

POLICY AND PERSPECTIVES

*Editor's note: In Fall 1995, the Journal published an article by Professors Gary Kleck and Marc Gertz, *Armed Resistance to Crime: The Prevalence and Nature of Self-Defense with a Gun*, 86 J. CRIM. & CRIMINOLOGY 150 (1995). As part of its Policy and Perspectives section, the Journal now publishes the views of Professor David Hemenway on the Kleck-Gertz paper, a reply by Professors Kleck and Gertz, and the views of Professor Tom Smith on both the Hemenway and Kleck-Gertz papers. As always, the views expressed here are those of the authors.*

SURVEY RESEARCH AND SELF-DEFENSE GUN USE: AN EXPLANATION OF EXTREME OVERESTIMATES

DAVID HEMENWAY*

I. INTRODUCTION AND SUMMARY

Gary Kleck and Marc Gertz conducted a survey of civilian defensive gun use in 1992. In 1993, Kleck began publicizing the estimate that civilians use guns in self-defense against offenders up to 2.5 million times each year.¹ This figure has been widely used by the National Rifle Association and by gun advocates. It is also often cited in the media² and even in Congress.³ The Kleck and Gertz (K-G) paper has now been published.⁴ It is clear, however, that its conclusions cannot be accepted as valid.

* Professor of Health Policy, Harvard School of Public Health. The author gives special thanks to Dimitrios Trichopoulos, Sue Swartz, Phil Cook, Deb Azrael, Eric Rimm, Frank Zimring, Jon Vernick, Daniel Webster, Tom Cole, and Al Biderman.

¹ J.N. Schulman, *Guns, Crimes and Self-defense*, ORANGE COUNTY REG., Sept. 19, 1993, at 3 (editorial section); *Talk of the Nation: Interview with Gary Kleck and Arthur Kellermann* (National Public Radio broadcast, Aug. 8, 1994).

² See Gordon Witkin, *The Great Debate: Should You Own a Gun?*, U.S. NEWS & WORLD REP., Aug. 15, 1994, at 24-31.

³ Cong. Res. Service, *Gun Control*, in CRS ISSUE BRIEF (Keith Bea ed., 1995).

⁴ Gary Kleck & Marc Gertz, *Armed Resistance to Crime: The Prevalence and Nature of Self-defense with a Gun*, 86 J. CRIM. L. & CRIMINOLOGY 150 (1995).

Two aspects of the K-G survey combine to create severe misestimation. The first is the likelihood of positive social desirability response, sometimes referred to as personal presentation bias. An individual who purchases a gun for self-defense and then uses it successfully to ward off a criminal is displaying the wisdom of his precautions and his capability in protecting himself, his loved ones, and his property. His action is to be commended and admired.

Some positive social desirability response bias, by itself, might not lead to serious overestimation. However, combined with a second aspect of the survey—the attempt to estimate a very rare event—it does. The search for a “needle in a haystack” has major methodological dangers, especially where researchers try to extrapolate the findings to society as a whole.

Until the K-G study, no one had estimated that even as many as 1% of adult civilians had used a gun in self-defense in the past year. Nevertheless, assume that the actual incidence is 1%. On average, for every 100 individuals asked a “Yes/No” question about the event, ninety-nine respondents will have a chance to be misclassified as a false positive. In ninety-nine answers there is the possibility of positive social desirability response bias. However, on average only one respondent—the one who actually did use a gun in self defense—could possibly be misclassified as a false negative (e.g., if she forgot about the event). Even if the chance of forgetting is high, as long as there is any possibility of positive response bias, it is very likely that the survey finding will be an overestimate.

The fact that the survey is trying to estimate a low probability event also means that a small percentage bias, when extrapolated, can lead to extreme overestimates. Consider a survey finding which contains a 1% overestimate of positive responses. If the true incidence of the event is 60%, estimating it at 61% would not be a problem. But if the true incidence is 1%, measuring it as 2% would be a doubling of the true rate; and if the true incidence is 0.1%, measuring it at 1.1% would be an eleven-fold overestimate.

The K-G survey design contains a huge overestimation bias. The authors do little to reduce the bias or to validate their findings by external measures. All checks for external validity of the Kleck-Gertz finding confirm that their estimate is highly exaggerated.

II. BACKGROUND

Previous data on self-defense gun use came from two sources—the large National Crime Victimization surveys (NCVS), and smaller private surveys (principally random-digit-dial telephone surveys). These two sources produce markedly different results.

