research in the Problems in Sex, a Rockefeller-funded grant-giving body that operated under the umbrella of the National Research Council. Operating from the premise that "there cannot be sound clinical practice, or sound planning of sex laws, until we understand more adequately the mammalian origins of human sexual be-

havior" (Kinsey et al. 1953:8), Kinsey's project focused on examining sexual behaviors such as masturbation, heterosexual petting, premarital sexual intercourse, homosexual contact, and animal contact (see Exhibit 2.4). The project resulted in interviews with approximately 18,000 males and females, ranging in age from 2 to 90 years old, and including people from a variety of educational, occupational, and religious backgrounds.

Kinsey (Kinsey et al. 1953) believed that the best method for obtaining information on these issues was to conduct personal interviews<sup>2</sup> with subjects. "We have elected to use personal interviews rather than questionnaires because we believe that face-to-face interviews are better adapted for obtaining such personal and confidential material as may appear in a sex history" (p. 58). He felt that it was easier for interviewers to establish rapport with subjects through this method and that personal interviews made it possible to adapt the wording of each question into the vocabulary and experience of each subject. Recognizing that respondents would be more truthful if they were assured of the confidentiality of their answers, Kinsey and his interviewers recorded answers on special sheets printed with a grid. Kinsey informed his respondents that the information was being recorded using unintelligible codes that only he and his two colleagues would be able to understand. On the other hand, he recognized that respondents might lie in personal interviews, so he embedded a number of checks into his interview schedules to detect individuals who were not being truthful. If contradictions in answers were revealed, then subjects were asked to explain them. If they refused to do so, the interview was terminated, and the information from it was not used (Bullough 1998). Perhaps somewhat unbelievably in the context of the questions being asked, Kinsey (Kinsey et al. 1953) claimed that "unlike the experience of those engaged in public opinion and some other surveys, we find no difficulty in getting our subjects to answer all of the questions in an interview. In the course of 14 years, there have not been more than half a dozen subjects who have refused to complete the records after they had once agreed to be interviewed" (p. 45).

The average interview encompassed approximately 300 questions and required between one and one-half to two hours, but for some respondents who had extensive sexual experience, the number of questions extended to 500 or more. Kinsey and his associates were aware of reliability and validity problems in their data, and used a number of techniques to address these issues. "Retakes" or reinterviews, in which the same questions were asked of respondents, were conducted with 124 males and 195 females 18 months (in most cases) after the original interview had taken place. For the adult females in the sample, the retakes modified the lifetime incidence of re-

**Exhibit 2.4.** Comparison of Male and Female Sexual Behavior: Selected Variables (Percentages)

<i>Activity</i>	<i>Males</i>	<i>Females</i>
Lifetime incidence of masturbation	93	62
Premarital sexual intercourse	84	50
Erotic responses to same sex	50	28
Sexual contacts with animals	30	15

SOURCE: Adapted from Kinsey et al. (1953).

ported sexual behavior by less than 2 percent; for adult males, there was no activity for which the retakes modified the incidences calculated on the original histories by as much as 3 percent. However, when examining reported frequencies of sexual behavior, there was less reliability. On seven out of the nine items reported by females and on eight out of the nine items reported by males, fewer than 70 percent of the subjects provided identical responses. As an additional check on reliability, Kinsey examined data from 706 pairs of spouses in the sample. He found that the number of identical responses ranged from 39 percent on the maximum frequency of coitus in any single week of the marriage to 99 percent on the use of the male superior position during coitus.

