Crime & Delinquency
1–24

DOI: 10.1177/0011128718763138
journals.sagepub.com/home/cad
SAGE

Response Errors in Surveys of Defensive Gun Use: A National Internet Survey Experiment

Gary Kleck[1]

Abstract

A national Internet survey of a probability sample of 5,550 U.S. adults was used to study possible sources of error in surveys of defensive gun use (DGU). Respondents (Rs) were randomly exposed to variant question wordings, question sequences, and combinations of questions. Rs were 70% more likely to report a victimization when they were instructed to report incidents involving offenders known to them, and 43% more likely to report a victimization if they were instructed to include incidents that resulted in no injury or property loss. Rs were 125% more likely to report DGUs if they were directly asked about DGU than if they were first asked about victimization experiences, then asked about DGU in connection with those experiences.

Keywords

defensive gun use, surveys, response errors, firearms, DGU prevalence

The costs and benefits of firearm ownership are of considerable importance in a nation with millions of guns in private hands, and huge numbers of violent acts committed with guns. By the end of calendar year 2015, there were

[1]Florida State University, Tallahassee, USA

Corresponding Author:
Gary Kleck, College of Criminology & Criminal Justice, Florida State University, 112 S. Copeland Street, Tallahassee, FL 32306-1127, USA.
Email: gkleck@fsu.edu

probably more than 380 million guns in private hands (Kleck, 1997, p. 97; U.S. Bureau of Alchohol, Tobacco and Firearms, 2017). In 2014, guns were used to attack or threaten victims in about 9,795 criminal homicides, 138,012 robberies, and 156,416 aggravated assaults, for a total of about 304,223 gun crimes known to the police (U.S. Federal Bureau of Investigation, 2014; total gun crimes based on victim survey data are lower; U.S. Bureau of Justice Statistics, 2014). Guns are also used by crime victims for self-protection, and it matters to gun control policy debates whether such use prevents harm such as injury or property, and how often such benefits are experienced.

In 2004, the National Research Council's Committee to Improve Research Information and Data on Firearms issued a report in which an entire chapter was devoted to "The Use of Firearms to Defend Against Criminals" (Wellford & Pepper Petrie, 2004, Chapter 5). The Committee concluded that the surveys on which estimates of the frequency of defensive gun use (DGU) were based were subject to numerous potentially serious errors, and that it was important to do "experimental evaluation" of the various sources of error, and that "the committee strongly believes that these types of studies can and should be undertaken" (p. 114). The present study is a response.

Issues in Survey Estimation of DGU Prevalence

In this section, a number of potential problems in survey research are discussed, but there is no attempt to cover every methodological problem that could conceivably influence DGU estimates. There is no discussion of coverage error, nonresponse error, or sampling error. Instead, the focus is primarily on sources of response error whose magnitude could be empirically estimated using random assignment of different question wordings and sequences in a single survey.

A DGU logically requires three elements: (a) a person *experiences a crime victimization* with a direct face-to-face confrontation with the offender, (b) the crime victim possesses or can retrieve a gun at the time of the victimization incident, and (c) the victim *uses* the gun for self-protection, that is, uses the gun to attack or threaten the offender(s). Correspondingly, whether survey respondents (Rs) report a DGU depends on whether they are willing and able to report the victimization, their possession of a gun at the time of the incident, and their use of a gun to threaten or attack their adversary. In the following section, we review prior methodological research bearing on response errors regarding each of these kinds of experiences, as well as the character and results of previous surveys of DGU prevalence.

Prior Research

Prior Surveys of DGU

At least 18 national surveys of probability samples of the U.S. adult population have asked Rs specifically about DGU (see Table 1). The estimates of the national DGU count per year vary sharply, from a low of 600,000 to a high of 3.7 million. These estimates were all based on telephone interviews with large random digit dial samples of noninstitutionalized adults age 18 or over, but are nevertheless not directly comparable with each other. One reason is that the surveys did not all attempt to estimate the prevalence of the same universe of events. In some surveys, Rs were asked only about DGUs in which the victim used a handgun, while in others, DGUs involving any type of firearm were covered. Rs in some surveys were told to exclude uses against animals, or uses linked with military or police duty, while Rs in other surveys were not told this, so the resulting DGU estimates included these kinds of uses. In some surveys, Rs were asked about DGUs they had personally experienced, while Rs in other surveys were asked to answer on behalf of all members of their households. This can cause differences in DGU estimates across surveys because Rs provide more complete information about their own experiences than they do about those of other people (see Census Bureau experiments on reporting of crimes in Kalish, 1981, p. 28; Ludwig, Cook, & Smith, 1998, regarding reporting of gun ownership; Kleck & Gertz, 1995, p. 165, regarding reporting of DGUs).

The surveys also differed regarding what segments of the sample were asked the DGU question. In some surveys, all Rs were asked the question, while in others, the question was asked only of those Rs who had reported personal or household gun ownership. This can distort DGU estimates because people can use guns that either belonged to other people, or that the Rs previously owned, but no longer owned by the time they were interviewed.

Finally, the recall period differed across surveys. Rs in some surveys were asked about DGU experiences in their entire lives, a recall period that would be both very long and that would differ for persons of different ages. Rs in other surveys were asked about the preceding 5 years or, in just two surveys, the previous year. This will lead to differences in estimates because the fraction of relevant events that are correctly recalled generally increases as the recall period is shortened (Groves, 2004, pp. 422-430).

Some critics have argued that sample surveys are inherently incapable of providing meaningful estimates of DGU frequency (Hemenway, 1997, 2004), though supporters of this approach have systematically responded to each of the criticisms (Kleck, 2001; Kleck & Gertz, 1997).