The NCVS employs a multistage design with a probability sample of some 50,000 housing units in the United States (e.g., in 1994 there were 47,600 housing units and 90,560 persons).⁵ The survey is conducted by the Census Bureau for the Bureau of Justice Statistics. Housing units remain in the NCVS for three years and residents are interviewed at six month intervals. Initial surveys are in-person, while subsequent ones are typically by telephone. Respondents who report a threatened, attempted or completed victimization for six crimes—rape, robbery, assault, burglary, non-business larceny and motor vehicle theft—are asked detailed questions about the incident.

NCVS results indicate that, nationally, victims use guns against offenders approximately 65,000 times per year.⁶ Kleck believes people under-report to the government NCVS interviewers, especially since the surveys are not anonymous. He also finds fault with the NCVS survey for asking about self-defense gun use only for individuals who have been victimized.⁷ Interestingly, it is this latter feature of the NCVS which dramatically reduces the overestimation bias found in the private surveys.

Based on eight national surveys, undertaken between 1976 and 1990, Kleck estimates that guns are used approximately 700,000 times per year in self-defense.⁸ However, all eight surveys have very serious limitations. Compared to the NCVS, the sample size of each of these surveys is small (600 - 1500) and interviewers typically asked only one vague question about gun use in self-defense (e.g., "Have you used a gun in self-defense in the previous five years?") with no follow-up questions.⁹ Only one of the surveys meets the minimum criteria of drawing from a representative national population, asking about a specific time frame, distinguishing civilian use from military or police uses, and distinguishing uses against humans from uses against animals.¹⁰

A review of Kleck's analysis argued that "Kleck's conclusions rest on limited data and strong assumptions. Small changes in the procedure produce large differences in the findings. The estimates are

⁵ UNITED STATES DEP'T OF JUSTICE, CRIMINAL VICTIMIZATION IN THE UNITED STATES, 1994: A NATIONAL CRIME VICTIMIZATION SURVEY REPORT (1997).

⁶ Philip J. Cook, *The Technology of Personal Violence*, in 14 CRIME AND JUSTICE: AN ANNUAL REVIEW OF RESEARCH (Michael Tonry ed., 1991); see also David McDowall & Brian Wiersma, *The Incidence of Defensive Firearm Use by U.S. Crime Victims, 1987 through 1990*, 84 AM. J. PUBLIC HEALTH 1982-84 (1994); Philip J. Cook & Jens Ludwig, *Guns in America: Results of a Comprehensive National Survey on Firearms Ownership and Use*, Police Foundation (1996).

⁷ GARY KLECK, POINT BLANK: GUNS AND VIOLENCE IN AMERICA 110 (1991).

⁸ See *id.* at 105-07.

⁹ Kleck & Gertz, *supra* note 4, at 157.

¹⁰ See *id.* at 155.

questionable, and it appears unwise to place much weight on them."¹¹ A National Research Council report also finds that Kleck's estimates appear exaggerated and says that it is almost certain that "some of what respondents designate as their own self-defense would be construed as aggression by others."¹²

III. THE KLECK-GERTZ SURVEY

In 1992, Kleck and Gertz conducted a national random-digit-dial survey of five thousand dwelling units, asking detailed questions about self-defense gun use.¹³ Their estimates of civilian self-defense gun use range from 1 million to 2.5 million times per year.¹⁴ The 2.5 million figure is the one they believe to be most accurate and the one Kleck has publicized, so that figure will be discussed in this paper.

K-G derive their 2.5 million estimate from the fact that 1.33% of the individuals surveyed reported that they themselves used a gun in self-defense during the past year;¹⁵ in other words, about 66 people out of 5000 reported such a use. Extrapolating the 1.33% figure to the entire population of almost 200 million adults gives 2.5 million uses.

Many problems exist with the survey conducted by Kleck and Gertz. A deficiency in their article is that they do not provide detailed information about their survey methodology or discuss its many limitations. For example, the survey was conducted by a small firm run by Professor Gertz. The interviewers presumably knew both the purpose of the survey and the staked-out position of the principal investigator regarding the expected results.

The article states that when a person answered, the interview was completed 61% of the time.¹⁶ But what happened when there was a busy signal, an answering machine or no answer? If no one was interviewed at a high percentage of the initially selected homes, the survey cannot be relied on to yield results representative of the population.

Interviewers do not appear to have questioned a random individual at a given telephone number, but rather asked to speak to the male head of the household.¹⁷ If that man was not at home, the caller

¹¹ David McDowall et al., *The Incidence of Civilian Defensive Firearm Use*, at 4 (University of Maryland Violence Group Discussion Paper, Nov. 10, 1992).

¹² UNDERSTANDING AND PREVENTING VIOLENCE 266 (Albert J. Reiss, Jr. & Jeffrey A. Roth, eds., 1992).

