

How the Hemenway Surveys Distorted Estimates of Defensive Gun Use
Frequency

Gary Kleck

College of Criminology and Criminal Justice

Florida State University

Tallahassee, FL 32306-1273

June 14, 2018

Abstract

David Hemenway and his colleagues have claimed that two national surveys that they conducted indicated there were few defensive gun uses (DGUs) in the U.S., and that the number of gun crimes committed with firearms greatly is far larger than the number of DGUs. This paper explains how these authors produced extreme underestimates of DGU frequency and gross overstatements of the number of gun crimes.

Underestimating DGUs was accomplished by (1) using an eccentric and biased wording of the DGU question, (2) using a trap question that misled respondents (Rs) with a DGU into thinking they had already reported the DGU, (3) employing a long recall period that increased memory loss, and the (4) selecting a biased sample that systematically underrepresented people likely to have a DGU. Overstating the number of gun crimes was accomplished by (1) mischaracterizing incidents falling into a largely meaningless “hostile display” category as gun crimes, when the authors’ own evidence indicated that most of the “victims” of these displays did not regard them as a part of gun crime, and (2) ignoring the far more sophisticated National Crime Victimization Survey estimates of gun crime.

One of the most contentious issues in the research literature on the links between firearms and violence is the question of the frequency of defensive gun use (DGU). It is critical to any assessment of the relative costs and benefits of widespread gun ownership. The only feasible way to measure the prevalence or incidence of DGU is with surveys, given that a large share of them are not reported to the police, and police do not maintain counts of DGUs aside from the tiny fraction (<1%) that result in the offender's death (Kleck and Gertz 1995).

While nearly all national surveys have yielded estimates of over a million DGUs per year, and the highest quality surveys have generated estimates over two million (Kleck and Gertz 1995; Cook and Ludwig 1996; Kleck 2001), two national surveys have yielded radically lower estimates – surveys done by David Hemenway and colleagues in 1996 (Hemenway and Azrael 2000) and 1999 (Hemenway, Azrael, and Miller 2000). This comment explains how Hemenway and his colleagues radically underestimated DGU prevalence.

How Discrepant Were the DGU Estimates Implied by the Hemenway Surveys?

First, it is worth noting just how different the results of these two surveys were from other national surveys. Table 1 summarizes the estimates of DGU frequency implied by the findings of 19 national surveys, including the two by Hemenway and his colleagues. All are professionally conducted surveys of probability samples of the U.S. adult population in which respondents were asked a question specifically pertaining to defensive uses of firearms. Producing an estimate of the number of annual DGUs often entailed nothing more complicated than computing the percent of Rs who reported a DGU in the survey sample and multiplying it by the U.S. adult (18+) resident population as of the year the survey was fielded. Many surveys, however, did not ask questions pertaining to the same universe of events that were covered in the better surveys. For example, some included uses against animals, though most did not. Some excluded uses by police and security guards, others did not; some included uses of all types of

guns, while others included only uses of handguns, and so on. The results of each survey were adjusted, where necessary, to make them comparable across surveys, i.e. so that all of the estimates pertained to the DGU frequency that would have been obtained if (1) the survey had used a one-year recall period, (2) excluded uses by police officers, security guards, and military personnel, (3) excluded uses against animals, (4) included uses of all gun types, not just handguns, (5) used information directly provided by the respondent about their own experiences rather than second-hand information about the experiences of other household members, (6) covered all DGUs regardless of their location, not just those occurring in the user's home, and (7) asked the DGU question of all respondents (Rs), not just those reporting current gun ownership. The full set of adjustment procedures are described in detail in (Kleck 2001, pp. 272-275). Readers who are dubious about these adjustments may rely instead on the results of the two surveys in Table 1 whose results did not need adjustment, designated Kleck and Gertz, and NSPOF. These surveys, fielded in 1993 and 1994, yielded extremely similar estimates of the annual number of DGUs, averaging 2.4 million

(Table 1 about here)

What did the two Hemenway surveys indicate? In the 1999 survey, after the authors excluded cases they did not regard as genuine DGUs, they were left with 29 Rs who reported a DGU, in the five-year recall period preceding the survey, out of 2,521 total Rs, for a five-year DGU prevalence estimate of 1.15%. This figure cannot simply be divided by five to get a one-year estimate, because Rs remember experiences within the past year better than those experienced two to five years earlier. Research directly comparing DGU estimates based on both 5-year and 1-year recall periods indicates that 1-year estimates are about 0.40 of 5-year estimates, not 0.20 (Kleck and Gertz 1995, p. 184). Therefore the 1.15% 5-year estimate was multiplied by 0.40 to establish what the estimate would have been had a past-year recall period been used. This produces a past-year prevalence of 0.46%. Multiplying this times the 1999

resident population age 18+ of 207,093,000 yields an estimated 952,907 DGUs per year. In the 1996 survey, 14 of 1,905 Rs or 0.73% reported DGUs against a person in the previous five years, excluding uses by police, security guards and military. Applying the same adjustment procedures to these data as with the 1999 data, the 1996 survey indicates that just 0.29% of the Rs reported a past-year DGU, implying 690,000 adults who had DGUs.