One of the standard methods of organizing data in this period was to create seven-point scales to classify behaviors, and Kinsey used a similar scale to classify individuals as homosexual or heterosexual. He did not trust individuals' self-classification as homosexual or heterosexual, so his only objective indicator was the kind of sexual activities that resulted in the respondent experiencing an orgasm. Although the use of this measure indicated that most in his sample were exclusively heterosexual, it also implied that homosexuality was just another form of sexual activity—a revolutionary assertion in this period, as well as the one that resulted in the most serious attacks on Kinsey and his data. Kinsey's other highly controversial findings challenged prevailing beliefs about the asexuality of women. He found that 40 percent of the females he interviewed had experienced orgasm within the first few months of marriage, 67 percent by the first six months, and 75 percent by the end of the first year. However, he also reported instances in which women failed to reach orgasm after 20 years of marriage (Bullough 1998).

Although critics have questioned several of Kinsey's methods and theories, to suggest that Kinsey's studies were revolutionary is by no means

overstating the case. His 1948 book, *Sexual Behavior in the American Male*, which sold for \$6.50 and was more than 800 pages long, was published by W. B. Saunders, a respected publishing company that specialized in medical texts. Due to the anticipated popular appeal of the book, the company ordered 25,000 copies, rather than the usual 2,000 or 3,000, and these quickly disappeared (Schwarz 1997).

Kinsey's findings that premarital and extramarital relations, homosexuality, oral sex, masturbation, and a host of other sexual practices were far more common than most people believed were groundbreaking. However, it was his method of data collection that is more important for our purposes. The Kinsey studies were among the first to show that people would report on their "deviant" activities. Other studies, whose researchers were contemporaries of Kinsey, were revealing a similar willingness to report socially "questionable" behaviors via self-administered questionnaires.

One example is the large and sophisticated study of college students' drinking behavior conducted in the late 1940s by researchers at Yale University's Center of Alcohol Studies (Straus and Bacon 1953). The study covered 27 colleges in the United States, which were selected to represent a number of different types of institutions, including public, private, and sectarian institutions; coeducational, exclusively male and female, white and black institutions; urban and rural; those with large and small enrollments; and institutions in different regions of the country. The researchers administered questionnaires to a total of 17,000 college students, and of the 16,300 who filled out the questionnaire, 96.6 percent were used in the analysis.

A few months after their survey operations had begun, Straus and Bacon issued a press release announcing the survey and describing its procedures and purposes. They noted how the popular media's reaction to this study tended to trivialize their efforts. The study was referred to as "Booze Kinsey," and an article in the New Haven, Connecticut, *Journal Courier* in 1949 suggested, "Yale, for some odd reason (maybe they haven't got enough work to do in New Haven), would like to amass a flock of statistics. The snoopers would like to know if rich kids drink more than poor kids, or if the sons of teetotalers lush it up more than the scions of soaks. . . . Every so often I despair of the work that Satan finds for the idle professional hands to do" (as cited in Straus and Bacon 1953:42-3). The researchers asserted, however, that it was important to conduct this study "against the background of stereotypes, conflicts, problems, and changing patterns of drinking and control" (Straus and Bacon 1953:45).

The questionnaire used by Straus and Bacon contained items about, among other things, students' religious affiliation, family characteristics,

frequency of drinking, age when first "tight," "times high, tight, drunk, and passed out," and their opinions about the association between drinking and sexual behavior.

Straus and Bacon (1953) were aware of the potential reliability and validity problems in asking students about their alcohol consumption. Similar to Kinsey, they engaged in extensive checks of each questionnaire to eliminate responses that indicated "insincerity" or inconsistency. They noted that "about 100 students made humorous or sarcastic comments. Most of the latter were male students from one school where a member of the faculty assisting with the distribution of the forms invited attempts at humor with a joking remark at the start" (p. 4). They also realized that there would be variation in the students' knowledge of some issues such as family income and the drinking practices of their parents. Answers to other questions about so-called "measurable facts," such as frequency of drinking, amounts consumed on each occasion or number of times intoxicated, were dependent on memory and other factors in perception that could vary significantly from individual to individual.

