Title

A retail sampling approach to assess impact of geographic concentrations on probative value of comparative bullet lead analysis

Journal
Law Probability and Risk, 4(4)

ISSN
1470-8396

Authors
Cole, Simon A
Tobin, William A
Boggess, Lyndsay N
et al.

Publication Date
2005-12-01

DOI
10.1093/lpr/mgl006

Law, Probability and Risk (2005) **4**, 199–216 doi:10.1093/lpr/mgl006
Advance Access publication on April 27, 2006

A retail sampling approach to assess impact of geographic concentrations on probative value of comparative bullet lead analysis

SIMON A. COLE†
Department of Criminology, Law & Society, 2329 Social Ecology II, University of California, Irvine, CA 92697-7080, USA

WILLIAM A. TOBIN‡
Forensic Engineering International, 2708 Little Gunstock Rd., Lake Anna, VA 23024-8882, USA

LYNDSAY N. BOGGESS§
Department of Criminology, Law & Society, 2329 Social Ecology II, University of California, Irvine, CA 92697-7080, USA

AND

HAL S. STERN¶
Department of Statistics, 346D Computer Science, University of California, Irvine, CA 92697-1250, USA

[Received on 13 October 2005; revised on 17 March 2006; accepted on 18 March 2006]

The probative value of comparative bullet lead analysis (CBLA), a now discontinued technique that was used by the Federal Bureau of Investigation for more than 30 years, has been hotly debated over the last several years. One issue that has received relatively little attention concerns the degree of geographic dispersion of bullets as they pass from manufacturers to retailers. Proponents and critics of CBLA alike agree that geographic distribution is such a major consideration, if not a predominant one, that it could significantly diminish, or completely erode, the probative value of a CBLA 'match' or, in some cases, even make a match counter-probative. The inattention to this issue to date appears to be a consequence of lack of data, rather than lack of importance. Until now, no datum concerning bullet distribution has been presented in the public domain, critically hampering the proper estimation of the probative value of a CBLA match. In this paper, we use manufacturer packing codes on boxes of bullets in retail outlets at four sites in the United States as a surrogate measure of bullet lead compositions to gauge local retail bullet distribution. Using a *weighted average packing code match probability*, we found very high degrees of *geographic concentration* of bullet packing codes. Although these findings can only offer a rough estimate of the degree of geographic concentration of actual chemical compositions of bullets, they are sufficient to establish that geographic concentration

† Email: scole@uci.edu
‡ Email: wtobin@nexet.net
§ Email: lboggess@uci.edu
¶ Email: sternh@uci.edu

does, in fact, exist. Such a concentration would have a significant impact on the probative value of any claimed CBLA match.

Keywords: comparative bullet lead analysis; compositional analysis of bullet lead; geographic concentration; geographic distribution; probative value; National Research Council; forensic bullet composition comparisons; forensic bullet evidence.

1. Introduction

Comparative bullet lead analysis (CBLA) is a forensic service that has been offered by the U.S. Federal Bureau of Investigation (FBI) in criminal investigations since the 1960s. Although for most of that period, the relatively esoteric technique attracted little attention, in recent years the technique generated controversy as research into the practice expanded and a number of studies critical of CBLA were published (Imwinkelried & Tobin, 2003; Randich *et al.*, 2002; Tobin, 2004; Tobin & Duerfeldt, 2002). These critiques were followed by successful legal challenges to the admissibility of CBLA under the *Daubert* standard for scientific evidence[1] and by press coverage of the controversy (Hansen, 2004; Mejia & Sample, 2002; Piller & Mejia, 2003). The controversy prompted the FBI to fund a National Research Council (NRC) panel, which issued a report in 2004. The NRC recommended changes in the practice of analysing and reporting CBLA evidence but stopped short of recommending discontinuing use of the evidence. After publication of the NRC report, the FBI imposed a temporary moratorium on the forensic practice, pending further study. On 1 September 2005, the FBI declared the moratorium permanent indicating that 'neither scientists nor bullet manufacturers' could explain the forensic significance of a claimed match:

> One factor significantly influenced the Laboratory's decision to no longer conduct the examination of bullet lead: neither scientists nor bullet manufacturers are able to definitively attest to the significance of an association made between bullets in the course of a bullet lead examination. (Federal Bureau of Investigation, 2005)

Since the FBI Laboratory was the only forensic laboratory in the world offering CBLA as a routine forensic service to law enforcement agencies, its decision effectively ends use of the evidence in criminal trials, at least for the foreseeable future. Nonetheless, we view continued research and discussion of CBLA as scientifically and legally important for several reasons. Most importantly, numerous cases were determined or influenced at the trial stage by use of proffered CBLA evidence. Many of these cases have post-conviction proceedings still pending, and scientific information regarding the reliability of CBLA may be relevant to these proceedings. In addition, the FBI's recent

[1] *Daubert* v. *Merrell Dow Pharmaceuticals, Inc.*, 113 S. Ct. 2786 (1993), 509 U.S. 573 (1993). *State* v. *Miguel Trujillo*, Case Nos. D-0101-CR-2000-284, D-0101-CR-2000-229, D-0101-CR-99-677 (N.M. Dist. Ct. 2000) (court admitted testimony of claimed bullet lead 'match' but did not allow testimony regarding opinion as to forensic significance); *People* v. *Marlon Smith*, 02CR3477 (Colo. Dist. Ct. 2002) (Court ruled CBLA failed most Daubert criteria, that CBLA evidence could suggest to the jury an improper basis for its decision, and completely rejected CBLA); *United States* v. *Ronald Mikos*, 2003 WL 22922197, (N.D. Ill. 2003) (Court ruled that, notwithstanding that the FBI has been allowed to testify to CBLA for many years, there is no body of data to support an inference of same batch or melt); *United States* v. *Cornell Winfrei McClure*, Criminal No. DKC 01-0367 (D. Md. 2001) (Court observed that despite *stare decisis*, the '. . . so-called expert testimony both lacks sufficient foundation and, even if it did not, should be excluded because of the danger that a jury would misuse or misconstrue the opinion by giving it more weight than it deserves.'); *New Jersey* v. *Michael Behn* (N.J. Super. Ct. App. Div. 2005) (Court observed '. . . we note that the integrity of the criminal justice system is ill-served by allowing a conviction based on evidence of this quality, whether described as false, unproven, or unreliable, to stand.').

announcement does not preclude the FBI or another law enforcement agency, in the United States or elsewhere, from bringing the technique back into use in the future.[2] Finally, CBLA raised important conceptual, scientific, and statistical issues concerning how to assess the probative value of evidence. Continued discussion and research on CBLA may yet inform discussions of the probative value of other forms of forensic evidence that are still very much in use.