Before discussing the gap between these estimates and those yielded by the other surveys in Table 1, it is worth first noting that the Hemenway estimates are seven to ten times the “estimate” of DGUs supposedly implied by the National Crime Victimization Survey (NCVS), of roughly 65,000 (McDowall and Wiersema 1994) – the one that Hemenway has repeatedly endorsed (Hemenway 1997; Hemenway and Azrael 2000; Hemenway et al. 2000). Even his own surveys’ results indicate that the NCVS-based “estimate” of DGUs is far too low.

The reasons for the extremely small numbers of Rs in the NCVS reporting a DGU have been detailed elsewhere (Kleck 2001), so only a brief outline is needed here. The NCVS has never asked even a single respondent specifically about defensive use of firearms. Instead, Rs are only asked an open-ended question about whether they did anything to protect themselves during a victimization incident, and are then asked to mention whatever specific self-protective actions they might be inclined to volunteer. Commenting on this problem in the NCVS, Tom Smith, Research Director of the National Opinion Research Center, stated 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” (Smith 1997, p. 1462). Hemenway’s own results confirm this insight, reinforcing the conclusion that the NCVS cannot provide meaningful estimates of DGU frequency.

Further, the NCVS is conducted by one federal government agency, the Census Bureau, and the Rs are told that the information will be provided to another federal agency, the Justice Department. The interviews are not anonymous – Rs know that the surveyors know their address

and telephone number, and thus that their answers can be linked to them as individuals. These are extremely unfavorable circumstances in which to get Rs to report a behavior as controversial as threatening or attacking another person with a deadly weapon, however justifiable the action might have been.

The DGU annual estimate implied by Hemenway's 1996 survey was the lowest ever obtained in a national survey, while the 1999 survey yielded an estimate only slightly higher than the next-lowest estimates. Thus, Hemenway's DGU estimates are radically inconsistent with the results of 19 other national surveys conducted by a variety of organizations, including news outlets with strongly antigun editorial policies like the *Los Angeles Times*, *Washington Post*, and CNN, with estimates of annual DGUs in the U.S. ranging from 1 to 4 million. Why did the surveys conducted by Hemenway and his colleagues yield DGU estimates that were so discrepant? The rest of the paper helps explain their deviant results, and how the flaws in their surveys led to the erroneous conclusion that gun crimes greatly outnumber DGUs.

Eccentric Wording of the DGU Question

This is exact wording of the question used in the 1996 survey that was intended to inquire about DGUs:

“In the past 5 years, have you used a gun in self-defense to protect yourself from a person or people?” [underlining in original]

This is the very similar question used in the 1999 survey:

“In the past 5 years, have you used, displayed or brought out a gun in self-defense to protect yourself from a person or people?” [underlining in original]

Many surveys had previously asked Rs about DGUs, but no one other than Hemenway and his colleagues has ever used the wording employed in their surveys, before or since. Why did they choose to reinvent the wheel? While one might speculate that the wording invented by

Hemenway somehow improved on the wordings used in the best previous surveys (Kleck and Gertz 1995; Cook and Ludwig 1996; CDC 1996-1998), a close examination indicates otherwise.

First, the use of the phrase “to protect yourself” without the phrase “or others” has the effect of arbitrarily excluding uses in which the gun was used to defend *others*. Contrast this with the wording used in the Center for Disease Control and Prevention’s (CDC) Behavioral Risk Factor Surveillance System (BRFSS) surveys for 1996, 1997, and 1998: “During the last 12 months, have you confronted another person with a firearm, even if you did not fire it, *to protect yourself, your property, or someone else?*” (CDC 2018; emphasis added).

Readers will notice another difference in the Hemenway wording from the CDC wording: Hemenway did not provide Rs with the instruction that they should report DGUs “even if you did not fire it.” Without this phrase, some people who had used a gun to threaten a criminal would think that their experience did not qualify because they did not fire the gun, assuming that the word “used” implied firing the gun. This could make a huge difference all by itself, since over 60% of DGUs do not involve the defender firing the gun. Instead, most DGUs involve the gun merely being used to threaten the criminal (Kleck and Gertz 1995, p. 185).