Although there is not sufficient space here to enter into a detailed discussion of their results, several of the findings from this study are worth noting. Straus and Bacon found that 20 percent of the males and 39 percent of the females identified themselves as abstainers from alcohol. Male students were much more likely to have become tight than females (see Exhibit 2.5). Consumption of alcohol for most of the students took place in homes or public places such as restaurants, taverns, bars, or night clubs, as opposed to college dorm rooms and other on-campus sites. Beer was the most common alcoholic drink consumed by males, and wine was more commonly consumed by females. Straus and Bacon also noted that most students began drinking before they entered college. Of those who drank, approximately half had begun drinking by the age of 17. The reasons for drinking varied somewhat according to the type of beverage consumed, although these were generally related to issues of sociability.

As survey methodology in general, and self-report methodology in particular, progressed in sophistication and acceptability, researchers began to focus more on the measurement of criminal and deviant behavior. One of the pioneering studies in this genre was that of Porterfield (1946), who compared the self-reported delinquent behavior of 337 college students with that of 2,049 "alleged delinquents" who had appeared in the juvenile court in Fort Worth, Texas. Porterfield noted that those in the officially delinquent sample had been charged with a total of 55 specific offenses, ranging from "shouting spit wads at a wrestling match" to murder. How-

**Exhibit 2.5. College Students' Drinking (1953): First Drink Before or After Entering College (Percentages)**

<u>First Drink</u>	<u>Men</u>	<u>Women</u>
Before entering college	79	65
After entering college	21	35

**Male Students Ascribing Importance to Each of 12 Reasons for Drinking, by Type of Beverage Most Frequently Used**  
(Percentage ascribing some or considerable importance to each reason)

<u>Reason for Drinking</u>	<u>Beer</u>	<u>Wine</u>	<u>Spirits</u>
To get along better on dates	36	15	36
To relieve fatigue or tension	56	47	55
To be gay	66	41	63
To relieve illness	26	26	28
To comply with custom	65	53	66
Because of enjoyment of taste	77	65	72
In order not to be shy	27	26	24
As an aid in meeting crises	8	12	12
For a sense of well-being	18	21	28
As an aid to forgetting disappointments	29	18	23
To get high	53	24	42
To get drunk	16	9	17

**Number of Times Student Drinkers Have Become Tight (Percentages)**

<u>Occasions</u>	<u>Men</u>	<u>Women</u>
Never	20	51
1-5 times	25	32
6-15 times	18	9
16-50 times	17	4
51-100 times	5	0
100 or more	4	—
Have been tight, frequency not stated	11	4

SOURCE: Adapted from Straus and Bacon, *Drinking in College*, Copyright © 1953. Used by permission of Yale University Press.

ever, the study also revealed a significant amount of delinquent activity on the part of the sample of college students. "One well-adjusted ministerial student said he had indulged in 27 of the 55 offenses," and a few of the college students allegedly even confessed to committing murder. Presaging the findings of self-reported studies conducted in the 1960s and later years,

however, Porterfield acknowledged that although the offenses of the college students were as serious as those committed by the official delinquents, they were probably not committed as frequently. Porterfield (1946) surmised that the official delinquents had been labeled as such due to inherent biases in the operations of the criminal justice system and characterized the juvenile delinquent as "a friendless young person who does not live in a good home or in a college dormitory . . . but who has offended some part of a rather peevish and irresponsible community, and been charged with the necessity for being responsible and other than peevish himself" (p. 205).

Murphy, Shirley, and Witmer (1946) similarly questioned the validity of official data on juvenile delinquency and took as their point of departure the fact that a considerable number of juveniles who violated the law did not appear in official criminal statistics. Studying a group of 114 officially delinquent boys, they grouped delinquency into three categories of seriousness: (1) violations of city ordinances, including such "offenses" as shining shoes or vending without a license, street ball playing, hopping streetcars, swimming or fishing or both in forbidden places, and violating curfew laws; (2) minor offenses, involving behaviors such as truancy, petty stealing, trespassing, running away from home, and sneaking into movies; and (3) more serious offenses, involving acts such as breaking and entering, larceny-theft, assault, drunkenness, and sex offenses.