¹³ Kleck & Gertz, *supra* note 4, at 160-62, 166.

¹⁴ *Id.*, as reported in 184 tbl.2.

¹⁵ *Id.*, as reported in 184 tbl.2.

¹⁶ *Id.* at 161.

¹⁷ *Id.*

interviewed the adult who answered the phone.¹⁸ Although this approach is sometimes used in telephone surveys to reduce expense, it does not yield a representative sample of the population.

The 2.5 million estimate is based on individuals rather than households.¹⁹ But the survey is randomized by dwelling unit rather than by individual, so the findings cannot simply be extrapolated to the national population. Respondents who are the only adults in a household will receive too much weight.

K-G oversampled males and individuals from the South and West.²⁰ The reader is presented with weighted rather than actual data, yet the authors do not explain their weighting technique. K-G claim their weighted data provide representative information for the entire country,²¹ but they appear to have obtained various anomalous results. For example, they find that only 38% of households in the nation possess a gun, which is low, outside the range of all other national surveys.²² They find that only 8.9% of the adult population is black,²³ when 1992 Census data indicate that 12.5% of individuals were black.²⁴

The above limitations are serious. However, it is two other aspects of the survey that, when combined together, lead to an enormous overestimation of self-defense gun use: the fact that K-G are trying (1) to measure a very low probability event which (2) has positive social desirability response bias. The problem is one of misclassification.

IV. MISCLASSIFICATION IN SURVEYS GENERALLY

All surveys have problems with accuracy.²⁵ Incorrect classifications come from a wide variety of causes including misunderstanding, miscoding, misremembering, misinterpretation of events, mischief or downright mendacity.

Some percentage of answers to virtually all survey questions are incorrect. Respondents substantially over-report their seat belt use,²⁶

¹⁸ Florida State Univ. Poll, Research Network of Tallahassee, Nat'l Self-Defense Survey, at 1.

¹⁹ Kleck & Gertz, *supra* note 4, at 184 tbl.2.

²⁰ *Id.* at 161.

²¹ *Id.* at 160-61.

²² KLECK, *supra* note 7, at 51 tbl.2.2.

²³ Kleck & Gertz, *supra* note 4, 187 tbl.4.

²⁴ UNITED STATES DEP'T OF COMMERCE, BUREAU OF THE CENSUS, STATISTICAL ABSTRACT OF THE UNITED STATES 1996, 18 tbl.18.

²⁵ See generally Mark S. Litwin, HOW TO MEASURE SURVEY RELIABILITY AND VALIDITY (1995).

²⁶ Frederick M. Streff & Alexander C. Wagenaar, *Are There Really Shortcuts? Estimating Seat Belt Use with Self-report Measures*, 21 ACCIDENT ANALYSIS AND PREVENTION 509 (1989).

for example, and inaccurately report whether they voted.²⁷ Not all people are completely truthful when reporting about such mundane details as their age,²⁸ height, or weight.²⁹ A book on survey response validity characterizes as "quite high" accuracy rates of 83% to 98% to questions about possession of an automobile, a home, a driver's license or a library card.³⁰

Respondents who misreport are not necessarily deliberately lying; they may be shading the truth or simply perceive and present themselves in a slightly more favorable light than a purely objective observer would. In addition, some Americans may simply have a different perception of reality than most of us. For example, the best estimates are that in any recent six month period, one to two million Americans suffered from schizophrenia, one to two million suffered from antisocial personality disorder,³¹ and another two million suffered from Alzheimer's disease and other cases of severe dementia.³²

V. SURVEY BIAS TOWARD SUBSTANTIAL OVERESTIMATION

Self-report surveys tend to overestimate rare events which carry no social stigma, and such surveys can wildly overestimate rare events which have some social desirability.

The overestimation problem is probably best explained in the context of the screening of diseases. Consider Table 1 (at the end of this article), which describes the results of a screening test. The screen could be any kind of test, for example breast cancer, but in this context, let it be the response to the Yes/No question "Have you used a gun for self-defense in the past year?" The two Rows are the response (the screen) and the two Columns are the truth (the actual fact). Each respondent can be placed in one of the four categories. In Table 1:

a = the number of people answering yes who actually had a self-defensive gun use in the past year (true positive).

b = the number of people answering yes who did not have a self-

²⁷ See Hugh Perry & Helen Crossley, *Validity of Responses to Survey Questionnaires*, 14 PUB. OPINION Q. 61, 73 (1950).

²⁸ Charles N. Weaver & Carol L. Swanson, *Validity of Reported Date of Birth, Salary and Seniority*, 38 PUB. OPINION Q. 69, 73 (1974).