2. Background

CBLA is generally a technique of last resort when recovered bullet fragments are too damaged to perform more familiar methods of bullet analysis, such as associating bullets with a firearm through microscopic analysis by matching striations on crime-scene bullets to striations on bullets test-fired from a suspect firearm. Since approximately 1993, the analytical procedure of choice for CBLA has been inductively coupled plasma–optical (or atomic) emission spectrometry. Bullet fragment samples, typically from a crime scene, are dissolved in an acid solution and ionized, a process that generates a ‘chemical signature’ for the questioned bullet fragments. Of forensic interest are the amounts of several elements commonly found in bullet lead in trace quantities: arsenic, antimony, tin, copper, bismuth, silver and cadmium.

CBLA was used in efforts to link the suspect to the crime by analysing the similarity of the chemical signature of the crime-scene bullet fragments to the chemical signature of bullets in the possession of a suspect. The fundamental premise of CBLA is that similarity of the chemical signatures between the crime-scene fragments and the bullets in possession of the suspect purportedly provides evidence that the bullets are from the same ‘molten source’, pour or lot of bullets made by the same manufacturer on the same day.[3] This, in turn, indicated that the crime-scene bullet may have come from the suspect’s cache of bullets.

Clearly, such an inference demands the resolution of a number of empirical questions. Perhaps the fundamental questions concern the internal homogeneity of sources of lead and the degree of heterogeneity of the universe of sources of lead. (These, in turn, raise questions of how ‘chemical similarity’ is defined and measured.) Critics argued that the universal assumptions of internal homogeneity and absolute heterogeneity were empirically unsupported. They collected empirical data that undermined these assumptions (Randich *et al.*, 2002). These data were vigorously debated in the literature (Peters, 2002; Imwinkelried & Tobin, 2003; Tobin, 2004; Tobin & Duerfeldt, 2002), and they were the subject of extensive discussion in the NRC report (National Research Council, 2004). The report also discussed statistical techniques for defining CBLA matches, forensic significance of claimed matches and the proper presentation of CBLA expert testimony if admitted by courts. These issues are currently being debated in the literature (Finkelstein & Levin, 2005; Thompson, 2005; Kaye, 2005; Spiegelman & Kafadar, 2006). In contrast to the issues of the underlying compositional variability of bullet lots and statistical inference, the seemingly more prosaic issue of what we call *geographic concentration* in the distribution of bullets has been the subject of relatively

[2] The German *Bundeskriminalamt* uses CBLA only for investigative purposes and only then when there are multiple shooters each accusing the other. It does not use it as evidence of criminal guilt in trials. It developed this policy after lengthy study in part because of some of the concerns we raise here. Dr. Wilfried Stoecklein, BKA-KT1, e-mail to W. A. Tobin on 7 May 2003.

[3] The terms ‘lot’ and ‘pour’ are ill-defined among bullet manufacturers and secondary refiners (the sources of the lead for bullet manufacturers). In this paper, the term ‘lot’ will refer to all bullets included in a day’s production or specific manufacturing process (mechanical or thermal); the term ‘pour’ will be any subset quantity from the day’s production originating from the same pot or crucible of molten lead.

little debate in the scholarly literature. A review of the literature, however, makes it clear that this reticence is attributable to the lack of data on geographic distribution rather than to perceived lack of importance.

It is clear that the geographic concentration of bullet distribution could have an enormous impact on the probative value of CBLA evidence. If, to take an extreme case, all bullets of a given brand and caliber in a region of interest were derived from the same pour or lot of lead, then a finding of a match between a crime-scene bullet and a bullet in a suspect's possession would have little, if any, probative value. In a less extreme case, if most bullets of a given brand and caliber in a region of interest derive from the same pour or lot, the probative value of claimed matches would need to be adjusted accordingly. And, since the same pour or lot may be used to manufacture bullets of different *calibers*, it is quite possible that bullets of other calibers with the same composition exist in the same region.[4] Finally, scholars have suggested that a CBLA match might—counterintuitively—*even be exculpatory* when matching bullets are more common in the victim's neighbourhood stock of bullets than in the suspect's (Thompson, 2005; Kaye, 2005).

Most scholars and scientists who have studied CBLA recognize the potential importance of geographic concentration. Critics of the FBI's application of CBLA noted the potential importance of geographical distribution early on.[5] Tobin and Duerfeldt (2002, p. 33) suggested that the delivery of 'perhaps hundreds of thousands of bullets from the same molten source . . . to the same geographic region' could result in 'geographic concentrations of specific compositions'. The importance of geographic concentration has been acknowledged by the NRC report (National Research Council, 2004, p. 6), NRC Committee members in their personal scholarship (Finkelstein & Levin, 2005, p. 130; Spiegelman & Kafadar, 2006), critics of the NRC report (Thompson 2005), statisticians (Carriquiry *et al.*, 2000; Carriquiry, 2006; Stern, 2004) and statistically minded legal scholars (Kaye, 2005) who have studied CBLA, and proponents of CBLA (Koons & Buscaglia, 2005, pp. 2–3). Because of the lack of data on geographic distribution, statisticians who have attempted to generate models for estimating the probative value of CBLA matches left geographic concentration out of their models, even as they acknowledged that it might have enormous impact on the probative value of the evidence (Carriquiry *et al.*, 2000). The NRC Committee called 'the absence of information on the geographic distribution of ... bullets' one of the two principal factors that precluded expert witnesses from turning 'a conclusion that two bullets came from the same melt' into evidence that 'this fact increases the odds that the crime bullet came from the defendant' (National Research Council, 2004, p. 102).[6] Similarly, the lack of data on geographic concentration is a principal reason for scientists' and bullet manufacturers' inability to attest to the probative value of a CBLA association, cited by the FBI (2005) as the primary reason for the discontinuation of CBLA.

[4] This leads to the counterintuitive observation that bullet composition may frequently be less probative than caliber. It should also be noted that caliber itself can be expected to be subject to geographic concentration, especially given that caliber choice and turnover are quite dependent on game hunted, duration of hunting seasons, weather, urban versus rural market, state and local laws and other marketing considerations.

[5] In testimony in 1988, Dr. Vincent Guinn, CBLA pioneer and critic of its application by the FBI, testified as follows:

. . . I don't know the number of gun shops that stock cartridges, but there must be plenty of them. And all the people in that area in general are buying from those, and if any one of these gun shops receives a shipment of this material, then anybody who buys it in that area is going to have the same type. And if there's another gun shop three miles down the road and they're getting Winchester Westerns of this type at the same time, the same thing will happen. *Huffington* v. *Nuth*, post-conviction matter arising from *Maryland* v. *Huffington*, 140 F.3d 572; 500 A.2d 272 (Md. 1985); 452 A.2d 1211 (Md. 1982), Tr. Trans., 10 May 1988, at 371–372, on file with author Tobin.

[6] The second factor was the large number of bullets produced by a single melt.