Table 1. National Surveys of DGU in the United States.[a]

Survey	Cambridge reports	DMI[a]	DMI[b]	Hart	Time/CNN	Mauser
Time of interviews	April-May, 1978	May-June, 1978	December 1978	October 1981	December 1989	March-April, 1990
Sample size	1,500	1,500	1,010	1,228	605	344
Population covered	Adults	Registered voters	Registered voters	Registered voters	"Firearm owners"	Residents
Gun type covered	Handguns	All guns	All guns	Handguns	All guns	All guns
Recall period	Ever	Ever	Ever	5 years	Ever	5 years
Excluded uses against animals?	No	No	Yes	Yes	No	Yes
Excluded military, police uses?	No	Yes	Yes	Yes	Yes	Yes
Defensive question asked of	Protection hgun owners	All	All	All	Gun owners	All
Defensive question refers to	R	Household	Household	Household	R	Household
Unadjusted % adults with DGU[b]	3	15	7	4	9	3.79

(continued)

Table 1. (continued)

Survey	Gallup	Kleck & Gertz	Gallup	*LA Times*	Tarrance	CDC
Time of interviews	May 1991	February-April, 1993	December 1993	April 1994	May 1994	April-September, 1994
Sample size	1,002	4,997	1,014	1,682	1,000	5,238
Population covered	Adults	Adults	Adults	Adults	Adults	Adults
Gun type covered	All guns	All guns	All guns	All guns	All guns	All guns
Recall period	Ever	1 year	Ever	Ever	5 years	1 year
Excluded uses against animals?	No	Yes	No	No	Yes	Yes
Excluded military, police uses?	No	Yes	Yes	Yes	Yes	Yes
Defensive question asked of	Rs in handgun households	All	Gun owners	All	Rs in gun-owning households	Rs in gun-owning households
Defensive question refers to	R	R	R	R	R	R
Unadjusted % adults with DGU[b]	8	1.326	11	8[c]	1	2.0

(continued)

Table 1. (continued)

Survey	NSPOF	Hemenway & Azrael	Hearst	Hemenway	Gallup	*Washington Post*
Time of interviews	November-December, 1994	May-June, 1996	August 1997	Spring 1999	May 2000	June 2000
Sample size	2,568	1,906	2,016	2,474	1,031	1,068
Population covered	Adults	Adults	Adults	Adults	Adults	Adults
Gun type covered	All guns	All guns	All guns	All guns	All guns	All guns
Recall period	1 year	5 years	Ever	5 years	Ever	Ever
Excluded uses against animals?	Yes	Yes	No	Yes	No	No
Excluded military, police uses?	Yes	Yes	No	Yes	No	No
Defensive question asked of	All	All	All	All	All	All
Defensive question refers to	R	R	R	R	R	R
Unadjusted % adults with DGU[b]	1.44	0.73	5	1.15	7	8

Source. Kleck (2001); Roper Center (2011) iPoll Databank online database.

Note. DMI = Decision Making Information; *R* = respondent; Hgun = handgun; DGU = defensive gun use; CDC = Centers for Disease Control and Prevention; NSPOF = National Survey of the Private Ownership of Firearms; NCVS = National Crime Victimization Survey.

[a]Table covers surveys of probability samples of the general U.S. population that directly asked Rs about DGU. It excludes the survey reported in McDowall, Loftin, and Presser (2000), which was instead based on samples of "commercial lists of likely gun owners" (p. 8), and the NCVS, which never asks Rs specifically about DGU.

[b]This percentage is the share of persons or households who reported a DGU for whatever recall period was used, for whatever subset of gun types or circumstances that happened to be specified in the survey's original question. Thus, these figures are generally not even minimally comparable across surveys.

[c]This survey inquired only about DGUs outside the home.

Underreporting of Victimizations

It is a logical necessity that Rs are able and willing to report a victimization experience if they are to report a victimization that resulted in the victim's use of a gun in self-protection. There is a wealth of evidence indicating that crime victims underreport their victimization experiences in surveys, even in the U.S. Census Bureau's very sophisticated National Crime Victimization Survey (NCVS). Reverse record checks conducted by Census Bureau staff found that even for crimes that victims had reported to the police, and even when asking about a fairly short recall period of 12 months, victims failed to report 33% of known victimizations to survey interviewers (Murphy & Dodge, 1981). The problem was especially severe for assaults, as victims failed to report to interviewers 63% of the assaults that they had reported to police. The degree of underreporting would almost certainly have been still worse if the researchers could somehow have included in their samples crime victims who did not report their victimizations to police.

Underreporting of crimes to the police has been found to be especially serious for crimes without injury to victim or property loss, and for crimes where the offender was a person known to the victim (U.S. Bureau of Justice Statistics, 1985). By definition, successful DGUs result in no harm to the victim. If underreporting of crimes to surveyors follows the same patterns as underreporting to the police, these patterns imply that crimes in which the victim's defensive actions were effective in preventing injury or property loss would be especially likely to be censored out of samples of survey-reported crimes (a problem noted long ago by early victimization researchers; Hindelang & Gottfredson, 1976).

In the NCVS, interviewers stress to Rs that they should include in their reports of victimizations crimes in which someone *attempted* to attack the R or *attempted* to steal something from the R, as well as completed crimes, thereby emphasizing that crimes that did not result in physical injury or property loss are also relevant to the interviewer's inquiries (see Q.42a, U.S. Bureau of Justice Statistics, 1996, pp. 119, 125). In contrast, none of the private surveys asking DGU questions explicitly encouraged Rs to report DGUs linked with uncompleted crimes. This may result in failures to trigger memories of these experiences, or failures of Rs to recognize that incidents that "turned out OK" for the victim are nevertheless crimes, and that victim defensive actions linked with them would therefore also be relevant to the interviewer's inquiries.

Reverse record checks have also shown that Rs are especially likely to fail to report victimizations in which the offenders were persons known to the victim. For example, reverse record checks in one study showed that, while

75% of violent crimes involving strangers were reported to interviewers, only 22% of those involving relatives were reported (Turner, 1981, p. 26). In the NCVS, Rs are specifically instructed, in a screener interview, to include incidents involving offenders known to them, but none of the private gun surveys did this. Rs in the latter surveys who failed to understand that victimizations by persons known to them, such as abusive husbands, qualified as relevant victimizations would be unlikely to report any defensive use of guns connected with those incidents. This could distort DGU prevalence estimates, as well as biasing samples of DGU incidents by underrepresenting uses against persons known to the victim.