²⁹ Charles Cannell et al., *A Summary of Studies*, 69 VITAL AND HEALTH STAT. 2 (1977).

³⁰ See ELLEN J. WENTLAND & K.W. SMITH, SURVEY RESPONSES: AN EVALUATION OF THEIR VALIDITY 19 (1993).

³¹ See HANDBOOK ON MENTAL HEALTH POLICY IN THE UNITED STATES (David A. Rochefort ed., 1989); Dairrel A. Regier et al., *The De Facto U.S. Mental and Addictive Disorders Service System: Epidemiologic Catchment Area Prospective One Year Prevalence Rates of Disorders and Services*, 50 ARCHIVES OF GEN. PSYCHIATRY 85 (1993).

³² OFFICE OF TECHNOLOGY ASSESSMENT, U.S. CONGRESS, LOSING A MILLION MINDS: CONFRONTING THE TRAGEDY OF ALZHEIMER'S DISEASE AND OTHER DEMENTIAS (1987).

defensive gun use in the past year (false positive).

c = the number of people answering no who did in fact have a self-defensive gun use in the past year (false negative).

d = the number of people answering no who had no self-defensive gun use in the past year (true negative).

Epidemiologists have names for various ratios in this table. Sensitivity is defined as $a/(a + c)$, or the percentage of all true positives accurately detected by the screening test. Specificity is defined as $d/(b + d)$, or the percentage of all true negatives accurately detected by the screening test. Positive Predictive Value is defined as $a/(a + b)$, or the percentage of the screened positives who were truly positive.

The figure K-G are trying to determine is the true cumulative incidence of self-defense gun use over a period of one year, or $(a + c)/(a + b + c + d)$. But the figure they present is the one derived by the survey, or $(a + b)/(a + b + c + d)$. Should we expect the size of "b" and "c" to differ markedly in the K-G survey? We definitely should. A basic epidemiology text helps explain why:

The predictive value of a screening test is determined not only by factors that determine validity of the test itself (i.e., sensitivity and specificity), but also by the characteristics of the population . . . in particular the prevalence of preclinical disease. . . . For rare diseases, the major determinant of the predictive value positive is the prevalence of the preclinical disease in the screened population. No matter how specific the test, if the population is at low risk for having the disease, results that are positive will mostly be false positives.³³

With a huge number of actual negatives, virtually any screen or screening question will pick up a sizable absolute number of false positives. With few actual positives, it is impossible for a screen to pick up many false negatives. It follows that, for events with low incidence (with neither positive nor negative connotations), the estimated incidence will tend to be greater than the true incidence.

Some numerical examples illustrate this point. In the various versions of Table 2, we take the K-G survey results as given, that 66 respondents out of approximately 5,000 reported a self-defense gun incident in the past year. In Table 2A, assume that the misclassification error is extremely small, only 1%. In other words, assume that a random 1% of respondents are categorized incorrectly by the screening test. In that case, while the K-G estimate of defensive gun use is 1.33% or about 2.5 million uses per year, the true incidence of defensive gun use would be only 0.32% or about 600,000 uses per year. The reported 2.5 million uses would be a four-fold overestimate.

The key point is that the K-G estimates are extremely sensitive to

³³ CHARLES H. HENNEKENS & JULIE E. BURING, *EPIDEMIOLOGY IN MEDICINE* 337 (1987).

miniscule changes in the Specificity rate. In Table 2B, it is assumed that 1.3% of respondents are randomly misclassified, which means there are still extremely high 98.7% Sensitivity and 98.7% Specificity rates. In that case the truth would be that just 0.04% of individuals actually used a gun in self-defense in the previous year, or about 76,000 uses per year for the entire adult population. This result comes very close to the NCVS estimate; the K-G figure of 2.5 million uses would be a thirty-three-fold overestimate!

Whenever the incidence is low, the ratio of the reported incidence to the true incidence depends almost entirely on the Specificity of the test, and very little on the Sensitivity. Table 2C shows the results if 1.3% of true negatives are misclassified (98.7% Specificity rate), while 50.0% of true positives are misclassified (50% Sensitivity rate, or half of those with an actual self-defense gun use in the past year did not report it). Given the K-G survey results, the true number of self-defense gun uses per year would still be only 150,000. The K-G figure would be a seventeen-fold overestimate.

K-G do not believe that 80% or more of the respondents who say they used a gun in self-defense in the past year could be misclassified. Note, however, that the 66 individuals in their survey who report a self-defensive gun use in the past year have not been chosen randomly—these 66 are the 1.3% whose screen result (i.e., their survey response) has attracted our attention.