An indication of the importance that the NRC panel afforded the issue of geographic concentration is the fact that it undertook efforts to obtain data on it, data that it described as potentially 'invaluable'. These efforts were ultimately unsuccessful. The NRC report states that this information 'either does not exist or is considered proprietary' (National Research Council, 2004, pp. 4–6). The NRC Committee responded to this lack of data by adopting an agnostic stance toward the question of geographic concentration of bullets. There was 'no evidence', the Committee opined, of geographic concentration (National Research Council, 2004, p. 82). But, neither did the Committee refer to any evidence against geographic concentration. One would imagine that, in the absence of any evidence either way, the Committee might have assumed that some geographic concentration exists. For bullet manufacturers, truly uniform geographical distribution of bullets would seem to be both difficult to achieve and counter to their economic self-interest. Instead, the NRC report simply set aside the potential impact of geographic concentration on the probative value of CBLA evidence.

3. Methods

We therefore have a complete absence of public data on bullet distribution coupled with nearly universal agreement that such data would shed important light on the probative value of CBLA evidence. Under these circumstances, the presentation of some measure of geographic concentration of bullet distribution, even a rough one, would inform scientific discussion of CBLA. This article reports the results of a study of packaging codes on bullet boxes on the shelves of retail bullet distributors. We use these packing codes as a measure of the degree of geographic concentration in the distribution of bullets and thus as a surrogate measure of the degree of geographic concentrations of bullet compositions.

Packing codes are serial alphanumeric identifiers stamped on the sides of bullet boxes by manufacturers. Although their significance varies by the conventions of the manufacturers (National Research Council, 2004, p. 84), generally speaking, identical packing codes on boxes of bullets are indicative of contemporaneous packaging. Packing codes are, therefore, reasonably correlated with specific, relatively contemporaneously produced, compositions. We call packing codes only a surrogate measure of chemical composition because packing codes are, *per se*, minimally indicators of a specific day's packaging, not (intentionally) of production. Therefore, they have only an indirect correlation to the chemical composition of the bullets within the boxes containing them. The degree of correlation varies according to manufacturer, among other parameters. For example, it is crucial to note that even within a single-box, distinct chemical compositions may be found (Peele, *et al.* 1991) due, in part, to the temporary inventory/staging practices of bullet manufacturers. The NRC report noted, 'It is *routinely* found that a single box contains multiple distinct compositional groupings—as many as 14' (National Research Council, 2004, p. 84, emphasis added). (Boxes typically contain 50 bullets.) However, it is also well documented that other boxes contain only one or two 'distinguishable' compositions. It should be expressly noted here, however, that because source composition homogeneity is largely unknown, different compositions do not necessarily indicate different molten sources, lots or pours, or even different days' productions. In sum, packing codes on boxes are likely to correlate loosely, but not necessarily tightly (depending on manufacturer), with the chemical composition of the bullets contained in those boxes.

The initial set of data reported in this article, consisting of packing codes on boxes of .22-caliber ammunition in Juneau, AK, was collected by one of the authors for purposes of litigation. It is important to note that the author was under no illusion that these data would allow for a definitive estimate of the value of a CBLA match in all locales and at all times. Rather, the impetus for the collection of the data were claims by FBI analysts, made during the course of litigation, implying that geographic concentration was not an issue of concern and would not have a significant impact on the probative value of a CBLA match. The author conceived the retail sampling study as a simple way to demonstrate the implausibility of the FBI's position.

A second set of data was collected for research purposes during three trips over 4 months to the same author's local Wal-Mart in Fredericksburg, VA. The Virginia site was visited on three different occasions over a 4-month period (in February, May and June). A third set of data, consisting of 00-buckshot shells in eastern Tennessee (Oneida, Cookville, Crossville, Jamestown and Byrdstown), was collected by a private investigator for litigation purposes. The 00-buck data were collected without the knowledge of the authors; the investigator subsequently provided the data to us. Treating these initial data collection efforts as pilot studies, we undertook a larger, though less comprehensive, data collection effort for scientific purposes in a separate region, selected by convenience. This generated a fourth set of data, consisting primarily of .22 but also a lesser amount of other caliber ammunition in Orange County, CA.

Researchers went to ammunition retailers and recorded detailed information from the boxes of bullets available for sale on the shelves of the retail outlets. The data obtained included the manufacturer's packing code, name of manufacturer, caliber, brand, 'feature' (such as magnum and hollow point), number of boxes, number of bullets per box, name and address of each retail outlet and date the information was obtained. The option of anonymity or acknowledgement was offered to each of the proprietors and outlets cooperating.

In Alaska, the retailers were 100% responsive (3/3). The data include all of the .22-caliber bullets available for retail sale in Juneau on the date of sampling. No response rate is associated with the data obtained from Virginia, inasmuch as only one outlet was targeted for data acquisition. Retailers in California were chosen as follows: Branches of a national discount store were visited, beginning with the branch closest to the University of California, Irvine, and working outward. A major regional sporting goods retailer was chosen in the same fashion, and independent retailers were selected using local yellow pages.

Undergraduate research assistants requested access to ammunition shelves in the sampled retail outlets. This approach generated moderate success. For various reasons, many stores declined researchers' requests. Generally, the explanations were that the researchers might interfere with customers, that managers were wary because the request concerned ammunition or that the store was too busy to spare an employee to supervise and assist the researchers. At the other end of the spectrum, some retailers were enormously cooperative. In California, the response rate was 37% (7/19). The lower response rate in California, relative to those in Alaska and Virginia, may have been due to the personalized approach taken by professional investigators or scientists, as opposed to the more casual, possibly less savvy, approach taken by undergraduate assistants.

The two most cooperative California retailers were sampled a second time, approximately 4 months after the initial sampling. We use these second samples to address consistency of geographic concentrations over time (in Section 4.1) and to measure the likelihood of cross-time packing code matches (in Section 4.3).

Sampling was comprehensive as to available (displayed or behind the counter) ammunition. In other words, researchers recorded the packing codes on *all* boxes of bullets on the shelves for a particular caliber. If all packing codes for a particular caliber could not be recorded, the data were considered incomplete and not included in the data presented here. Researchers initially sampled

.22-caliber ammunition and then sampled larger calibers after completion of the .22-caliber display inventory. In some cases, it was possible to gather data on only one or several calibers before the patience of the retailer reached its limit. Accordingly, the majority of the data in our study is that of .22-caliber ammunition, which also represents the highest turnover caliber marketed (Imwinkelried & Tobin, 2003; Randich *et al.*, 2002; Tobin, 2004; Tobin & Duerfeldt, 2002; National Research Council, 2004, p. 72).

Our data collection sites were therefore generated by convenience, and no effort was made to ensure that they were representative of either the country as a whole or of specific regions individually for any period other than the ones under study. The first three sources of data were convenience samples because they were not undertaken solely for scientific purposes. The fourth data collection effort was limited to a single convenient site due to limited resources. As we will discuss below, our findings do not depend on the 'representativeness' of our sites, but future researchers may want to undertake more representative sampling.