Underreporting of Gun Possession

Victim possession of a gun is also a necessary element of a DGU. No one has tested the validity of survey self-reports of victim gun possession during crime incidents, but tests of the validity of reporting of gun ownership in surveys consistently indicate substantial numbers of false negative responses. For example, it has been found that as much as 12.7% of persons registered with government agencies as gun owners—presumably an especially legitimate, law-abiding set of gun owners—will deny owning guns when questioned in a survey (Kellermann, Rivara, Banton, Ready, & Fligner, 1990; Rafferty, Thrush, Smith, & McGee, 1995). Furthermore, some Rs who report having a legally required gun owner's *license* at one point in an interview will nevertheless deny owning guns when questioned at a later point in the same interview (Kleck, 1991, pp. 455-457). Concealing of gun ownership would presumably be higher among people who own guns illegally. Researchers have also found that in husband–wife households, wives are significantly less likely to report household gun ownership than if husbands are interviewed, even though the share who report a household gun (as distinct from individual gun ownership) should be identical for husbands and wives (Kleck, 1997, pp. 66-67; Ludwig et al., 1998). In sum, the evidence consistently indicates that many gun owners fail to correctly report gun possession in surveys. We have no evidence on whether any significant number of nonowners falsely report gun ownership in surveys.

Underreporting of Illegal Behavior

DGU in a crime incident that occurred in a public place necessarily entails victim gun possession in a locale other than the victim's home. Except for the 3% of adults who have firearm carry permits (U.S. Government Accountability Office, 2012), gun possession in such a location would generally be a

violation of laws forbidding unlicensed carrying of firearms. Although carry laws have become less restrictive in recent years, in the 1978-2000 period when the DGU surveys were conducted, it was unlawful to possess a gun in a public place without a carry permit in nearly all states (Kleck, 1997). Most DGUs occur in public places (Kleck & Gertz, 1995, p. 185). Thus, the reporting of DGUs by persons without a carry permit would require the R to confess to the crime of unlawful gun possession, even if their defensive actions themselves were legally justifiable.

There is no definitive evidence directly bearing on the validity of survey reports of unlawful gun carrying or possession in public places in particular, but there is considerable evidence of substantial underreporting of illegal behaviors in general. The most convincing evidence on underreporting of illegal acts by survey Rs concerns the use of illegal drugs, as it is possible to reliably determine, through urine tests and hair assays, the reality of whether a person has indeed used illegal drugs, and thus committed the crime of possession of illicit drugs. This body of research is especially valuable because it allows detection of both underreporting and overreporting, that is, both false negative responses and false positive responses. The evidence consistently indicates that false positives are rare, that false negatives are far more common than false positives, and thus that surveys, on net, underestimate illicit drug use.

One of the largest scale tests of drug reporting validity ever conducted was performed in connection with the Drug Use Forecasting program of the U.S. Justice Department. Nearly 22,000 arrestees were first interviewed about their drug use (without being told of the possibility of later urinalysis), then their urine was tested for the presence of the metabolites of illicit drugs. For every drug, false positives were rare, and greatly outnumbered by false negatives. False negatives outnumbered false positives by a factor of 15.4 for cocaine, 3.1 for opiates, 3.3 for amphetamines, and 1.3 for marijuana (Harrison, 1995, p. 94). In sum, Rs underreport this type of criminal behavior more than they overreport.

This pattern in self-reports is not confined to drug use. A unique study applying polygraph exams to assess validity found, across a wide variety of illegal behaviors, that overreporting of illegal or controversial behaviors was rare, and that underreporting was more common. Of 35 deviant behaviors covered, overreporting exceeded underreporting for just four behaviors (Clark & Tifft, 1966, pp. 517-518). Violence clearly labeled as aggressive rather than defensive (“started a fist fight”) and weapon carrying (“carried a razor, switchblade, or gun as weapon”) were among the few behaviors showing a net overreporting. Nevertheless, the researchers judged 92.5% of the weapon carrying responses, and 80% of the aggressive violence responses to

be accurate. These instances of overreporting might be due to the exclusively male character of the study sample (p. 519). For self-reporting of criminal behavior in general, underreporting is the dominant pattern.

Specificity of Self-Protection Questions

There is evidence from victimization surveys that more specific prompts yield higher rates of reports of crime experiences. For example, the estimated prevalence of rape victimization tripled when NCVS interviewers asked specifically about sexual assault ("any rape, attempted rape or other type of sexual assault") in screener questioning, compared with previous NCVS procedures in which Rs were asked about assaults in general and then later asked for details about the nature of the assault (U.S. Bureau of Justice Statistics, 1996, p. 151).

A similar problem could result from the way the NCVS asks about victim defensive actions. Interviewers never ask Rs specifically about DGU or any other specific type of self-protection action. Rather, they only ask Rs who had reported a victimization a general question about measures they had taken to protect themselves: "Did you do anything with the idea of protecting yourself or your property while the incident was going on?" Those who respond "yes" are then asked an open-ended question: "What did you do?" (U.S. Bureau of Justice Statistics, 1996, p. 133). Commenting on this aspect of the NCVS, the research director of the General Social Surveys, Tom Smith (1997) observed that "indirect questions that rely on a respondent volunteering a specific element as part of a broad and unfocused inquiry uniformly lead to undercounts of the particular of interest" (pp. 1462-1463). The absence of a prompt specifically pertaining to DGU may fail to trigger memories of such experiences or fail to communicate their relevance to the interviewer's inquiries.