Thus the eccentric Hemenway wording deviated significantly from wording of DGU questions used in all other surveys in two crucial ways, and the deviations were not neutral regarding their effect on estimates of DGU frequency: the peculiarities in the wording both had the effect of reducing the number of people likely to report a DGU.

Use of a “Trap Question”

Hemenway et al. drew their conclusion about DGU frequency based on the combined results of the 1996 and 1999 surveys (Hemenway et. al. 2000). In the 1996 survey, before asking the question intended to inquire whether the respondent (R) had committed a defensive gun use (DGU), the authors asked another, highly ambiguous question supposedly intended to measure what the authors called “hostile gun displays.” Here is the exact wording:

“In the past five years, has anyone displayed or brought out a gun in a hostile manner, even if this event did not take place during the commission of a crime?”

This curious question had never been asked in any previous survey, and appears to have been invented by Hemenway and/or his co-authors. Likewise, the ambiguous concept of “hostile gun displays” had never been used by any other survey researchers. The question is extraordinarily ambiguous in at least three ways. First, the question could be interpreted as inquiring about either something the R *did*, or something that was done *to* the R. Second, the curious use of the phrase “has anyone ...” opens up the possibility that the question was not even limited to asking about something directly involving the R, but might also encompass things the R merely witnessed, or heard about second-hand. Indeed, of 122 Rs in the 1996 initially reporting a “hostile gun display,” 31 revealed that they were reporting something they had merely witnessed (Hemingway et al. 2000, p. 264). If the question were interpreted literally, *everyone* in the sample should have answered “yes,” since it is obviously true that someone, sometime, has displayed a gun in a hostile manner. Third, and most importantly, what does the phrase “displayed or brought out a gun in a hostile manner” mean? If the R interpreted the question to be asking about things that the R did instead of things done to him, it would clearly encompass defensive uses of guns by the R when she or he was a crime victim, since any defensive use of a gun is surely done “in a hostile manner” by the crime victim towards the offender. Even Hemenway and Azrael partially conceded the problems with this question, admitting that “we also do not know whether or not the hostile gun use was in fact a self-defense gun use against the respondent” (Azrael and Hemenway 2000, p. 290). They did not, however, admit that such a gun use could also have been a self-defense gun use *by* the respondent.

Hemenway and his colleagues (2000) claimed to have reduced the ambiguity of this question in the 1999 survey, by making it clearer that they were asking about things done “against” the R, but this is not what their documentation for their 1999 survey indicates.

Question 22 reads “In the past five years, approximately how many times has anyone used, displayed or brought out a gun in a hostile manner?” The interviewer then told the R “I am going to ask you some question about the most recent time someone displayed or brought out a gun in a hostile manner” (ICPSR 2007). Thus, the ambiguity of the 1996 question was not reduced in the 1999 survey. In both surveys, the R could easily interpret the question as asking about either acts the R had committed with a gun or DGUs directed at the R, and, more specifically, could interpret the question as encompassing DGUs that Rs themselves had committed.

Here is the significance of the authors asking this highly ambiguous question first, and then asking the question that was intended to inquire about DGUs. Many of the Rs in this survey who had used a gun defensively would have answered “yes” to the first (“hostile display”) question with the DGU in mind. Then when they were asked the question that was supposedly intended to measure DGUs, those who had interpreted the “hostile display” question as encompassing DGUs would have thought they had already reported the DGU in response to the earlier question. Since it would not make sense for the surveyors to ask the same question twice, the most reasonable interpretation of the DGU question, for those that had committed a DGU, would be that it was asking about any *additional* DGU experiences the R might have experienced, beyond the one they thought they had already reported in response to the “hostile display” question. Other than the few Rs who had multiple DGUs in the recall period, defensive gun users who interpreted the question this way would have to answer “No” to the DGU question, intending to indicate that they had experienced no *additional* DGUs beyond the one they believed they had already reported. Hemenway et al., however, interpreted “no” responses to the second question to always indicate that the R did not have any DGU experiences.

In sum, to the extent that the “hostile display” question served to mislead Rs into thinking they had already reported their DGUs, it would contribute to a serious underestimation of DGU

prevalence. So it was not surprising that the authors obtained results implying an estimate of DGU frequency that was far lower than estimates obtained in every other survey on the topic – none of which used a trap question (Kleck 2001; 2018).