4. Findings

4.1 *Geographic concentration*

The findings presented here are based on date–store–caliber–manufacturer–feature combinations.[7] Our statistical approach first considers each box of bullets in turn and derives the probability of finding a packing code match for that box. To understand the calculation, imagine *a particular box of bullets* in the data set as being purchased by Customer A. Then consider Customer B, who enters the same store on the same day seeking to purchase bullets of the same caliber by the same manufacturer under the same brand name with the same 'feature' (e.g. magnum and jacketed hollow point). We can calculate the probability that Customer B would purchase a box of bullets containing the same packing code that appeared on the box purchased by Customer A. This is just the number of boxes with that packing code remaining divided by the total number of boxes remaining. We carry out this calculation for each box of bullets in the data set. In our calculations, the box purchased by Customer A has been removed from the shelf and is not available to Customer B. Therefore, date–store–caliber–manufacturer–feature combinations that represented only a single box of bullets were eliminated from our calculations.

To illustrate the calculation consider the data for one combination (data collected on 29 February from 'Store X' in Orange, CA, for .22 WMR (Winchester Magnum Rimfire) bullets). There were nine boxes of bullets for this particular combination with one box having packing code E24E06, five boxes having packing code L13E05, and three boxes having packing code M13E10. For the one box with packing code E24E06, it is not possible to obtain a matching box so the match probability is zero. For each of the five boxes having packing code L13E05 we know that if Customer A has bought this box, then four of the eight remaining boxes have a matching packing code. Thus, the probability of a match is 0.50 for these boxes. Finally, for each of the three boxes having packing code M13E10 we note that if Customer A buys such a box, then two of the eight remaining boxes have matching packing codes. For these three boxes the probability of a match is 2/8 or 0.25.

The four panels in Fig. 1 show the distribution of probabilities generated by this approach for .22-caliber bullets in California (Panel b), Alaska (Panel c), Virginia (Panel d) and total sample (Panel a).

[7] Our approach analysing the data in terms of such combinations is supported by the high degree of brand loyalty among bullet purchasers (Petty, 2003).

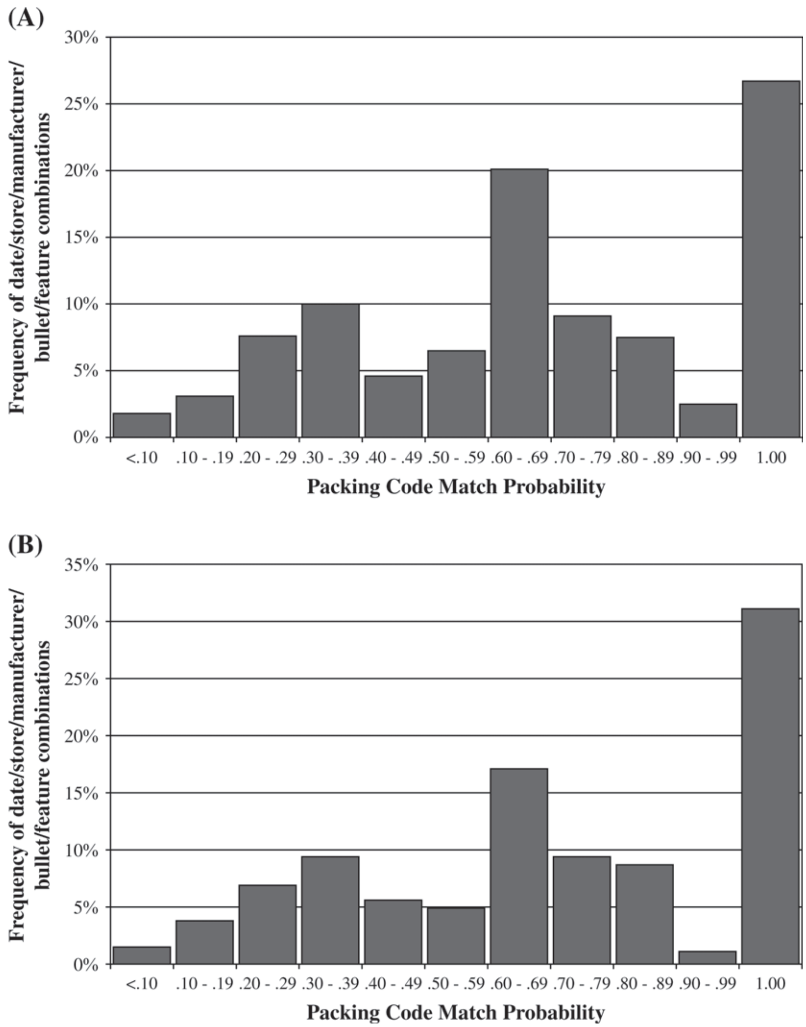

FIG. 1. Match probabilities for boxes of .22-caliber bullets in (A) total sample ($n = 3189$), (B) California ($n = 2479$), (C) Alaska ($n = 494$) and (D) Virginia ($n = 216$).

Each box contributes one probability measurement to the relevant figure. If the probability is 1 for a particular box, then all boxes with the given caliber–manufacturer–feature in that store on that day share the same packing code. For such cases Customer B has no choice but to buy a box of bullets bearing the same packing code as the box purchased by Customer A. As these panels make clear, packing codes are highly geographically concentrated. The modal probability of one of the sites is 1. For the other two, the modal probabilities are greater than 0.5.

One possibly undesirable feature of Fig. 1 is that a large sample (i.e. a large number of boxes) in a particular store–caliber–manufacturer–feature combination would lead to many repeated values of the same probability. For example, a combination which has 80 boxes of a single packing code