To measure the extent to which their subjects engaged in these offenses, Murphy, Shirley, and Witmer (1946) engaged in a group consultation with youth case workers and reviewed the youth's self-reported delinquency to determine whether he engaged in these offenses rarely (denoting a frequency span of from 1 to 3 offenses per year), occasionally (4 to 9 offenses per year), or frequently (more than 10 times per year). Based on these data, they estimated that the 114 boys had committed a minimum of 6,416 infractions of the law during a five-year period (covering the ages between 11 and 16). Supporting their contention that there existed a large dark figure of unrecorded delinquency, the researchers noted that only 95 of these violations became a matter of "official complaint." The authors noted the implications of their study for official measures of crime: "So frequent are the misdeeds of youth that even a moderate increase in the amount of attention paid to it by law enforcement authorities could create a semblance of a 'delinquency wave' without there being the slightest change in adolescent behavior" (p. 696).

Wallerstein and Wyle (1947) conducted a self-report study of delinquent behavior using a sample of 1,698 adult men and women in New York City,

focusing on the delinquent behavior these subjects had committed before they reached the age of 16. The mailed questionnaire listed 49 separate offenses, and 99 percent of their sample reported committing at least 1 delinquent act. Men admitted to an average of 18 crimes and women to an average of 11. Perhaps even more surprising, 64 percent of the males and 29 percent of the women in the sample reported committing at least 1 of the 14 felonies included in the list of offenses.

Further developments in self-report methodology were associated with the work of James F. Short Jr. and Ivan Nye (Nye and Short 1957; Nye, Short, and Olson 1958; Short and Nye 1957-58). These researchers began their project with a critique of existing criminological theories that had examined the relationship between juvenile delinquency and socioeconomic status based on official data. Similar to the earlier observations of Robison (1936), they asserted that studies using court records, police files, and other official measures were adequate for measuring official delinquency but were unreliable as indexes of delinquent behavior in the general population.

In one of their studies, Short and Nye (1957-58) administered a questionnaire to samples of the general population of adolescents who were attending school and to a sample of "official delinquents" who were institutionalized in training schools. The questionnaire included a total of 23 items, and from these, they created a delinquency index consisting of 7 items. Respondents were asked if they had committed the following acts since beginning grade school: (1) defied parents' authority (to their face), (2) taken little things (worth less than \$2) you didn't want or need, (3) driven a car without a driver's license or permit, (4) skipped school without a legitimate excuse, (5) bought or drank beer, wine, or liquor (including drinking at home), (6) purposely damaged or destroyed public or private property that did not belong to you, and (7) had sexual relations with a person of the opposite sex. For each one of these items, involvement in the behavior was divided into four categories: (1) did not commit the act, (2) committed the act once or twice, (3) committed the act several times, and (4) committed the act very often.

In analyses of their data, Nye, Short, and Olson (1958) found only weak relationships between delinquency and social class. For example, they noted that heterosexual relations were most frequently engaged in by lower-class boys in their sample, but purposely damaging or destroying property was committed most frequently by upper-class boys and girls. However, the researchers did recognize the limitations of their study in that not all adolescents were in school. School dropouts, who would not have been questioned in their surveys, may well have been more delinquent than

those in school and may have been disproportionately concentrated in the lower class.

Short and Nye were also aware of potential reliability and validity problems in measuring crime and delinquency through self-reports. They were especially concerned that the institutionalized delinquents they included in their samples would attempt to manipulate the interview situation. To assess the extent of lying on the questionnaires, Short and Nye (1957-58) included a number of trap questions, which were designed to identify both overreporting and underreporting of delinquency. They argued that if respondents indicated they had never told a lie and had never disobeyed their parents, they were presenting an "over-conformist" image; respondents who were identified as such were excluded from the analyses. On the other hand, some noninstitutionalized respondents in their study reported that they had committed all the offenses on the checklist; such individuals were also excluded because it was believed that "such a person would not be at large." Such attention to the potential methodological weaknesses in these early self-report studies has been of considerable value to those who continue to conduct research using this methodology.