Long Recall Period in Both Surveys

For both surveys, the interviewers asked Rs to report DGUs that occurred in the five years preceding the interviews - a very long recall period. Although some might think that victimization experiences are surely memorable enough for victims to recall all or nearly all of them even many years after the event, research by the Census Bureau during the development of the National Crime Victimization Surveys (NCVS) showed that this is clearly not so. Most victimizations are minor, and a large share of them are not recalled in surveys. For example, reverse record checks conducted during the development of the NCVS found that, even among crimes serious enough to have been reported to the police, *most* assaults were not reported to the Census Bureau victim survey interviewers (Dodge 1981; Murphy and Dodge 1981).

More specifically, it has been established that questions asking about the previous five-years yield far fewer reports of DGUs per year than would be expected based on the number reported as occurring in the past year (Kleck and Gertz 1995, p. 184). Use of a five-year recall period by Hemenway and his colleagues served to increase recall failure due to memory problems and to push the DGU estimates down. Hemenway et al. never explained why they used so long a recall period, or acknowledged that it would tend to reduce estimates of DGU frequency.

Sample Bias – Under-representation of the Subpopulations Most Likely to Have a DGU

Both the 1996 and 1999 surveys were based on grossly biased samples of the U.S. adult population. The bias was documented for the 1996 survey in Hemenway and Azrael (2000, p.

260), but the article describing the 1999 survey did not report any evidence bearing on the issue of whether the sample was representative of the population, and said nothing in the text about the issue (Hemenway et al. 2000). More importantly, neither article acknowledged the most important implication of the sample biases – they systematically underrepresented groups that were most likely to suffer direct-contact criminal victimizations, and thus sets of people most likely to use a gun for self-protection.

Both surveys severely underrepresented males. While the percent male should have been 49% if the sample were representative of the national adult population with respect to gender, the 1996 survey sample was only 42% male and the 1999 sample was just 41.5% male. The samples were even more seriously biased regarding race. While 12% of the U.S. population is African-American, only 8.5% of the 1996 sample and 9.0% of the 1999 sample were African-American. Worse still, there were way too few low-income people included in either sample. While 23% of the U.S. population in 1996 had a household income under \$15,000, only 12.1% of the 1996 survey's sample, and 9.3% of the 1999 sample had household incomes under \$15,000 (Hemenway and Azrael 2000, p. 260; author's analysis of ICPSR 2007)

In sum, the Hemenway samples systematically underrepresented the people most likely to suffer a direct-contact criminal victimization, and thus most likely to experience an occasion for using a gun for protection against crime – an obvious implication of the sample biases that Hemenway et al. did not share with their readers. They could have partially compensated for these biases by using post-stratification weighting, but did not.

The Meaning of “Hostile Gun Displays” and the Balance of Gun Crimes and DGUs

Based on the results of their 1996 survey, Hemenway and Azrael (2000, p. 257) claimed that “criminal gun use is far more common than self-defense gun use.” Is there any empirical foundation for this assertion? In the five-year period from 1995 through 1999, the National

Crime Victimization Survey estimated that there were 3,355,025 violent crime incidents in which offenders possessed, but did not necessarily use, firearms (U.S. Bureau of Justice Statistics 2000). Research by the Bureau of Justice Statistics (BJS) found that 46.8% of NCVS-reported victimizations classified as handgun crimes involved the handgun merely being present, but not actually used to attack or even threaten the victim (Rand 1994). In many violent crimes in which the victim believed the offender possessed a gun, the victim was indeed attacked or threatened, but not with the gun. Thus, only 53.2% of the NCVS-reported events that are loosely labelled by BJS as “gun crimes” involve a gun actually being used to threaten or attack the victim. Kleck and Gertz (1995) only classified an event as a DGU if the victim used the gun to threaten or attack the criminal, so only crimes in which the offender *used* a gun to threaten or attack the victim should be counted if one is to make a meaningful comparison of offensive and defensive gun *uses*.

If only 53.2% of the 3,355,025 gun-linked violent crimes committed from 1995 through 1999 involved the offender using the gun to threaten or attack the victim, this means there were 1,784,873 such crimes – an average of 356,975 gun crimes per year. Even if every incident involved a different victim, spreading the victimizations across the maximum number of different victims, these figures imply that 0.176% of Americans were the victim of a violent crime in which the offender used a gun to threaten or attack. Recall that the 1999 Hemenway et al. survey, ignoring its critical flaws for the moment, indicated that there were 952,907 defensive uses of guns per year, which is 0.471% of the adult (18+) resident population. Thus, even this unusually low DGU estimate was 2.68 times as large as the NCVS-estimated number of gun crimes. In short, Hemenway’s own data indicated that DGUs occur far more often than gun crimes – exactly the opposite of what he and his colleagues claimed.