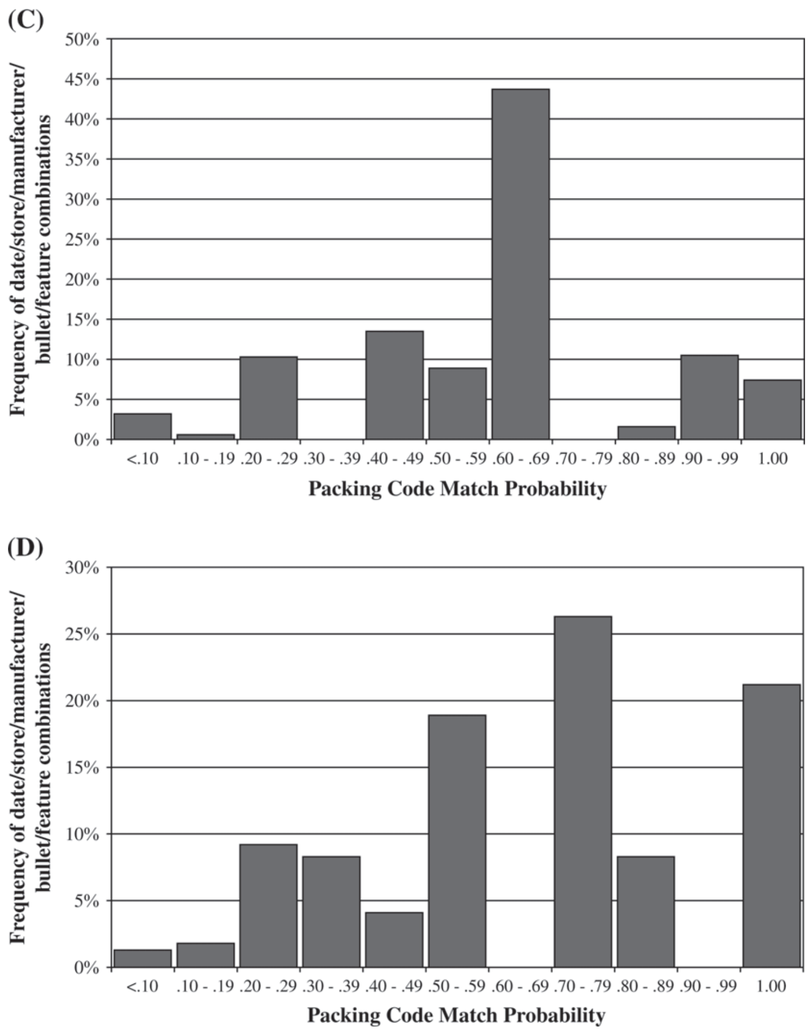

FIG. 1. Continued

is reported as 80 box probabilities equal to one. We can summarize the degree of concentration of packing codes in a particular combination by reporting the average probability for the boxes (of the given date–store–manufacturer–feature combination) as a *weighted average packing code match probability*. This probability, which turns out to be equivalent to a quantity measuring the likelihood of a coincidental match in Carriquiry *et al*. (2000), is just the probability that two boxes of a given date–store–caliber–manufacturer–feature combination will share the same packing code. The weighted average packing code match probability is thus a number between 0 and 1 that characterizes all the data for a particular date–store–caliber–manufacturer–feature combination (rather than just a single box). In our example above, the average match probability for the nine boxes is $(0.00 + 0.50 + 0.50 + 0.50 + 0.50 + 0.50 + 0.25 + 0.25 + 0.25)/9 = 0.36$. The higher the number, the more likely

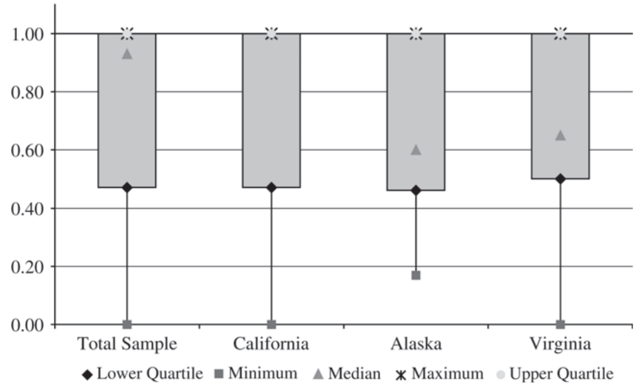

FIG. 2. Boxplots of the concentration of .22-caliber bullets, all sites.

TABLE 1 *Measure of concentration† for .22-caliber bullets, all sites‡*

	n	Mean	Standard deviation	Concentration score of 1 (%)
California	138	0.73	0.31	31.1
Virginia	18	0.71	0.30	21.2
Alaska	17	0.71	0.29	7.4
Total sample	173	0.73	0.30	26.7

†Concentration ranges from 0 to 1. The closer the score is to 1, the more concentrated the sample (i.e. more box codes are the same and the probability of getting two boxes with matching codes is higher).
‡By date–store–caliber–manufacturer–feature combinations.

Customer B is to buy a box of bullets bearing the same packing code as Customer A. Figure 2 shows the probabilities for each date–store–caliber–manufacturer–feature combination for each site and overall.

By examining the weighted average packing code match probabilities in a given region, we can generate summaries that give an idea of the degree of concentration of bullet box packing codes in each region at the time of data acquisition. Table 1 shows the mean probabilities for each site and overall. Again, the geographic concentration hypothesis is supported. The median probabilities are quite high: 1 in California, around 0.6 in Virginia and Alaska and 0.93 overall (a figure which is heavily influenced by the greater size of our sample in California). This means that two purchases of a particular manufacturer–caliber–feature combination from the same store on the same day will typically share the same packing code in California, and have a good chance of sharing the same packing code in Virginia and Alaska.

We also calculated the weighted average packing code match probability for .38- and .357-caliber bullets in California and shot shell data collected from Tennessee. Figure 3 shows these probabilities. Table 2 shows the mean probabilities for the non-.22-caliber samples. Again, we see quite a high concentration.

It should be noted that our data represent snapshots of geographic concentrations in particular stores at arbitrarily chosen times. The actual degree of geographic concentration may be expected to

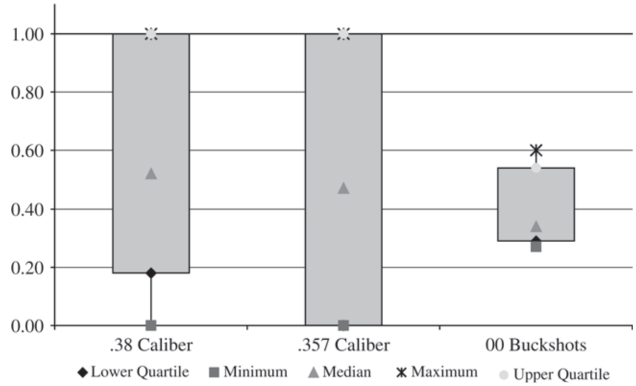

FIG. 3. Boxplots of concentration of other caliber bullets.

TABLE 2 *Concentration† for other bullet calibers, California and Tennessee‡*

	n	Mean	Standard deviation	Concentration score of 1 (%)
.38 Caliber§	53	0.58	0.41	39.6
.357 Caliber	30	0.52	0.44	40.0
00 Buckshot	24	0.39	0.14	0.0

†Concentration ranges from 0 to 1. The closer the score is to 1, the more concentrated the sample (i.e. more box codes are the same and the probability of getting two boxes with matching codes is higher).
‡By date–store–caliber–manufacturer–feature combinations.
§The .38- and .357-caliber bullets were sampled in California; the 00 buckshots were sampled from Tennessee.

vary at different times. We used the second samples of two California retailers to address this point. Figure 4 shows the weighted average packing code match probabilities for the two twice-sampled stores in California on different sampling dates, and Table 3 shows the average probability for each store on each sampling occasion. Although all the samplings show high concentrations, the overall distribution of probabilities are generally higher on the later dates. In other words, packing codes may be more or less concentrated on different dates in the same store, and the probative value of a CBLA 'match' is highly time sensitive.