Similarly, an extremely accurate medical test, for a disease such as breast cancer, when performed on the general population of women, will yield a pool of individuals with positive test results, the vast majority of whom do *not* have cancer. Virtually no test or screen or question will classify everyone correctly, and when almost all individuals tested are actual negatives, you inevitably get a large number of false positives relative to the number of true positives.

VI. AN ANALOGY

Since a small percentage of people may report virtually anything on a telephone survey, there are serious risks of overestimation in using such surveys to measure rare events. The problem becomes particularly severe when the issue has even a remote possibility of positive social desirability response bias.

Consider the responses to a national random-digit-dial telephone survey of over 1500 adults conducted in May 1994 by ABC News and the Washington Post.³⁴ One question asked: "Have you yourself ever seen anything that you believe was a spacecraft from another planet?"

³⁴ ABC News/WASH. POST Poll, Roper Center for Public Opinion Res., May 1994.

Ten percent of respondents answered in the affirmative. These 150 individuals were then asked, "Have you personally ever been in contact with aliens from another planet or not?" and 6% answered "Yes."

By extrapolating to the national population, we might conclude that almost 20 million Americans have seen spacecraft from another planet, and over a million have been in personal contact with aliens from other planets. That more than a million Americans had contact with aliens would be incredible news—but not the kind actively publicized by reputable scientists. Yet the ABC News/WASHINGTON POST data on aliens are as good as or better than that from any of the thirteen surveys cited by K-G as supporting their conclusions about self-defense gun use.

VII. MISCLASSIFICATION IN THE KLECK-GERTZ SURVEY

Using a gun in self-defense, like having contact with an alien, is an interesting, potentially exciting event that might well be heroic. In the K-G survey, many of those who report a self-defense gun use apparently see themselves as quite heroic. Were we to accept their claims, people using guns in self-defense are saving about 400,000 people each year from being murdered. Yet most people do not have guns and there were only a total of 27,000 homicides in 1992.³⁵

Survey respondents, like most mortals, like to present themselves in the best light. Many respondents who claim to have had contact with alien life forms are probably not deliberately lying, but are putting an interesting interpretation on circumstances which were not clear cut. Similarly, many respondents who claim to have used a gun successfully in the past year may be unconsciously improving on the truth—e.g., on situations in which they were afraid, they retrieved a gun, and nothing bad happened. It would not be surprising if respondents tended to embellish their stories of potentially dangerous events which occurred many months in the past. Their replies to the questions about the benefits of their gun use and how many "bad guys" they shot support that expectation.

The likelihood of social desirability response bias (self-presentation bias) is clear. For example, many respondents own firearms for self-protection. The successful use of a gun in self-defense shows their foresight as well as their competence in protecting themselves. The vast majority of self-reported self-defense gun uses in the K-G study appear to have been successful.³⁶

³⁵ UNITED STATES DEP'T OF COMMERCE, BUREAU OF THE CENSUS, STATISTICAL ABSTRACT OF THE UNITED STATES 1995, 92 tbl.125.

³⁶ Kleck & Gertz, *supra* note 4, at 174.

Strategic reporting is another possible reason for response bias.³⁷ Some respondents were undoubtedly aware of the debate over the incidence and utility of gun use in self-defense. A few might actually deliberately lie on a telephone survey to help boost the numbers for the sake of their political beliefs concerning the dangers of gun control.

False positives can also come from external "telescoping," the reporting of events that actually occurred, but were outside the time frame in question.³⁸ Unlike the one-shot K-G survey, the NCVS questions the same household every six months.³⁹ Analysis of the NCVS unbounded (or first-time) panel in comparison with the second indicates a substantial amount of telescoping of criminal victimization.⁴⁰ Unbounded rates of reported victimization are typically 30% to 40% higher than bounded rates.⁴¹

K-G believe that most of their respondents who claim a self-defense gun use in the past year are reporting accurately. To obtain estimates similar to the NCVS surveys, they argue, "one would have to suppose that 29 out of every 30 people reporting a defensive gun use in the present survey was lying. There is no precedent in criminological survey research for such an enormous level of intentional and sustained fabrication."⁴² While K-G do not believe that 60 out of 5000 respondents in their own survey might be misclassified, they are quite willing to speculate that over 95% of the individuals who used their guns in self-defense in the past year deliberately lie to the NCVS surveyors. Were we to accept the 2.5 million figure as accurate, then 1,200 of the approximately 90,000 adults interviewed by NCVS in 1994 had a self-defense gun use. However, only about 34 report any such use. So, according to K-G, it appears that 1,166 out of 1,200 are not telling the truth. In addition, none of the 88,800 NCVS individuals who have not had a gun use are reporting having had one. There is certainly no precedent for this extreme pattern of lying. Yet K-G believe this pattern of lying occurs continuously on the semi-annual NCVS surveys.