Long Recall Periods

It has long been recognized that long recall periods result in more recall failure (much of it presumably due to memory failure), and thus underreporting of the behaviors and experiences of interest (Groves, 2004). This has been documented for a wide variety of kinds of experiences, but has been specifically shown to affect recall of crime victimization experiences—The more time that passes from the occurrence of the crimes to interviewing, the lower the share that is recalled. A reverse record check study found that even with recall periods of 12 months or less, 19% of crime incidents reported to police were not recalled at all by the victims when interviewed (Dodge, 1981). Many DGU surveys used either lifetime recall periods, or asked about the 5 years preceding the interview (Table 1), so recall failure was worse than this,

as underreporting of experiences increases as the recall period is extended back in time (Groves, 2004, pp. 422-430). Recall failure and forward telescoping (reporting an experience that happened prior to the recall period) are about equally common in victimization surveys using a 1-year recall period (Dodge, 1981), but the size of recall failure effects relative to telescoping effects grows as recall periods lengthen (Sudman & Bradburn, 1973; Woltman, Bushery, & Carstensen, 1984). Thus, the estimate-reducing effects of recall failure are likely to outweigh the estimate-increasing effects of telescoping to an increasing degree, as recall periods increase in length past 1 year.

Survey Mode Effects

We did not randomly assign survey mode in the present research, as only one mode (Internet) could be used. Nevertheless, survey mode is clearly relevant to the present study as it uses a mode different from that used in all previous national DGU surveys. All prior national surveys of DGU have relied on the telephone mode for conducting interviews—Questions are spoken by human interviewers who contacted Rs by phone, and the Rs speak their answers to interviewers. Prior experimental research consistently indicates that Rs are much less likely to reveal illegal or controversial behaviors or attributes when surveyed via the telephone mode (see review in Kleck & Roberts, 2012). It is possible that telephone interviews with human callers do not provide a strong enough sense of privacy for Rs to be willing to report illegal or controversial actions. It may also be more embarrassing to directly speak admissions of such behaviors to human interviewers.

Random assignment studies have found that Rs are substantially more likely to report socially undesirable information about themselves in Internet surveys than in telephone surveys (Eaton et al., 2010; Kleck & Roberts, 2012; Kreuter, Presser, & Tourangeau, 2008; van de Looij-Jansen & de Wilde, 2008). These findings are consistent with a broader range of experimental studies indicating that Rs are more willing to report illegal, socially undesirable, controversial, or sensitive behaviors, experiences, and attributes in *self-administered* modes in general than in surveys using interviewers, such as telephone surveys (see reviews by Tourangeau & Smith, 1998). Indeed, the telephone mode appears to be the least effective of common survey modes in eliciting self-reports of illegal behavior (Kleck & Roberts, 2012).

The NCVS is especially problematic in this regard, because a large share of the interviewing is conducted by telephone. Furthermore, unlike most telephone surveys, the NCVS cannot promise its Rs anonymity because they are named and most were interviewed in their homes when they first become part

of the NCVS sample (McDowall, Loftin, & Presser, 2000). The private DGU surveys summarized in Table 1 also used telephone interviewing, but these interviews were more anonymous than those in the NCVS because callers working for the private survey firms did not ask for the identities of Rs.

Early experimental findings indicate that the web mode is even more effective at eliciting reports of sensitive or controversial behaviors than other self-administered modes. For example, of 16 illegal behaviors that Eaton et al. (2010) asked about, the web mode elicited more admissions than self-administered paper-and-pencil questionnaires for all 16 crimes, and eight of these differences were significant at the .05 level. Of particular relevance to the present research, both Eaton et al. (2010, p. 146) and van de Looij-Jansen and de Wilde (2008, p. 1715) found that Rs randomly assigned to the web mode were significantly more likely to admit to carrying a weapon, compared with those who filled out self-administered paper-and-pencil questionnaires. As most DGUs involve gun carrying due to their public location (Kleck & Gertz, 1995), these findings suggest that use of the web mode instead of telephone interviewing would increase Rs' willingness to admit the unlawful gun carrying that commonly accompanies a DGU.

Question Sequence Effects

A DGU survey can (a) ask about DGU first, then ask about details of the crime involved, as was done in the National Self-Defense Survey (Kleck & Gertz, 1995) and the National Survey of the Private Ownership of Firearms (NSPOF; Cook & Ludwig 1996; 1998), or (b) ask about crime victimization first and then ask about defensive actions taken by the victim, as is done in the NCVS. Asking first about DGU focuses the R's attention specifically on the R's use of a gun, and may be more likely to trigger recall of a DGU. In a survey of "likely gun owners," McDowall et al. (2000) randomly assigned Rs to either being asked about DGU first, or victimization first, and found that Rs were significantly more likely to report a DGU if asked about DGU first. Unfortunately, the randomly assigned conditions in this study also differed regarding question wordings and other attributes, so the study was not able to isolate the effect of question sequence by itself (pp. 6, 14).

A number of other possible sources of error have been identified, though they cannot be addressed in detail here. For example, the surveys summarized in Table 1 all share coverage errors—They do not cover adolescents or persons in households without telephones. Furthermore, it has been hypothesized that false negative response errors about DGUs are more common if a survey is sponsored by an agency of the federal government such as the Justice Department, or the survey is conducted by a government agency like

the Census Bureau, as is the case with the NCVS (Kleck & Gertz, 1995). Also, estimates derived from private one-time surveys are affected by forward telescoping, in which experiences that occurred before the recall period are reported (Kleck & Gertz, 1995). In contrast, the panel design of the NCVS largely eliminates this problem. Methodological studies of the NCVS indicate that telescoping in a survey with a 1-year recall period could inflate estimates of victimization frequency by a maximum of 21%, but that this effect is canceled out by a roughly equal amount of recall failure with a 1-year recall period (Dodge, 1981; Kleck & Gertz, 1995, p. 171). Scholars have also speculated that some incidents reported as DGUs in private surveys may actually have involved the R using a gun when there was no actual criminal threat. McDowall and his colleagues (2000) argued that in some claimed DGUs, Rs may have used guns preemptively against persons that the Rs erroneously believed were going to victimize them. In such cases, the R's gun use was not really defensive. Others have argued that some purported DGUs were illegal (Hemenway, Azrael, & Miller, 2000). Finally, it has also been noted that estimates derived from the NCVS do not include DGUs linked with less serious crimes, such as trespassing or vandalism.