We also note that in both stores the packing codes were less concentrated on the earlier dates—the sampling dates that were included in our overall data set discussed above. We note this not to suggest a temporal trend, but rather to emphasize that we are using the data least favourable to our own hypothesis. Therefore, our data may *underestimate* the geographic concentration of packing code distribution.

4.2 *Cross-town and cross-region matches*

Thus far, we have measured geographic concentration according to repeated packing codes occurring within a single store. A further issue concerns the probability of purchasing boxes of bullets bearing

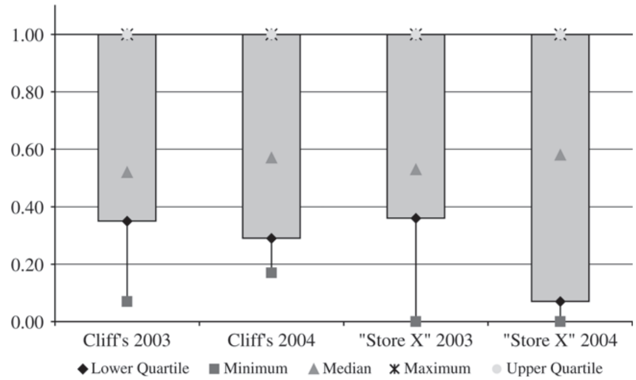

FIG. 4. Boxplots of concentration of .22-caliber bullets for repeated visits, Cliff's Gunsmithing and 'Store X', California.

TABLE 3 *Concentration† for repeated visits, Cliff's Gunsmithing and Store Y, California‡*

	n	Mean	Standard deviation	Concentration score of 1 (%)
Cliff's 2003	16	0.60	0.32	31.3
Cliff's 2004	21	0.62	0.32	33.3
Store Y 2003	24	0.62	0.34	37.5
Store Y 2004	21	0.59	0.43	47.6

†Concentration ranges from 0 to 1. The closer the score is to 1, the more concentrated the sample (i.e. more box codes are the same and the probability of getting two boxes with matching codes is higher).
‡By date–store–caliber–manufacturer–feature combinations.

identical packing codes in different stores in the same region or town. We call these *cross-town matches*. Two of our sites, Alaska (.22 caliber) and the Tennessee sampling (00 buckshot) showed cross-town matches. California showed almost no cross-region matching, and in Virginia we had insufficient data to measure cross-town matching because only one store was sampled. The finding of greater cross-region matching in Alaska and Tennessee than in California conforms to our expectation that cross-region matching would be higher in rural areas with fewer transportation channels into them than in larger, more metropolitan areas. Indeed, it is difficult to imagine an American city with a narrower transportation network than Juneau, which is solely accessible by water or air (McPhee, 1977, p. 103). Therefore, it seems more likely that bullet boxes intended for neighbouring stores in Juneau might be delivered in the same shipment than would be the case for neighbouring stores in Orange County, CA, which is served by a variety of transportation channels. If the boxes are delivered in the same shipment, this presumably raises the likelihood that they were manufactured contemporaneously and bear identical packing codes.

4.3 *Cross-time matches*

The above discussion assumes customers shopping in the same store, or the same town or region, on the same date. An additional consideration is the extent to which customers shopping on different

TABLE 4 *Probability† of coincidental match, twice-sampled stores‡*

	n	Median	Mean	Maximum	Standard deviation
Cliff's Gunsmithing	27	0.0	0.09	0.78	0.20
Store Y	36	0.0	0.09	1.00	0.25

†Probability ranges from 0 to 1. The closer the score is to 1, the more likely it is that customers purchasing a box of bullets at time 2 (March 2003) would purchase a box with the same packing code as time 1 (November 2004).
‡By caliber–manufacturer–feature combinations.

dates may yet purchase boxes of bullets bearing identical packing codes. As noted above, we returned to the same outlet in Virginia, three times over a 4-month period. The investigator found that some packing codes found on both the first and second visits also appeared on the third visit (in the fourth month of the study). Since this finding raised the possibility of *cross-time matches*, we used the second samples obtained from two of the stores in the California data set approximately 4 months after the first sampling to measure the likelihood of such matches.

For these two stores, we computed the probability that a customer with a particular store–caliber–manufacturer–feature preference purchasing a box of bullets in November would purchase a box bearing the same packing code as a customer with the same preference making a purchase in March. For each store–caliber–manufacturer–feature combination, we multiplied the number of boxes available in November by the number available in March to give the number of possible pairs. We then divided the number of pairs that yield a coincidental packing code match by the number of possible pairs. Like our measure of geographic concentration (see Section 4.1), this process yields a single-number measure of the likelihood of cross-time matches for each store–caliber–manufacturer–feature combination.

Table 4 shows all of these probabilities (across caliber–manufacturer–feature combinations) for each of the two stores. The probabilities of cross-time matches, over the arbitrary span of 4 months, are fairly low: zero for more than three-quarters of the caliber–manufacturer–feature combinations for both stores. But the probability of cross-time match is not entirely negligible. In the case of one store–caliber–manufacturer–feature preference, a cross-time match is unavoidable. It should be noted that these measurements of the likelihood of cross-time matches concern the caliber of ammunition (.22 caliber) that is known to have the highest turnover. In other words, we have presented data on the caliber most favourable to proponents of CBLA. For ammunition with lower turnover, such as shot shells, the potential for cross-time matches may have an even greater impact on probative value.

5. Discussion

We wish to emphasize the modesty and limited scope of our claims. CBLA expert witnesses have implied that bullet lots are uniformly distributed throughout the United States.[8] The NRC chose to assume that no geographic concentration exists (National Research Council, 2004). And, scholars have acknowledged that it may well exist but has not been measured (Carriquiry *et al.*, 2000). Since

[8] Expert witness in *State* v. *Jose (Che) Mateu*, 1JU-D03-58B; 1KE-00-869CR (D. Alaska 2003), Tr. Trans. April 17 at 10:13:22 in audio transcript, on file with author Tobin (indicating that packaged boxes of 50 rounds are put into a brick and 'distributed throughout the United States').

our aim is merely to falsify the assumption of uniform distribution, we need not show that geographic concentration exists or is significant in all places, at all times for all CBLA issues. We merely need show only that geographic concentration exists in one place, at one time, for one type of bullet to show that geographic concentration must be considered when calculating the probative value of a CBLA match.