K-G claim they have a large false negative rate, that many re-

³⁷ Philip J. Cook et al., *The Gun Debate's New Mythical Number: How Many Defensive Gun Uses Per Year?*, 16 J. POL'Y ANALYSIS & MGMT. 463-69 (1997).

³⁸ Norman Bradburn, *Response Effects*, in HANDBOOK OF SURVEY RESEARCH (P.H. Rossi et al. eds., 1989).

³⁹ UNITED STATES DEP'T OF JUSTICE, CRIMINAL VICTIMIZATION IN THE UNITED STATES, 1993: A NATIONAL CRIME VICTIMIZATION SURVEY REPORT (1996).

⁴⁰ David Cantor, *Substantive Implications of Longitudinal Design Features: the National Crime Survey as a Case Study*, in PANEL SURVEYS 25, 28 (D. Kasprzyk et al. eds., 1989).

⁴¹ *Id.* at 33.

⁴² Kleck & Gertz, *supra* note 4, at 179.

sponders who actually did use a gun in self-defense in the past year forgot to report it on their survey. K-G offer no compelling evidence as to why many people should forget such a recent memorable event. The crucial point, though, is that the Sensitivity rate (the problem of false negatives) is almost irrelevant concerning the accuracy of the K-G results. It is the Specificity rate (the problem of false positives) that really matters for rare events. When estimating very rare events for which there is a positive social desirability response bias, the number of false positives almost inevitably will be far larger than the number of false negatives.

VIII. RELIABILITY

Reliability in surveys refers to the reproducibility of results. K-G claim that we should accept their findings as accurate because their results are consistent with those of other private surveys while the NCVS estimates are not.

[O]ne might suppose that the gross inconsistency of the NCVS-based estimates with all other known estimates, each derived from sources with no known flaws even remotely substantial enough to account for nine-to-one, or more, discrepancies, would be sufficient to persuade any serious scholar that the NCVS estimates are unreliable.⁴³

The K-G arguments about the reliability of the private surveys compared to the NCVS surveys are misleading and largely irrelevant. Leaving aside the issue of whether or not the results of the private surveys are really consistent, three facts need to be understood: (1) if all these private surveys were combined, including the K-G survey, they would not be half the size of a single National Crime Victimization Survey; (2) the NCVS is not just one survey, but is given every six months. No one has ever claimed that the results from the various NCVS are inconsistent; and (3) consistency of findings is irrelevant when the methodology among all the private surveys is similar and when that methodology is biased toward substantially overestimating the event in question.

K-G argue that “[t]he strongest evidence that a measurement is inaccurate is that it is inconsistent with many other independent measurements or observations of the same phenomenon; indeed, some would argue that this is ultimately the *only* way of knowing that a measurement is wrong.”⁴⁴ However, reproducing the same result over and over with the same flawed measurement tool does not provide much evidence about anything. The best way of knowing that a measurement is wrong is that it does not correspond to reality. Thus, we

⁴³ *Id.* at 153.

⁴⁴ *Id.*

should not accept as fact that men have intercourse with women far more often than women have intercourse with men simply because surveys consistently report such a result.⁴⁵

IX. EXTERNAL VALIDITY

We expect that the number of false positives in the K-G survey will vastly exceed the number of false negatives. Therefore, if the survey findings are to be considered accurate, we need some strong evidence of external validity, rather than mere consistency with other surveys using the same methodology. However, all reality checks indicate that the K-G victimization and gun use results are highly exaggerated.

In his analyses, Kleck often uses the NCVS victimization data, which he sees as the gold standard for estimates of crime victimization in the United States. The NCVS "survey instrument has been carefully refined and evaluated over the years to do as good a job as possible in getting people to report illegal things which *other* people have done *to* them."⁴⁶

The K-G survey estimates of victimization rates are far higher than the NCVS victimization results, indicating a substantial problem of telescoping of events. For example, in the K-G survey, 2.5% of U.S. adults report being robbery victims in the past year. That corresponds to 5 million robberies. But NCVS results show only 1.2 million attempted or completed robberies in 1992.⁴⁷ The victimization rate was 0.6% for individuals 12 years and older, and only 1% of households were victimized.⁴⁸

Combining the K-G gun use estimates with the gold standard NCVS victimization rates leads to completely implausible conclusions. For example, K-G find that 34% of the time a gun was used in self-defense, the offender was committing a burglary. If we use their 2.5 million estimate, we would conclude that, in 1992, a gun was used by defenders for self-defense in approximately 845,000 burglaries. However, from the NCVS, we know that there were fewer than 6 million burglaries in 1992.⁴⁹ Over 55% of the time the residence was definitely unoccupied at the time of the burglary (in another 23% it was not known whether the dwellings was occupied or not). Only 22% of the time was someone certainly at home (1.3 million burglaries). Kleck accepts as valid the claim that the dwellings were occupied in

⁴⁵ R.C. Lewontin, *Sex, Lies and Social Science*, N.Y. REV. OF BOOKS, Apr. 20, 1995, at 29.