Despite the inclusion of relatively trivial "offenses," Short and Nye's work and other early self-report studies were important both methodologically and substantively. They demonstrated that people would report having committed delinquent acts and that the alleged negative relationship between social class and delinquent behavior was not as strong as extant criminological theory purported it to be. As Hindelang, Hirschi, and Weis (1981) noted, "Much like the Kinsey studies before them, the Short/Nye studies revolutionized ideas about the feasibility of using survey procedures with a hitherto taboo topic. They also eventually led to a revolution in thinking about the substance of the phenomenon itself" (p. 23).

Erickson and Empey (1963) also took issue with the use of official data to measure delinquency. They argued that such behavior was not an attribute but was instead a phenomenon that was distributed along one or more continua. Their study involved personal interviews with males aged 15 to 17 in Utah and included four subsamples: (1) 50 high school boys who had never appeared in court, (2) 30 boys who had been to court once, (3) 50 repeat offenders who were on probation, and (4) 50 incarcerated offenders. Erickson and Empey suggested that the face-to-face interview method was the most effective in uncovering delinquent behavior because it allowed interviewers to provide more complete and reliable data, especially given a lack of literacy among some of their subjects. They also asserted that this method allowed for more accurate estimates of the frequency of involvement in delinquent

behavior than the standard method of having subjects respond to predetermined categories such as "none," "a few," or "a great many times." However, interviewers in the Erickson and Empey study encountered problems because some respondents were reluctant to reveal their involvement in offenses and the fact that some more habitual offenders had committed the offenses so frequently that they could not accurately estimate the number of times.

Similar to the findings of previous self-report studies, Erickson and Empey (1963) noted that the number of violations admitted to by their respondents was "tremendous." Three types of offenses were most common: theft (a total of 24,199 offenses), traffic violations (23,946), and the purchase and drinking of alcohol (21,698). In more than 90 percent of the cases, these offenses were undetected and not acted on by any official agency. Although the amount of hidden delinquency in this sample was significant, they found that boys who had been labeled as officially delinquent had committed a far greater number of delinquent acts than those who had not been so labeled.

Gold's (1966) self-report study began by noting that in one Michigan city, boys who lived in the poorer sections of town and were apprehended by police were four to five times more likely to be officially labeled delinquent than boys from the wealthier sections of town who were involved in the commission of the same types of offenses. This study relied on interviews with 522 13- to 16-year-old boys and girls living in the school district, who were matched with interviewers on the basis of race and sex. The self-report instrument contained a total of 51 questions that asked youth about offenses they had committed during the previous three years. Aware of the possibility that his respondents might conceal their delinquent activities, Gold also interviewed a "criterion" group of 125 young people for whom he already had collected reliable information on their delinquency from official data. In addition, he interviewed peers of the respondents in order to gather independent information on delinquent acts that they had witnessed or had been described to them by the respondents.

Using these data, Gold concluded that 72 percent of the youth provided self-reported delinquency information that was consistent with information provided by the informants, 17 percent were identified as "outright concealers," and the remaining 11 percent were "questionables." More important for measures of delinquency and for theories of delinquent behavior, and in contrast to several of the earlier self-report studies, Gold (1966) concluded that crime was, in fact, inversely related to social status. The lower-status young people in this study were found to commit delinquent

acts more frequently than higher-status adolescents. However, this relationship existed only among boys.