How then did Hemenway and his co-authors manage to invert the reality of their DGU findings? Instead of using the results of the sophisticated NCVS regarding the number of gun

crimes, they chose to compare the frequency of DGUs with the frequency of their ambiguous “hostile gun displays.” They falsely claimed to know that these experiences were “criminal gun uses” (Hemenway et al. 2000, p. 266) even though that was not at all what the “hostile displays” question had actually asked about. The authors even insisted that their results regarding the frequency of these “criminal gun uses” were “consistent with results from the NCVS” (p. 266). In reality, their 1999 survey indicated that 3.7% of the population (94 of 2,521 Rs) experienced a “hostile gun display” in the previous five years, or about 0.75% in any one year assuming an even distribution across years – a rate that was *more than four times* the 0.176% NCVS estimate of the prevalence of gun crime.

This is how Hemenway and his colleagues summarized their main conclusion: “We find that criminal gun use is far more common than self-defense gun use” (Hemenway et al., p. 257). Note that the authors did not claim that their survey provided a better estimate of the prevalence of gun crime victimization than the NCVS. Instead, they simply ignored the gun crime estimates of the far more sophisticated NCVS in favor of their own meaningless estimates of “hostile gun displays.”

Worse still, the authors withheld from readers information that directly contradicted their characterization of “hostile gun displays” as “criminal gun uses.” When I analyzed the raw data from their 1996 survey, I found that 75% of the Rs reporting a “hostile gun display” *did not consider the display to be “part of the commission of a crime”* (ICPSR 2006, frequencies on variable vc35). Although the authors claimed that, after exclusions of ineligible cases, there were 58 Rs in the 1996 survey who had a “hostile gun display” (Hemenway et al. 2000, p. 264), their data indicate that there were only 26 Rs who thought the display was part of crime (ICPSR 2006, frequencies on variable vc35). Thus, most of the Rs that the authors were claiming as victims of “hostile gun displays” (at least 32 of the 58) did not regard the displays as part of a gun crime. Use of the ambiguous “hostile gun display” concept, however, allowed the authors to

enormously increase the number of gun-related experiences that they could characterize as somehow “against the interests of society” (Hemenway et al. 2000, p. 263).

To summarize, Hemenway and his colleagues managed to create the impression that DGUs were greatly outnumbered by gun crimes by grossly overstating gun crimes and underestimating DGUs. The former was accomplished by (1) mischaracterizing incidents falling into the largely meaningless “hostile display” category as gun crimes, when their own evidence indicated that most of the “victims” of these displays did not regard them as a part of gun crime, and (2) ignoring the far more sophisticated NCVS estimates of gun crime. Underestimating DGUs was accomplished by (1) using an eccentric and biased wording of the DGU question, (2) using a trap question that misled Rs with a DGU into thinking they had already reported the DGU, (3) employing a long recall period that increased memory loss, and the (4) selecting a biased sample that systematically underrepresented people likely to have a DGU.

Criminal Court Judges’ Opinions of the Legality of 35 Reported DGUs

Hemenway and his colleagues argue that it doesn’t really matter how many DGUs are reported in surveys because most of these uses were “probably illegal” and “against the interests of society” (Hemenway, et al. 2000, p. 263). One of the more eccentric analyses performed by the authors involved asking five unnamed criminal courts judges for their opinions on the legality of 35 DGUs reported in the two surveys, based solely on summaries of the incidents crafted by the authors. It is impossible to know exactly how this curious exercise was carried out since the authors provided only a very sketchy outline of their methods - a description that was more noteworthy for what it did not say than for what it did say. The authors did not say why these particular five judges were picked (was there prior information that the judges had unusually narrow conceptions of lawful self-defense?), did not provide readers with a copy of the instructions provided to judges describing their task, did not provide copies of the summaries of

purported DGU incidents provided to the five judges, and did not provide transcripts of Rs' descriptions of their claimed DGUs (all easily posted in an online appendix). Further, the judges provided their opinions on just 35 claimed DGUs, far too few to yield meaningful estimates of the fraction of DGUs that are lawful.