⁴⁶ Kleck & Gertz, *supra* note 4, at 156.

⁴⁷ BUREAU OF JUSTICE STATISTICS, U.S. DEP'T. OF JUSTICE, CRIMINAL VICTIMIZATION IN THE UNITED STATES, 1992 (1994).

⁴⁸ *Id.*

⁴⁹ *Id.*

only 9% of U.S. burglaries.⁵⁰ Since fewer than half of U.S. households have a firearm of any kind and since the victims in two-thirds of occupied dwelling were asleep, the K-G result asks us to believe that burglary victims in gun owning households use their guns in self-defense more than 100% of the time, even though most were initially asleep.

An analysis of Atlanta police department reports of home invasion crimes provides further information on self-defense gun use during burglaries.⁵¹ Examining a three-month period in 1994, researchers identified 198 cases of unwanted entry into a single-family dwelling while one or more individuals were present in the home.⁵² Most of these cases met the Uniform Crime Report criteria for burglary (the others were classified as robberies or assaults).⁵³ In 16% of the cases of home invasion, at least one of the offenders carried a firearm. In only three cases (1.5%) was a victim able to use a firearm in self-defense.⁵⁴ To accept the K-G results requires a belief that, of the 198 events, all the families with guns actually used them in self defense, but only three bothered to tell the police.

K-G find that guns were are in self-defense in 205,000 rapes and sexual assaults each year. The new NCVS, specifically revised to produce more accurate reporting of rape and sexual assault, estimates approximately 500,000 cases of rape and sexual assault in 1993. The NCVS data include not only rape and attempted rape, but also other sexual assault and verbal threats of sexual assault.⁵⁵ Some 150,000 of these cases involve individuals younger than 18 years of age, who would not be included in the K-G survey of adults. So if we believe the K-G results, women (the vast majority of whom do not own guns) defend themselves with guns in almost 40% of all sexual assaults.⁵⁶

K-G report that 207,000 times per year the gun defender thought he wounded or killed the offender.⁵⁷ However, only about 100,000 people are treated in emergency rooms each year for non-fatal firearm-related injuries;⁵⁸ almost all of these are victims of assault, suicide

⁵⁰ See KLECK, *supra* note 7, at 140.

⁵¹ See Arthur L. Kellermann et al., *Weapon Involvement in Home Invasion Crimes*, 273 JAMA 1759 (1995).

⁵² *Id.* at 1759-60.

⁵³ *Id.* at 1760.

⁵⁴ *Id.* at 1761.

⁵⁵ RONET BACHMAN & LINDA E. SALTZMAN, U.S. DEP'T. OF JUSTICE, VIOLENCE AGAINST WOMEN: ESTIMATES FROM THE REDESIGNED SURVEY (1995); BUREAU OF JUSTICE STATISTICS, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS, 320 tbl.3.1 (1994).

⁵⁶ David Hemenway, *Guns, Public Health and Public Safety*, in GUNS AND THE CONSTITUTION 49, 64 (Dennis A. Henigan et al. eds., 1995).

⁵⁷ Kleck & Gertz, *supra* note 4, tbl.3 extrapolated.

⁵⁸ See Joseph A. Annest et al., *National Estimates of Nonfatal Firearm-Related Injuries: Beyond the Tip of the Iceberg*, 273 JAMA 1749 (1995).

attempts and unintentional gun shootings rather than criminals shot by defenders.

K-G report that 392,000 times per year a gun defender thought that someone almost certainly would have been killed had the gun not been used; that another 355,000 times someone probably would have been killed; and another 405,000 times someone might have been killed if the gun had not been used for protection. The K-G results imply that many hundreds of thousands of murders should have been occurring when a private gun was not available for protection. Yet guns are rarely carried, less than a third of adult Americans personally own guns, and only 27,000 homicides occurred in 1992.⁵⁹

Given the number of victims allegedly being saved with guns, it would seem natural to conclude that owning a gun substantially reduces your chances of being murdered. Yet a careful case-control study of homicide in the home found that a gun in the home was associated with an increased rather than a reduced risk of homicide.⁶⁰

Finally, the 2.5 million figure would lead us to conclude that, in a serious crime, the victim is three to four times more likely than the offender to have and use a gun. Although the criminal determines when and where a crime occurs, although pro-gun advocates claim that criminals can always get guns, although few potential victims carry guns away from home, the criminal, according to K-G, is usually out-gunned by the individual he is trying to assault, burglarize, rob or rape.