Methods of the Present Study

This study focuses on some potential sources that are amenable to experimental assessment in a one-time survey. The experiments were carried out under the aegis of the Time-Sharing Experiments for the Social Sciences (TESS; 2011) program sponsored by the National Science Foundation. Research projects supported under this program all involve random assignment experimentation in surveys conducted on the Internet. At the time this study was done, survey work was carried out by the GfK (formerly Knowledge Networks) survey firm; the TESS program has recently changed from GfK to another survey firm.

GfK's "Knowledge Panel" Internet surveys are unusual because they are based on true probability samples of virtually the entire U.S. population (97% coverage), including the approximately 25% who would not otherwise have home Internet access were it not for their recruitment into the GfK panels of survey Rs. GfK uses address-based sampling to select potential Rs, using the U.S. Postal Service's Delivery Sequence File of residential addresses, which provides virtually complete coverage of the U.S. population other than the homeless. Thus, the samples even cover households without telephone service. Persons randomly selected from this file are contacted by GfK via mail and (if necessary and possible) by telephone, and invited to participate in the GfK "KnowledgePanel." Those without home Internet access are provided with a

free computer and Internet service, thereby extending coverage of the samples to a segment of the population not covered in other Internet surveys.

For any one survey, GfK randomly selects a subset of their panel members, who are invited via email to participate in the survey, whose topic is unknown to them. Those who agree to participate use a link in the email invitation message that takes them to the survey's webpage, where they are provided with instructions and the survey's questionnaire, which is available in English, Spanish, or any of many other languages. Rs read the questions on their computer monitor screen, and provide their responses using their mouse or keyboard (GfK, 2015).

The present research was based on a large (n = 5,550 completions) nationally representative sample of adults (age 18+)—the largest ever used in a survey asking Rs about DGU (recall that the NCVS never asks Rs specifically about DGU). The data reported in Table 2 demonstrate that the GfK sample matches up extremely closely with the U.S. adult population. Even before weighting, the GfK sample was very similar to the population, and weighting (using the GfK post-stratification weights) made the sample essentially identical to the U.S. adult population with regard to its distribution across age, sex, race, education, and region categories.

The survey employed an experimental design with random assignment of Rs to survey conditions. In combination with the large sample size, this design effectively rules out the influence of confounding factors, making it unnecessary to introduce multivariate controls to isolate the effect of the survey conditions on responses to DGU-related questions. The study is therefore unusual in combining both the strong external validity (generalizability) produced by the use of a national probability sample, and the strong internal validity produced by random assignment manipulation of survey conditions.

The survey was fielded from November 25 to December 3, 2009. Response rates for Internet panel surveys are difficult to compare with those for one-time surveys, but the most comparable rate is probably the "study completion rate"—of 8,188 GfK panel members randomly selected and invited to participate, 5,550 or 67.8% completed the questionnaire. The break-off rate was 2.1%—2.1% of those who started the survey failed to complete it.

Rs were randomly assigned to

1. either receive an explicit instruction to report victimization incidents involving offenders known to them, or to receive no such instruction;
2. either receive an explicit instruction to report incidents that did not result in any injury or property loss, or to receive no such instruction;

Table 2. The Internet Sample Compared With the U.S. Adult Population (percentages).

	U.S. 2009	Sample	
	Population, 18+	Unweighted	Weighted[a]
Age			
18-29	22.4	17.7	21.7
30-44	26.4	24.8	26.5
45-59	27.3	28.5	27.8
60+	23.8	29.0	23.9
Sex			
Male	48.7	49.7	48.5
Female	51.3	50.3	51.5
Race/ethnicity			
White non-Hispanic	65.1	71.0	68.7
Black non-Hispanic	12.3	11.3	11.5
Other non-Hispanic	5.4	3.5	5.4
Hispanic	15.8	10.3	13.4
Multiple races, non-Hispanic	1.5	3.9	1.1
Education			
Less than high school	14.1	12.5	12.8
High school	30.9	27.6	30.9
Some college	28.0	30.2	28.5
BA or higher	27.0	29.7	27.8
Region			
New England	4.7	4.8	4.7
Mid-Atlantic	13.3	13.3	13.8
East North Central	15.1	16.7	15.1
West North Central	6.6	7.4	6.8
South Atlantic	19.3	19.2	19.3
East South Central	6.0	6.3	6.6
West South Central	11.7	10.7	10.7
Mountain	7.2	6.6	7.4
Pacific	16.1	14.8	15.6

[a]Cases weighted by GfK post-stratification weights.

3. either be asked about DGU and victimization experiences in the previous 12 months, or in the previous 5 years;
4. either be directly asked a question about DGU similar to the one used in Kleck and Gertz (1995), or to be first asked about victimization experiences, then asked about any self-protective actions they may have taken during the incident, as in the NCVS;

5. either be asked a generic question about any self-protective actions (as in the NCVS), or a question specifically asking about DGU in particular (as in the private surveys).

Each of these conditions was randomly varied independent of the other experimentally manipulated conditions. Thus, any one condition is uncorrelated with the other conditions, as well as uncorrelated with attributes of the Rs. Some combinations of conditions were never applied to any one R because they would not make sense to an R. Specifically, it would not have made sense to first ask the R about DGU in response to a victimization, and then—after they answered—to expose them to the instructions to (a) report victimizations involving no injury or property loss or (b) to report victimizations involving known offenders. Although the total sample size was 5,550, the number of Rs exposed to any one combination of conditions was necessarily smaller.

The exact wording of all the questions can be found in the online Technical Appendix. The numbers of Rs assigned to each condition are shown in Table 3.

Results

Each panel of Table 3 displays the findings pertaining to a randomly varied survey condition. All percentages in the rightmost column are based on cases weighted by the GfK post-stratification weights. That column also displays the two-tailed significance of the difference in percentages between the two survey conditions compared.