It is unclear why the authors thought that the legal status of these actions was important, as distinct from their defensive or moral character, but this analysis is nevertheless worth a closer look for its own sake. The authors reported that their five judges came from Massachusetts, Pennsylvania, and California, but did not say how many came from each of these three states, leaving open the possibility that three of the five judges served in MA. Convenience could even have led the Harvard-based researchers to recruit three judges from their local Cambridge, MA area – one of the most liberal areas in the nation. Massachusetts and California have two of the lowest state gun ownership rates in the nation (Okoro et al. 2005, p. e372), and all three states have gun control laws that are far stricter than the average in the U. S. (Giffords Law Center to Prevent Gun Violence 2018). One could bias judgements in favor of interpreting DGUs as unlawful merely by picking judges from states where few law-abiding citizens own guns, where firearms are tightly controlled, and that have unusually narrow definitions of lawful self-defense.

Further, criminal courts judges ordinarily would not be familiar with the laws of states other than their own. The laws governing self-defense are not identical in the three states where the judges worked, which means that the set of five judges used at least three different standards to assess the legality of reported DGUs. Further, the authors did not report any effort to match the state in which a reported DGU occurred with a judge from that state, nor would that have been possible given that the researchers only used judges from three states. Thus, most of the time a judge was assessing the lawfulness of a purported defensive act using legal standards of self-defense that did not even apply in the state where the reported DGU occurred.

A naive observer might be tempted to think that this is not a serious problem because self-defense laws and the way they are interpreted do not vary significantly from one state to another. One strong indication that courts are far less likely to accept claims of self-defense in some states than in others is the highly varying percent of homicides that are treated as justifiable homicides. In Houston, Texas, 6.1% of all homicides were treated as civilian (nonpolice) justifiable homicides, and the share in Dade County (Miami), Florida was even higher, at 12.7%. In sharp contrast, just 1.3% of all homicides in Philadelphia, Pennsylvania were legally classified as justifiable (summarized in Kleck 1991, p. 147). The assessments of the authors' five judges could therefore be the product of nothing more than a biased selection of judges and of the unusually narrow legal definitions of lawful self-defense that prevailed in the states where they served.

The authors did not report verbatim the instructions they provided to the judges describing their task, so there is no way to tell whether the instructions could have biased the judges' opinions about the legality of the reported DGUs. Likewise, the authors did not share with readers the summaries of DGU incidents provided to judges, or the full transcripts of Rs' descriptions of their DGUs, so there is no way to tell whether the summaries fairly represented the incidents or, on the other hand, whether they excluded details that would have supported the view that the DGU was lawful. For example, 85% of those claiming a DGU in the 1996 survey stated that they had tried to avoid a confrontation with their adversary (my analysis of ICPSR 2006), but there is no way to tell if the summaries that were provided to the judges included this crucial information.

It is worth pointing out one-sided use that the authors' made of these five judges. The judges were asked to assess the legal status of claimed DGUs, but not the experiences that the authors characterized as "hostile gun displays." This one-sided procedure had the effect of the judges effectively "disqualifying" some of the DGUs as lawful self-defense acts, while not

“disqualifying” any of the “hostile gun displays” as “gun crimes.” The analysts could have made use of the judges’ legal expertise to assess how many of “hostile gun displays” actually were gun crimes, as the authors claimed, but they did not do so (or at least did not report the results of any such assessments).

Nonoperative Caveats Regarding the Limits of the DGU Findings

Hemenway often engages in the practice of what could be called the “nonoperative caveat.” He draws a dubious conclusion from flawed research, then in the “limitations” section of his research report weakly acknowledges a few of the flaws in his research (the caveats), but finally states conclusions that make sense only if one completely ignores the caveats. The caveats seem to have been stated for the sake of superficially satisfying academic norms and pacifying journal reviewers, but have no influence on the conclusions that Hemenway and his coauthors draw. The report in their 2000 article in *Injury Prevention* about the relative balance of DGUs and gun crimes provides an especially stark example. The authors conceded that their sample of five judges was a convenience sample rather than a “random sample,” and also was “too small to be confident of the stability of the aggregate ratings we report here.” This logically should mean that they could not generalize the judges’ rating to the full national population of reported DGUs. Yet two sentences later the authors did precisely that: “Our results indicate that gun use against adults to threaten and intimidate is far more common than self-defense gun use by them, and that most self reported self defense (sic) gun uses are probably illegal” (p. 267). This remarkable *non sequitur* was not stated so as to describe only the nonrandom 35 DGU incidents the judges rated, but rather was stated in unqualified terms to describe DGUs in general.

This is not the only example of Hemenway reporting his findings about this handful of nonrandomly selected incidents as if they were applicable to the nation as a whole. He is the

Director of the Harvard Injury Control Research Center, and presumably responsible for the content of its website regarding DGU frequency. That website makes statements about the frequency of DGU and its supposedly illegal character, based on Hemenway's surveys, that are expressed so as to indicate that they apply to DGUs in the U.S. as a whole, not just the unrepresentative handful studied by Hemenway and his colleagues (Hemenway 2018). There is nothing on the website to indicate that the results of the Hemenway surveys cannot be generalized to the U.S. population.