The explanation K-G offer for this finding is nonsensical. That defensive gun use is substantially more common than criminal gun use, they write, "should not come as a surprise, given that there are far more gun-owning crime victims than there are gun-owning criminals and that victimization is spread out over many different victims, while offending is more concentrated among a relatively small number of offenders."⁶¹ In fact, criminals are more rather than less likely than victims to possess guns. The statistics in question are the number of *times* criminals and victims use guns, not how many different individuals use guns. If a single criminal uses a gun in 20 robberies and in four of these cases a victim uses a gun in self-defense, the usage rates are 5-1 in favor of criminals, not 4-1 in favor of victims.

X. CONCLUSION

Self-report surveys of rare events easily lead to huge overestimates

⁵⁹ KLECK, *supra* note 7, at 51 tbl.2.2 (1991).

⁶⁰ See Arthur L. Kellermann et al., *Gun Ownership As a Risk Factor for Homicide in the Home*, 329 NEW ENG. J. MED. 1084, 1087 (1993).

⁶¹ Kleck & Gertz, *supra* note 4, at 180.

of the true incidence of such events, particularly if the event in question has some potential social desirability. Researchers who claim that such survey incidence data are accurate must show how they have eliminated the enormous problem of false positives. Kleck and Gertz do not accept, let alone meet, this burden of proof. Their survey methodology does not ensure a Specificity rate of well over 99%. Attempts to determine the external validity of their estimates only buttress the presumption of massive overestimation. The conclusion seems inescapable: the Kleck and Gertz survey results do not provide reasonable estimates about the total amount of self-defense gun use in the United States.

TABLE 1
RESULTS OF A SCREENING TEST

	Truth	Truth	
Screen: Response to Self-Defense Gun Question	Positive: A Self-Defense Gun Use	Negative: No Self-Defense Gun Use	Total
Positive	a	b	a + b
Negative	c	d	c + d
Total	a + c	b + d	a + b + c + d

TABLE 2A
RESULTS OF A SCREENING TEST

	Truth	Truth	
Screen: Response to Self-Defense Gun Question	Positive: A Self-Defense Gun Use	Negative: No Self-Defense Gun Use	Total
Positive	16	50	66
Negative	0 (actually 0.16)	4934	4934
Total	16	4984	5000

Assumptions: 66/5000 screens report a positive finding
 Test Sensitivity: 99%
 Test Specificity: 99%

(Or a random 1% of respondents are misclassified)

Predicted Incidence:	66/5000 = 1.33
True Incidence:	17/5000 = 0.32
Survey Overestimation	4 times too high
Estimated total Defensive Gun Use:	2.5 million
True total Defensive Gun Use:	.6 million

TABLE 2B
RESULTS OF A SCREENING TEST

	Truth	Truth	
Screen: Response to Self-Defense Gun Question	Positive: A Self-Defense Gun Use	Negative: No Self-Defense Gun Use	Total
Positive	2	64	66
Negative	0	4934	4934
Total	2	4998	5000
Assumptions:	66/5000 screens report a positive finding Test Sensitivity: 98.7% Test Specificity: 98.7%		
Or a random 1.3% of respondents are misclassified			
Predicted Incidence:	66/5000 = 1.33		
Actual Incidence:	2/5000 = 0.04		
Survey Overestimation	33 times too high		
Estimated Total Defensive Gun Use:	2.5 million		
True Total Defensive Gun Use:	.076 million (or 76,000)		

TABLE 2C
RESULTS OF A SCREENING TEST

	Truth	Truth	
Screen: Response to Self-Defense Gun Question	Positive: A Self-Defense Gun Use	Negative: No Self-Defense Gun Use	Total
Positive	2	64	66
Negative	2	4932	4934
Total	4	4996	5000
Assumptions:	66/5000 screens report a positive finding Test Sensitivity: 50.0% Test Specificity: 98.7%		
Or individuals who are actual positives are 38 times more likely to be misclassified as individuals who are actual negatives			
Predicted Incidence:	66/5000 = 1.33		
Actual Incidence:	4/5000 = 0.08		
Survey Overestimation	17 times too high		
Estimated Total Defensive Gun Use:	2.5 million		
True Total Defensive Gun Use:	.15 million (or 152,000)		