Panel A indicates that Rs were 70% more likely to report a victimization when they were explicitly instructed to report incidents involving offenders known to them than when they were not so instructed. This supports the hypothesis that the failure of past DGU surveys to include such an instruction contributed to an underestimation of DGUs due to failures to trigger recall of victimization experiences.

Panel B indicates that Rs were 43% more likely to report a victimization if they were instructed to include incidents that resulted in no injury or property loss to the victim. As previous DGU surveys did not include such instructions, this suggests, first, that they missed substantial numbers of DGUs and, second, that they were especially likely to miss successful DGUs, that is, those involving neither injury nor property loss to the victim.

Panel C reports results concerning the impact of using longer recall periods. Prior research and common sense suggests that there should be more DGU reports with a longer recall period. More specifically, if we assumed

Table 3. Reports of Victimization and DGU by Experimental Condition.

Panel A: Rs were instructed to include incidents involving offenders known to them	% reporting a victimization
No instructions (K2A) (n = 1,254)	4.4
Instruction (K2B) (n = 1,231)	7.5 ($p < .01$)
Panel B: Rs were instructed to include incidents with no injury or property loss	**% reporting a victimization**
No instructions (K2A) (n = 1,254)	4.4
Instruction (K2C) (n = 1,250)	6.3 ($p < .01$)
Panel C: Recall period—Past 12 months vs. past 5 years	**% reporting a DGU**
Past 12 months (n = 2,780)	2.9
Past 5 years (n = 2,770)	2.6 ($p > .10$)
Panel D: Question directly asking about DGU vs. victimization question first, then question about self-protection	**% reporting a DGU, past 12 months**
Question about DGU first as in Kleck and Gertz (1995) (K1)	6.3
Question about victimization (some version of K2), then question about self-protective actions taken as in NCVS (K4a and K5A)	2.8 ($p < .01$)
Panel E: General vs. specific question about DGU	**% reporting a DGU past 12 months**
Open-ended question about any self-protection actions (n = 2,784) (K4A and K5A)	2.7
Question specifically asking about DGU (n = 2,732) ((K4B)	2.8 ($p > .10$)

Note. Question numbers are provided in parentheses—see Technical Appendix. DGU = defensive gun use; NCVS = National Crime Victimization Survey.

roughly constant actual DGU rates over the 5-year period preceding the survey (roughly 2004-2009, a period when trends in crime rates and gun ownership were fairly flat) and no repeat DGU experiences, we would expect reports of DGUs to be roughly 5 times larger with a 5-year recall period than with a 1-year recall period. The experimental data indicate that Rs were no more likely to report a DGU for the previous 5 years than for the previous 12

months. Thus, 5-year recall seems to be only one fifth of what was expected based on a 1-year recall period. This anomaly could be partly due to telescoping of DGUs when the 12-month recall period was used but, as previously noted, telescoping is far too infrequent to explain this large an inconsistency. Instead, it is more likely that the anomaly is primarily due to the greater recall failure that occurs when longer recall periods are used. The problem would be even worse with lifetime recall periods. As 10 of the 18 previous DGU surveys used lifetime recall periods (Table 1), this problem affects estimates from most of the prior DGU surveys.

Findings reported in Panel D indicate that Rs were 125% more likely to report DGUs if they were (a) directly asked about DGU than if they were (b) first asked about victimization experiences, then asked about DGU in connection with those experiences. This finding supports the hypothesis that few DGUs are reported in the NCVS because of the question sequence, in which interviewers first ask about victimization, then about self-protective actions.

Panel E of Table 3 indicates that there was no significant difference in reporting of DGU depending on whether the R was asked a general question about any kind of self-protection action, followed by a question asking what type of self-protective action was taken, versus a question asking specifically about DGU. Thus, this particular feature does not seem to contribute to distortions in estimates of DGU prevalence.

Readers are cautioned against comparing the reported frequency of DGUs (Panels D and E) with the reported frequency of criminal victimization in general (Panels A and B). The relevant questions were not worded so as to yield comparable results. For example, the scope of the DGU questions is broader than the general victimization questions because the latter were confined to incidents in which the R saw the offender(s), while former were not.

Discussion

These findings suggest that the best survey estimate of DGU prevalence would be produced by a survey that (a) used a 1-year recall period; (b) used a question sequence asking about DGU first, then victimization; (c) included explicit instructions to report victimizations involving offenders known to the R; and (d) included explicit instructions to report victimizations involving no injury or property loss.

Due to the experimental design, however, our Rs were randomly assigned to receive *either* the instruction about incidents with known offenders *or* the instruction about those with no injury or property loss. No Rs were given both instructions, even though this would have improved recall.

Unlike with the National Self-Defense Surcey (Kleck & Gertz 1995), it was not possible, due to limits imposed by the TESS program, to ask a long series of follow-up questions establishing details of the events that Rs regarded as DGUs. Further information would almost certainly disqualify some of them.

The fact that a given Method A yields more reports of DGU or of the victimization experiences that could result in a DGU than Method B does not necessarily mean Method A produces more accurate estimates than Method B. This inference would be inaccurate if false positive responses concerning DGUs outnumbered false negative responses, and survey estimates therefore tended to be too high. If that were true, more reported DGUs using Method A might indicate that this method generates estimates that are more unduly high than when Method B was used. Prior research on survey methodology, however, has failed to obtain any empirical evidence that false positives are in fact more numerous than false negatives, with regard to any crime-related experience, in surveys of the general adult population. Both types of response errors occur in surveys, but there is no evidence that the former exceeds the latter in samples of the general adult U.S. population, that is, samples of the sort used in the DGU surveys. For the types of illegal behavior for which it has been possible to detect both kinds of erroneous responses, false negatives have consistently been found to outnumber false positives (Clark & Tifft, 1966; Harrison, 1995). Therefore, given that underreporting has been the more frequent kind of response error heretofore documented in surveys of crime-related experiences, methods generating higher reported rates of such experiences are at present best regarded as the method generating more valid results. The typical view was expressed in an extensive review of self-report methodology by Thornberry and Krohn (2000), who noted that