The website states that “Most purported self-defense gun uses are gun uses in escalating arguments, and are both socially undesirable and illegal,” an assertion based entirely on the five judges' opinions about 35 incidents reported in the 1999 survey. It also unambiguously claims that “guns are not used millions of times each year in self-defense,” a claim based on the two Hemenway surveys plus what the website loosely describes as “epidemiological theory.” (Hemenway 2018). The latter amounts to nothing more than a fallacious argument that surveys are bound to yield more “false positive” responses than “false negative” responses, simply because there is more *potential* for the former than the latter. This bizarre argument is dissected in detail in Kleck (2001, pp. 254-257). As to the 19 other professionally conducted national surveys of probability samples of the U.S. population (Table 1) that unanimously indicate far more DGUs than his two deeply flawed surveys, Hemenway simply ignored them.

Thus, when it came to making propaganda statements about DGUs aimed at large Internet audiences for his website, Hemenway ignored his own caveats. Having superficially satisfied scientific norms for the journal article in which he reported his survey results - stating a few caveats to the effect that his DGU results could not be generalized to DGUs in general - he then dropped the caveats and made public statements that were only true if those caveats were ignored. In short, these caveats were nonoperative, in that they ultimately did not in any way limit the conclusions Hemenway drew from his research.

Conclusions

The biased methods, radically deviant results, and *non sequitur* conclusions that characterize the authors' two surveys and their reports of the research would not be worrisome if only the referees and editors of the journals that published the reports had been deceived. One might optimistically hope that all this is of little concern because the authors' crude tactics surely must have been recognized by acute observers, and their conclusions rejected. Unfortunately, many sophisticated and well-known scholars who study guns and violence have uncritically accepted these conclusions. For example, Philip Cook cited Hemenway et al. as follows: "when the same respondents in the same sort of one-time survey are asked about both DGUs and victimizations by guns, they report many more victimizations than DGUs" (Cook 2013, p. 43). Cook did not explain why it was essential that both of these very different quantities be measured in the same survey. David McDowall even went Hemenway and his colleagues one better by making the bizarre claim that the DGU and "hostile gun use" questions were "similarly worded," before agreeing with them that "hostile gun displays occurred almost twice as often as did uses in self-defense" (McDowall 2005, 258). Jens Ludwig has likewise echoed this uncritical acceptance of the Hemenway claim (Cook and Ludwig 2015, p. 171). If even the specialist experts were fooled, surely more than a few nonspecialist academics and lay people were also misled.

References

- Azrael, Deborah, and David Hemenway. 2000. "In the safety of your own home': results from a national survey on gun use at home." *Social Science and Medicine* 50:285-291.
- Center for Disease Control and Prevention. 2018. Behavioral Risk Factor Surveillance System website at <https://www.cdc.gov/brfss/questionnaires/pdf-ques/96brfss.pdf>. Accessed 2-22-18.
- Cook, Philip J. 2013. "The great American gun war." *Crime and Justice* 42:19-73.
- Cook, Philip J. and Jens Ludwig. 2015. "Elusive facts about gun violence: where good surveys go bad." In *Envisioning Criminology* edited by Michael D. Maltz and Sephen K. Rice. NY: Springer.
- Dodge, Richard. 1981. "The Washington, D.C. recall study." In *The National Crime Surveys*, vol. 1, *Working Papers, Volume I: Current and Historical Perspectives*, ed. Robert G. Lehnen and Wesley G. Skogan (Washington, D.C.: U. S. Government Printing Office.
- Giffords Law Center to Prevent Gun Violence. 2018. Summary of state gun control laws available at <http://lawcenter.giffords.org/search-gun-law-by-gun-policy/>.
- Hemenway, David. 1997. "Survey research and self-defense gun use: an explanation of extreme overestimates." *Journal of Criminal Law and Criminology* 87:1430-1444.
- Hemenway, David. 2018. Statement about character and frequency of DGUs on website of Harvard Injury Control Research Center, at <https://www.hsph.harvard.edu/hicrc/firearms-research/gun-threats-and-self-defense-gun-use-2/>.
- Hemenway, David, and Deborah Azrael. 2000. "The relative frequency of offensive and defensive gun uses: results from a national survey." *Violence and Victims* 15:257-272.
- Hemenway, David, Deborah Azrael, and Matthew Miller. 2000. "Gun use in the United States: results from two national surveys." *Injury Prevention* 6:263-267.
- Inter-university Consortium for Political and Social Research (ICPSR). 2006. Study 1750.