Even in light of the indirect nature of our data, this study confirms other scholars' suspicions that geographic concentration is a significant issue worthy of consideration in assessing probative value. For approximately one quarter of the boxes of bullets sampled, the probability that another customer seeking bullets of the same caliber by the same manufacturer in the same store on the same date ends up with a box having an identical packing code is 1. In such cases, our putative second consumer has *no choice* but to purchase bullets with the same packing codes as our putative first consumer. For more than half the boxes of bullets sampled, the probability that a second consumer ends up with an identical packing code is more than 0.5.

The legal implications of these findings are quite serious. In most crimes including homicides, it is generally accepted that most perpetrators are 'local' (Tita & Griffiths, 2005).[9] In many cases, therefore, the chances that an innocent suspect may have bullets analytically indistinguishable from those used to perpetrate a crime may be quite high. Or, put another way, the probative value of a finding of a match between bullet fragments recovered from a crime scene and bullets seized from a suspect who lives in the neighbourhood in which the crime was committed may be significantly undermined, may be completely eroded or may even become counter-probative by geographic concentration.

There are several potential biases associated with our use of packing codes as surrogate measures of bullet composition. For instance, the probabilities of innocent purchases with matching packing codes that we have calculated may actually underrepresent the true probabilities because they implicitly assume that different packing codes represent different molten source or lot compositions. This implicit assumption is certainly incorrect. Different packing codes from the same manufacturer frequently contain indistinguishable compositions, in part because a 'melt' is not all used within a single day's packaging of bullets at the bullet manufacturer, and packing codes usually represent discrete dates and shifts (and usually caliber) of bullet packaging. The Randich *et al.* (2002) study has revealed frequent adventitious production of compositions considered 'analytically indistinguishable' or 'similar'. On the other hand, this effect may, to some extent, be offset by the fact that boxes with the same packing code usually have subgroups of bullets with discernible compositions, which would lead us to overrepresent the likelihood of innocent purchase.[10]

[9] This is, of course, a broad generalization that depends on the type of crime observed and how the notion of 'local' is defined.

[10] Though we acknowledge that our method of treating packing codes as surrogate measures of chemical composition could bias our results in either direction, we think that it is more likely that our study biases the results in favour of the proponent of CBLA evidence for several reasons.

First, the effects of compositional 'repeats' (different melts that produce indistinguishable compositions) and 'overlaps' (subregions of different melts that produce indistinguishable compositions) probably outweigh the effects of each packing code containing multiple compositions. Second, whenever possible, decisions with regard to data treatment were resolved in the light most favourable to CBLA proponents. For example, identical packing codes for Remington and for 'Thunderbolt' bullets were treated as separate entities even though the Thunderbolt brand is manufactured by Remington, and we did the same with Federal and for CCI brands.

Third, there are several scenarios in which compositionally indistinguishable bullets might exist and yet not be detected using our surrogate measure. Boxes from *different manufacturers* (which necessarily would have different packing codes)

To fairly and accurately assess the probative value of a claimed CBLA match, it would be necessary to understand not only the degree of geographic concentration that existed for the locale in which the crime-scene bullet was purchased for the caliber of bullet used, for the brand of bullet used, *at the time the bullet was purchased*, but also the incidence of identical composition bullets available in the area both prior to, and subsequent to, purchase of the crime-scene bullet. Thompson (2005) suggests that CBLA evidence might only be used 'On a proper showing that the prevalence of matching bullets in the defendant's stock of bullets is substantially higher than the prevalence of matching bullets in the surrounding community.'

It has often been suggested that it would be useful to create a national database of bullet compositions. The argument is that such a database might be used to calculate relative frequency determinations for probative significance of purported matches. Our findings, however, indicate that the creation of such a database would be unlikely to solve the underlying problem of fairly and accurately assessing the probative value of a CBLA match. It should be obvious at this point in our discussion that probative value or significance should be derived from relative frequency assessments *locally* rather than nationally, a very dynamic parameter, indeed. Even if an accurate estimate could be obtained for a local region based on any such study, it would change with time and still only constitute a *minimum* frequency because of the phenomenon of adventitious production ('repeats') of bullet lead compositions at the smelting 'sources' and in some bullet manufacturers' remelting operations shown to exist by Randich *et al*. (2002).

6. Conclusion

As noted above, the data reported here are not necessarily representative of any greater population. Sites were selected by convenience and not randomized. Further studies that select sites more deliberately would be informative. Although our convenience sample includes rural and suburban demographic information in disparate regions of the country, future researchers might wish to vary their sites along a variety of significant dimensions.

Dimensions describing a particular geographic region (weather, urban/rural, transport) or the population in that region (affluence, education, age) may be worth considering in designing later studies. For example, weather can play a role in that it may serve to limit or encourage outdoor shooting activities, reducing or increasing demand for ammunition locally and, thus, retail turnover. There are almost certainly seasonal demand variations in ammunition sales, in part because of local and state hunting regulations. Existence and length of hunting seasons, as well as type of game hunted, may be factors affecting product turnover and shelf-life exposure. Regional affluence is

may yet contain bullets with indistinguishable compositions. And, bullets of *different caliber* are commonly comprised of indistinguishable compositions, since bullet manufacturers (e.g. American Ammunition and Remington) are known to make up to 15 different calibers of bullets from the same pallet of source lead (ingots) (Tobin, 2004).

We believe that the major variables that almost certainly affect whether our estimates constitute underestimates or overestimates are the specific production and inventory practices of each bullet manufacturer. At one extreme, Federal's product typically has only one or two different compositions per box. At the other, Winchester's ammunition boxes often have 10–12 analytically distinguishable compositions. However, even this broad generalization can be further complicated. Because Winchester boxes tend to be so heterogeneous, compositionally indistinguishable bullets are distributed across a correspondingly larger number of boxes, bearing a correspondingly larger number of different packing codes, appearing on retail shelves over a correspondingly longer period of time! This effect may be expected to *broaden* (geographically and temporally) availability of an incriminating chemical composition to the 'innocent' population—i.e. the population purchasing or possessing bullets that are compositionally indistinguishable from a putative crime-scene bullet.

likely a factor. It seems intuitively correct that shooting and hunting activities are less prevalent in affluent areas of Connecticut and California, than in most areas of Alabama, Mississippi, Tennessee, Kentucky, etc. Similarly, the urban versus rural nature of a region also likely influences bullet product turnover, particularly in places where local laws restrict the discharge of a firearm within jurisdictional limits. Education level may be an influence on choice of outdoor activity, and the demography of age may also play a role in hobbies and outdoor shooting sport preferences. Regional transport accessibility is certainly a major consideration in retail distribution and shipping practices, which in turn can affect product inventory stockpiling, shelf availability, and turnover. The high geographical concentration evident in the Alaska data is possibly primarily the result of transport accessibility issues.