> based on the results of the tests of criterion validity, there appears to be a substantial degree of either concealing or forgetting past criminal behavior. Although the majority of respondents report their offenses and the majority of all offenses are reported, there is still considerable underreporting. (p. 58)

In contrast, these authors did not cite any evidence that there was any significant amount of overreporting or false positives. Even when noting the existence of false positives, scholars routinely note them to be much less common than false negatives: “While far less likely to occur than under-reporting or concealing of offenses, there is evidence that there is also some over-reporting in self-reports” (Loughran, Paternoster, & Thomas 2014, pp. 681-682)

There is no definitive evidence on the true relative balance of false positives and false negatives concerning DGU in particular, but there is some

relevant evidence. In the NSPOF, Rs were asked whether they had used a gun defensively, along with an extensive series of follow-up questions inquiring about details of reported DGUs. After each interview was completed, the interviewers were required to answer these questions: (a) "Did the respondent hesitate to answer the question about having used a gun for self-defense, or otherwise express reluctance to do so, and then answer "no"? indicating to the interviewer that the "respondent seemed to be concealing defensive gun use" and (b) "Did you get any impression that the R was either making up a gun self-defense incident or trying to mislead you about their role in the incident?" (Questions 116 and 117 in Police Foundation, 1998). Although the interviewers obviously could not know with certainty what R's true experiences were, they did have the advantage of actually hearing the R's tone of voice, any extraneous remarks they might have made, word choices, hesitations, and audible indications of nervousness or evasiveness. Secondary analysis of the NSPOF data reveals that interviewers believed that Rs were concealing a DGU in 60 cases, while Rs were making up a DGU or misstating their role in it in just 13 cases. Thus, based on the interviewers' judgments, suspected false negatives outnumbered suspected false positives by a margin of 4.6 to 1 (Kleck, 2001, p. 253).

Conclusion

The experiments reported here provide strong evidence that DGUs are more likely to go unreported if Rs are not explicitly told to report crime incidents in which the R escaped injury and property loss, or incidents involving offenders known to the R. Likewise, Rs are less likely to report DGU experiences if they are asked to recall events over longer time periods—the effects of recall failure apparently outweigh the effects of forward telescoping to a greater extent with a longer recall period. Asking directly about DGU elicits more reports of DGU experiences than asking first about victimization experiences, and then asking about self-protective actions taken in connection with reported crime experiences. Furthermore, though survey mode was not experimentally manipulated in this study, web surveying appears to elicit a higher rate of DGU reports than telephone interviews do. While it is possible that higher DGU reports might reflect greater overreporting due to more false positive reports, there is no foundation in prior survey research on response errors in crime surveys to expect more false positive responses than false negatives, and interviewer perceptions of possible misreporting suggest that false negatives outnumber false positives.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the National Science Foundation, under its Time-sharing Experiments for the Social Sciences program.

References

Clark, J. P., & Tifft, L. (1966). Polygraph and interview validation of self-reported deviant behavior. *American Sociological Review*, *31*, 516-523.

Cook, P. J., & Ludwig, J. (1996). *Guns in America*. Washington, DC: The Police Foundation.

Cook, P. J., & Ludwig, J. (1998). Defensive gun uses: New evidence from a National Survey. *Journal of Quantitative Criminology*, *14*:, 111-131.

Dodge, R. W. (1981). The Washington D.C. recall study. In R. G. Lehnen & W. G. Skogan (Eds.), *The National Crime Survey: Working Papers*, Volume I (pp. 12-15). Washington, DC: U.S. Government Printing Office.

Eaton, D. K., Brener, N. D., Kann, L., Denniston, M. M., McManus, T., Kyle, T. M., . . . Ross, J. G. (2010). Comparison of paper-and-pencil versus web administration of the Youth Risk Behavior Survey (YEARBS): Risk behavior prevalence estimates. *Evaluation Review*, *34*, 137-153.

GfK. (2015). Retrieved from http://www.gfk.com/products-a-z/us/knowledgepanel-united-states/

Groves, R. M. (2004). *Survey errors and survey costs*. Hoboken, NJ: John Wiley.

Harrison, L. (1995). The validity of self-reported data on drug use. *Journal of Drug Issues*, *25*, 91-111.

Hemenway, D. (1997). Survey research and self-defense gun use: An explanation of extreme overestimates. *Journal of Criminal Law and Criminology*, *87*, 1430-1445.

Hemenway, D. (2004). *Private guns, public health*. Ann Arbor: University of Michigan Press.

Hemenway, D., Azrael, D., & Miller, M. (2000). Gun use in the United States: Results from two national surveys. *Injury Prevention*, *6*, 263-267.

Hindelang, M., & Gottfredson, M. (1976). The victim's decision not to invoke the criminal justice process. In W. F. McDonald (Ed.), *Criminal justice and the victim* (pp. 72-98). Beverly Hills, CA: SAGE.

Kalish, C. B. (1981). The Dayton-San Jose recall study. In R. G. Lehnen & W. G. Skogan (Eds.), *The National Crime Survey: Working Papers*, Volume I (pp. 28-29). Washington, DC: U.S. Government Printing Office.

Kellermann, A. L., Rivara, F. P., Banton, J., Ready, D., & Fligner, C. L. (1990). Validating survey responses to questions about gun ownership among owners of registered handguns. *American Journal of Epidemiology*, *131*, 1080-1084.

Kleck, G. (1991). *Point Blank: Guns and Violence in America*. NY: Aldine de Gruyter.

Kleck, G. (1997). *Targeting guns: Firearms and their control*. New York, NY: Aldine de Gruyter.

Kleck, G. (2001). The frequency of defensive gun use. In G. Kleck & D. B. Kates (Eds.), *Armed* (pp.213-238). Buffalo, NY: Prometheus Books.

Kleck, G., & Gertz, M. (1995). Armed resistance to crime: The prevalence and nature of self-defense with a gun. *Journal of Criminal Law and Criminology*, *86*, 150-187.

Kleck, G., & Gertz, M. (1997). The illegitimacy of one-sided speculation: getting the defensive gun use estimate down. *Journal of Criminal Law & Criminology*, *87*, 1446-1461.