An additional development with respect to reliability checks of self-reports of deviant behavior was associated with the work of Clark and Tift (1966). These researchers conducted a study of 45 white males enrolled in a sociology course at a midwestern U.S. university in which respondents were asked to report their frequency of commission of a number of delinquent behaviors. The subjects were also given a polygraph test to check the veracity of their reports. Clark and Tift found that self-reports of delinquent behavior were accurate when a wide range of behaviors was considered simultaneously, but that there was differential validity on specific questionnaire items.

The first National Youth Survey (NYS), conducted in 1967 (Williams and Gold 1972), drew on interviews and official records of 847 13- to 16-year-old boys and girls. In this study, respondents were given the following instructions: "Here is a set of things other kids have told us they have done. Which of them have you done in the past three years, whether you were caught or not?" Respondents were asked to report whether they had never engaged in the activity, whether they had done it just once in the previous three years, or whether they had done it more than once. Williams and Gold found that 88 percent of the teenagers they interviewed confessed to committing at least one chargeable offense in the three years prior to the interview. They concluded that "if the authorities were omniscient and technically zealous, a large majority of American 13- to 16-year-olds would be labeled juvenile delinquents" (p. 213).

The Williams and Gold study was one of the first to focus in some detail on racial differences in self-reported delinquency. The researchers discovered that black females were not more frequently or seriously delinquent than white females, and black boys were not more frequently delinquent than white boys. However, black males reported involvement in more serious forms of delinquency than whites. For example, when involved in theft, blacks stole more expensive items, and when involved in assaults, they tended to inflict more serious injury.

## Victimization Surveys

Another method of measuring crime that arose in response to the limitations of official data is the victim survey. Although not commonly identified as such in the literature, one of the first studies of this nature was con-



ducted by Fitzpatrick and Kanin (1957), who investigated sexual aggressiveness in dating relationships on a university campus. Questionnaires were distributed to 291 females in 22 university classes, asking them about their experiences with males in dating relationships. Of the respondents, 55.7 percent reported that they had been "offended at least once during the academic year at some level of erotic intimacy," 20.9 percent had experienced forceful attempts at intercourse, and 6.2 percent had experienced "aggressively forceful attempts at sexual intercourse in the course of which menacing threats or coercive infliction of pain were employed."

The primary motivation for the development of comprehensive national surveys of crime victims in the United States was the recognition of the limitations in official measures of crime. In response to apparently escalating crime rates and urban unrest in several U.S. cities in the late 1960s, the President's Commission on Law Enforcement and Administration of Justice (1968) was impaneled in 1965 to develop policy and recommendations concerning the crime problem. The President's Commission noted that "one of the most neglected subjects in the study of crime is its victims" and found that much of the information needed to formulate policy recommendations with respect to the rising crime problem was not available. The President's Commission also noted that official statistics were problematic because many crimes were not reported to the police, and a number of administrative and organizational factors were believed to affect the reporting of these statistics in particular jurisdictions.

Both the deficiencies of official data and developments in the methodology of large-scale sample surveys provided the President's Commission with the impetus for developing victimization surveys. The initial efforts involved three separate studies: a pilot study in Washington, D.C., a second-stage study in three U.S. cities, and a national survey.

The Washington, D.C., pilot study was conducted during the spring of 1966. Working from a probability sample of homes in three police precincts, 511 interviews were completed with individuals who were asked to report whether they had been a victim of a list of crimes since New Year's Day, 1965. This study contributed to the methodological sophistication of subsequent victimization studies. It also demonstrated that household surveys would provide a different picture of crime than that derived from police statistics. Depending on the type of crime, the pilot study found that there were from 3 to 10 times as many criminal incidents reported by victims than were recorded in official data (President's Commission 1968).

The second-stage study was designed to elicit criminal victimizations experienced by businesses and organizations in selected high-crime areas in

Boston, Chicago, and Washington, D.C., and to measure household victimizations among residents in Boston and Chicago. This study similarly found that there was much more crime than was reported in official statistics.