Caliber preference might also be an important issue. Caliber and product turnover rates are certainly factors in assessing probative value. Caliber can, depending on manufacturer, affect both projectile composition and inventory stockpiles at retail outlets. The highest turnover caliber of ammunition used by consumers is the .22-caliber family of bullets. It is much more likely that a higher percentage of bullets manufactured from a specific molten source will be geographically concentrated for .22-caliber bullets than for shotgun shells, for example. Aside from the high-frequency use (and turnover) attributed to the .22 family of calibers, large-scale merchandisers such as Wal-Mart and K-Mart purchase or contract for very large quantities (the two combined reportedly acquired over half the product manufactured for retail sale during one survey and have limited and localized regional storage and distribution centres, increasing the propensity for regionalized concentrations of specific compositions). Conversely, the lower use of shotgun shells allows for a greater range of geographic distributions, based on subpoenaed records reviewed by one of the authors, but would also be marketed with less selection because of increased cost and diminished use.

The data presented here support various scholars' emphasis on the critical importance of knowledge about the local population of bullets for any responsible effort to estimate the probative value of any particular CBLA match. Thompson (2005) e.g. has shown that the probative value of a CBLA match is highly sensitive to the composition of bullets in local retail outlets. Indeed, he shows, counterintuitively, that under some conditions, a CBLA match may be 'counter-probative'. Specifically, when the defendant's stock of bullets contains fewer 'matching' bullets than the surrounding a neighbourhood's population of bullets, a CBLA match 'should make the trier-of-fact *less* confident that the crime scene bullet came from the defendant (compared to having no bullet-lead evidence), even though the defendant owned one or more "matching" bullets' (Thompson, 2005, p. 81, original emphasis). Much the same point is made by Kaye (2005).

Recall that the FBI finally discontinued the use of CBLA because 'neither scientists nor bullet manufacturers are able to definitively attest to the significance of an association made between bullets in the course of a bullet lead examination' (Federal Bureau of Investigation, 2005). Our findings suggest that this was not an unreasonable decision. What is surprising is that it took nearly 40 years to make the appropriate decision. We also find it surprising that it has taken nearly 40 years to begin the empirical study of bullet distribution.

Acknowledgements

Special thanks to the invaluable contribution of Mr. H. H. 'Tres' Lewis III, Mendenhall Investigations, who 'silver-tongued' our way to 100% retailer cooperation in our data collection efforts for

Alaska. For research assistance at the California site, we are grateful to Ana A. Jimenez, Rachel L. Lee, Robert S. Liu, Vivian H. Mirabueno, Erich S. Reyburn, Vanessa L. M. Yap, and especially Tina Wan and Dana M. Engstrom. We are grateful to the helpful employees of Cliff's Gunsmithing and other cooperating retailers who preferred to remain anonymous. For comments and assistance on drafts of this paper, we are grateful to George Tita, William C. Thompson, Ryan Fischer, David Kaye, and Carroll Seron. This material is based in part upon work supported by the National Science Foundation under Grant No. 0347305. Responsibility for any errors rests with the authors.

References

CARRIQUIRY, A. L. (2006) Good riddance: further arguing against CABL as a forensic tool. *Chance*, **19** (in press).

CARRIQUIRY, A. L., DANIELS, M. & STERN H. S. (2000) Statistical treatment of class evidence: trace element concentrations in bullet lead. Report prepared for the Federal Bureau of Investigation, Department of Statistics, Iowa State University, Ames, IA. Available at www.public.iastate.edu/alicia.

FEDERAL BUREAU OF INVESTIGATION (2005) FBI Laboratory announces discontinuation of bullet lead examination. Available at http://www.fbi.gov/pressrel/pressrel05/bullet_lead_analysis.htm.

FINKELSTEIN, M. O. & LEVIN, B. (2005) Compositional analysis of bullet lead as forensic evidence. *Journal of Law Policy*, **13**, 119–142.

HANSEN, M. (2004) Bullet proof. *ABA Journal*, **Sept.**, 58–63.

IMWINKELRIED, E. J. & TOBIN, W. A. (2003) Comparative bullet lead analysis (CBLA) evidence: valid inference or ipse dixit? *Oklahoma City University Law Review*, **28**, 43–72.

KAYE, D. H. (2005) Bullet-proofing the NRC bullet-lead report: should science committees make legal findings? *Jurimetrics*, **46**, 91–105.

KOONS, R. D. & BUSCAGLIA, J. (2005) Forensic significance of bullet lead compositions. *Journal of Forensic Science*, **50**, 1–11.

MCPHEE, J. (1977) *Coming into the Country*. Farrar, Straus & Giroux, New York.

MEJIA, R. & SAMPLE, I. (2002) Bite the bullet. *New Scientist*, **Apr. 20**, 4.

NATIONAL RESEARCH COUNCIL (2004) *Weighing Bullet Lead Evidence*. The National Academies Press, Washington, DC.

PEELE, E. R., HAVEKOST, D. G., PETERS, C. A., RILEY, J. P., HALBERSTAM, R. C. & KOONS, R. D. (1991) Composition of bullets using the elemental composition of the lead component. *Proceedings of the International Symposium on the Forensic Aspects of Trace Evidence*; 24–28 June; Quantico, Virginia.

PETERS, C. A. (2002) The basis for compositional bullet lead comparisons. *Forensic Science Communications*, **4**, 3. Available at http://www.fbi.gov/hq/lab/fsc/backissu/july2002/peters.htm.

PETTY, C. E. (2003) Prime your ammo & reloading sales. *Shooting Industry*, **Jan.** Available at http://www.findarticles.com/p/articles/mi_m3197/is_1_48/ai_97232944.

PILLER, C. & MEJIA, R. (2003) Science casts doubt on FBI's bullet evidence. *Los Angeles Times* (Feb. 3), pp. A1–A17.

RANDICH, E., DUERFELDT, W., MCLENDON, W. & TOBIN, W. (2002) A metallurgical review of the interpretation of bullet lead compositional analysis. *Forensic Science International*, **127**, 174–191.

SPIEGELMAN, C. & KAFADAR, K. (2006) Data methods, materials, and scientific responsibility. *Chance*, **19** (in press).

STERN, H. S. (2004) *The Bullets Match. But who Cares? Decisions and Justice*. Institute for Mathematical Behavioral Sciences, University of California, Irvine.

THOMPSON, W. C. (2005) Analyzing the relevance and admissibility of bullet-lead evidence: did the NRC report miss the target? *Jurimetrics*, **46**, 65–89.

TITA, G. & GRIFFITHS, E. (2005) Traveling to violence: the case for a mobility-based spatial typology of homicide. *Journal of Research on Crime and Delinquency*, **42**, 275–308.

TOBIN, W. A. (2004) Comparative bullet lead analysis: a case study in flawed forensics. *Champion*, **27**, 12–22.

TOBIN, W. A. & DUERFELDT, W. (2002) How probative is comparative bullet lead analysis? *Criminal Justice*, **17**, 26–34.