- Hemenway, David, and Azrael, Deborah. *Survey of Gun Owners in the United States, 1996*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-03-30. <https://doi.org/10.3886/ICPSR02750.v1>.
- Inter-university Consortium for Political and Social Research (ICPSR). 2007. Study 4552 *National Firearms Survey, 1999*. David Hemenway. Data Collection Instrument. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-07-06. <https://doi.org/10.3886/ICPSR04552.v1>.
- Kleck, Gary. 1991. *Point Blank: Guns and Violence in America*. NY: Aldine de Gruyter.
- Kleck, Gary. 2001. "The frequency of defensive gun use: evidence and disinformation." Chapter 6 in Gary Kleck and Don B. Kates (eds.), *Armed: New Perspectives on Gun Control*. N.Y.: Prometheus Books.
- Kleck, Gary. 2018. "What do CDC's surveys say about the frequency of defensive gun uses?" Copy available online at the Social Science Research Network.
- Kleck, Gary, and Marc Gertz. 1995. "Armed resistance to crime: the prevalence and nature of self-defense with a gun." *Journal of Criminal Law and Criminology* 86:150-187.
- McDowall, David. 2005. "John R. Lott, Jr.'s defensive gun brandishing estimates." *Public Opinion Quarterly* 69:246-263.
- McDowall, David, and Brian Wiersema. 1994. "The incidence of defensive firearm use by US crime victims, 1987 through 1990." *American Journal of Public Health* 84:1982-1984.
- Murphy, Linda R., and Richard W. Dodge. 1981. "The Baltimore Recall Study." In *The National Crime Surveys: Working Papers, Volume I: Current and Historical Perspectives*, ed. Robert G. Lehnen and Wesley G. Skogan. Washington, D.C.: U. S. Government Printing Office.
- Okoro, Catherine A., David E. Nelson, James A. Mercy, Lina S. Balluz, Alex E. Crosby, and Ali H. Mokdad. 2005. "Prevalence of household firearms and firearm-storage practices in the 50 states and the District of Columbia: findings from the Behavioral Risk Factor

Surveillance System, 2002.” *Pediatrics* 116:e370-e376.

Rand, Michael R. 1994. *Guns and Crime*. Bureau of Justice Statistics. Crime Data Brief.

NCJ-147003. Washington, D.C.: U. S. Government Printing Office.

Smith, Tom W. 1997. “Call for a truce in the DGU war.” *Journal of Criminal Law and Criminology* 87

U.S. Bureau of Justice Statistics. 2000. *Criminal Victimization in the United States, [year] – Statistical Tables*, Table 66. At <https://www.bjs.gov/content/pub/pdf/cvus97.pdf>.

Table 1. National Survey Estimates of the Prevalence of Defensive Gun Use

<u>Survey</u>	<u>When Fielded</u>	<u>% DGU Past Year^a</u>	<u>Annual DGUs (in millions)</u>
Cambridge Reports	1978	0.74	1.1
DMIa	1978	2.22	1.7
DM Ib	1978	1.14	0.9
Hart	1981	2.01	1.7
Time/CNN	1989	4.5	2.6
Mauser	1990	1.5	1.4
Gallup	1991	1.93	3.6
Kleck and Gertz	1993	1.326	2.55
Gallup	1993	1.63	1.4
L. A. Times	1994	2.08	4.0
Tarrance	1994	0.51	1.0
NSPOF	1994	1.44	2.8
Hemenway and Azrael	1996	0.29	0.69
Hearst	1997	0.60	1.2
Hemenway, Azrael, & Miller	1999	0.46	0.95
Gallup	2000	0.84	1.70
Washington Post	2000	0.96	2.0
CNN	2014	1.693	4.03
Pew Research Center	2017	1.037	2.60

Sources:

Kleck (2018).

Hemenway and Azrael, and Hemenway, Azrael & Miller – see text of present paper.

CNN, Pew Research Center: Roper iPoll online database of survey results.

All other surveys: Kleck (2001) and sources cited therein.

Note:

- a. All estimates pertain to the percent of U.S. adults age 18+ who used a gun defensively against a human being in the preceding 12 months, based on direct reports by respondents about their own experiences, excluding uses against animals or by police officers, security guards, or members of the military. It was assumed that each person reporting a DGU had experienced just one DGU. When necessary, survey results were adjusted to make them comparable across surveys. See Kleck (2001, pp. 272-275) for a detailed explanation of what adjustments are needed and how they were applied.