The third victimization study sponsored by the President's Commission was a national survey conducted by the National Opinion Research Center, in which one respondent was interviewed in each of 10,000 households. This study revealed that approximately twice as many incidents of personal violence, and more than twice as many individual property victimizations, were estimated to have occurred than were recorded in the UCR. These data indicated that although many people experienced crime, many chose not to report it to the police. This survey also examined nonreporting of victimizations and the reasons individuals did not report offenses they had experienced. Nonreporting was found to vary across offenses, ranging from a high of 90 percent for consumer frauds to a low of 11 percent for automobile thefts. Most of those who did not report their victimizations to the police felt either that the incident was private or that the police could not do anything about the offense (President's Commission 1968).

The National Opinion Research Center victimization study reported a number of findings that were of interest to criminologists and policy makers. For example, the highest rates of victimization were found in the lower-income groups, nonwhites were victimized disproportionately by all index crimes except theft over \$50, and the rates of victimization for men were almost three times higher than those for women.

Early reports also recognized the need for caution in the interpretation of victimization data, however, and pointed to a number of methodological problems. For example, in noting the higher rates of burglary, larceny, and auto theft committed against men, the President's Commission suggested that this was primarily an artifact of the survey methodology, whereby offenses committed against the household were assigned to the head of the household, which, in most cases, was a male. There was also recognition of the problems of telescoping and recency effects in the interview data. For instance, Biderman (1967) noted that the distribution of incidents reported for the national survey had a "bulge" at the beginning of the 12-month period for which the respondents were asked to report, as well as a larger bulge at the recent end of the distribution. These bulges suggested that respondents were remembering some crimes that they had been the victims of before the 12-month reference period as well as being more likely to recall crimes that they had experienced recently. The studies also revealed that people interviewed about crimes affecting their households mentioned incidents they had experienced personally in considerable disproportion to

incidents affecting others who lived with them—that is, they were more likely to report their own victimization experiences as opposed to those of their family members. As Biderman (1967) noted, the inefficiency of asking about others' victimization experiences was underscored by the finding that there was not a positive relationship between the number of individuals in the household and the number of incidents reported by the respondent.

The initial victimization surveys also revealed an interesting finding with respect to the education of respondents: those who were college-educated reported more frequent victimization experiences than others. But as Biderman (1967) noted, this may have been an artifact related to the "productivity" of respondents: those with higher levels of education may have had better memories of events.

Although these initial victimization surveys were thus characterized by several methodological problems that will be explored in more detail in Chapter 5, they were important in establishing the victimization survey as an alternative to official measures of crime.

## Summary and Conclusions

Accurate counts of various social phenomena have been important to policy makers and citizens alike since at least early Greek and Roman times. Still, the century just passed may well be "the first measured century," because more and more numbers are used to characterize more and more facets of our lives. Historically, measures of crime and delinquency are among those that have the greatest potential for generating controversy.

As described in this chapter, the earliest measures of crime were derived from official statistics. Concern over the reliability and validity of official counts led to the development of self-report and victimization measures—the earliest examples of which likewise evidenced limitations. Nonetheless, those pioneering official, self-report, and victimization studies served as the basis for more recent developments and refinements in crime measurement.

Chapters 3, 4, and 5 critically examine each of these data sources—official statistics, self-report studies, and victimization surveys—in turn, with the goal of identifying its role in providing accurate measures of crime and delinquency.

## Notes

1. When Ellis published this book, he was arrested by the London police and all copies of the volume were seized. It was not until 1936 that individuals other than doctors and lawyers in the United States were allowed to legally purchase the book (*Time*, March 9, 1936).
2. Apparently some individuals, falsely claiming that they were part of the Kinsey research project, attempted to conduct telephone interviews with unsuspecting subjects. "At many points in the United States within the past five years, there have been imposters who posed as interviewers connected with the present project. In most instances they operated through telephone calls" (Kinsey et al., 1953:58). In response to this problem, Kinsey notes, "It should be understood that our staff never conducts interviews over the telephone."