Kleck, G., & Roberts, K. (2012). What survey modes are most effective in eliciting self-reports of criminal or delinquent behavior? In L. Gideon (Ed.), *Handbook of survey methodology* (pp. 415-439). New York, NY: Springer.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys. *Public Opinion Quarterly*, *72*, 847-865.

Loughran, T. T., Paternoster, R., and Thomas, K. J. (2014). Incentivizing responses to self-report questions in perceptual deterrence studies: an investigation of the validity of deterrence theory using Bayesian truth serum. *Journal of Quantitative Criminology*, *30*, 677-707.

Ludwig, J., Cook, P.J. and Smith, T.W. (1998). The gender gap in reporting household gun ownership. *American Journal of Public Health*, *88*, 1715-1718.

McDowall, D., Loftin, C., & Presser, S. (2000). Measuring civilian defensive firearm use: A methodological experiment. *Journal of Quantitative Criminology*, *16*, 1-19.

Murphy, L. R., & Dodge, R. W. (1981). The Baltimore recall study. In R. G. Lehnen & W. G. Skogan (Eds.), *The National Crime Survey: Working Papers* (pp. 16-21), Volume I. Washington, DC: U.S. Government Printing Office.

Police Foundation. (1998). *National study of private ownership of firearms in the United States*, 1994 (Codebook for study 6955). Available from https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/6955

Rafferty, A. P., Thrush, J. C., Smith, P. K., & McGee, H. B. (1995). Validity of a household gun question in a telephone survey. *Public Health Reports*, *110*, 282-288.

Roper Center. (2011). *iPoll online database*. Available from https://ropercenter.cornell.edu/ipoll-database/

Smith, T. (1997). A call for a truce in the DGU war. *Journal of Criminal Law and Criminology*, *87*, 1462-1469.

Sudman, S., and Bradburn, N.M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, *68*, 805-815.

Tark, J., & Kleck, G. (2004). Resisting crime: The effects of victim action on the outcomes of crimes. *Criminology*, *42*, 861-909.

Thornberry, T. P. & Krohn, M. D. (2000). The Self-Report Method for Measuring Delinquency and Crime. *In Criminal Justice 2000*, Volume 4. Washington, D. C.: U. S. Government Printing Office.

Time-Sharing Experiments for the Social Sciences. (2011). Available from http://www.tessexperiments.org/

Tourangeau, R., & Smith, T. W. (1998). Collecting sensitive information with different modes of data collection. In M. Couper, R. Baker, J. Bethlehem, C. Clark, J. Martin, W. Nicholls & J. O'Reilly (Eds.), *Computer assisted survey information collection* (pp. 431-454). New York, NY: John Wiley.

Turner, A. G. (1981). The San Jose recall study. In R. G. Lehnen & W. G. Skogan (Eds.), *The National Crime Survey: Working Papers, Volume I: Current and historical perspectives* (pp. 22-27). Washington, DC: U.S. Government Printing Office.

U. S. Bureau of Alcohol, Tobacco and Firearms. 2017. *Firearms Commerce in the United States: Annual Statistical Update 2017*. Retrieved from https://www.atf.gov/resource-center/data-statistics.

U.S. Bureau of Justice Statistics. (1985). *Reporting crimes to the police* (BJS special report). Washington, DC: U.S. Government Printing Office.

U.S. Bureau of Justice Statistics. (1996). *Criminal victimization in the United States, 1993*. Washington, DC: U.S. Government Printing Office.

U.S. Bureau of Justice Statistics. (2014). *Criminal victimization, 2013*. Retrieved from http://www.bjs.gov/content/pub/pdf/cv13.pdf

U.S. Federal Bureau of Investigation. (2014). *Crime in the United States, 2013*. Retrieved from https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013

U.S. Government Accountability Office. (2012). *Gun control: States' laws and requirements for concealed carry permits vary across the nation*. Washington, DC: U.S. Government Printing Office.

van de Looij-Jansen, P. M., & de Wilde, E. J. (2008). Comparison of web-based versus paper-and-pencil self-administered questionnaire. *Health Services Research*, *43* (5, Pt. 1), 1708-1721.

Wellford, C. J. V., & Pepper Petrie, C. V. (2004). *Firearms and violence: A critical review* (Committee to improve research information and data on firearms). Washington, DC: The National Academies Press.

Woltman, H., Bushery, J., & Carstensen, L. (1984). Recall bias and telescoping in the National Crime Survey." In Robert G. Lehnen and Wesley G. Skogan (Eds.), *The National Crime Survey: Working Papers, Volume II: Methodological Studies* (pp. 90-93). Washington, D.C.: U.S. Government Printing Office.

Author Biography

Gary Kleck is the Emeritus David J. Bordua professor of criminology and criminal justice at Florida State University. His research has focused on the topics of the impact of firearms and gun control on violence, deterrence, crime control, and violence. He is the author of *Point Blank: Guns and Violence in America*, which won the 1993

Michael J. Hindelang Award of the American Society of Criminology. He also wrote *Targeting Guns* (1997); with *Don B. Kates, Jr.*, *The Great American Gun Debate* (1997) and *Armed* (2001); and, with Brion Sever, *Punishment and Crime* (2017). His articles have been published in the *American Sociological Review, American Journal of Sociology*, *Social Forces*, *Social Problems*, *Criminology*, *Journal of Criminal Law and Criminology*, *Law & Society Review*, *Journal of Research in Crime and Delinquency*, *Journal of Quantitative Criminology*, *Crime and Delinquency*, and many other journals. He has testified before Congress and state legislatures on gun control issues and his work has been cited by the U.S. Supreme Court. He has served as a consultant to the National Research Council, National Academy of Sciences Panel on the Understanding and Prevention of Violence, as a member of the U.S. Sentencing Commission's Drugs-Violence Task Force, and a member of the National Research Council Committee on Priorities for a Public Health Research Agenda to Reduce the Threat of Firearm-Related Violence.