AFFIDAVIT OF PROFESSOR MICHAEL J. SAKS

I, MICHAEL J. SAKS, declare as follows:

Affiant's Background

1. At the Arizona State University, I am a Regents Professor on the faculty of the Sandra Day O'Connor College of Law and a fellow of the Center for Law, Science and Innovation, with a further appointment to the faculty of the Department of Psychology. Previous full-time academic appointments include Boston College, Georgetown University Law Center, University of Iowa, and Ohio State University. My doctoral training in experimental social psychology emphasized research methodology and statistical analysis. One of my areas of professional interest has been the law's use of scientific evidence, which I have taught for nearly 30 years to law students, law professors, attorneys, and judges (at the University of Virginia's LL.M. program for appellate judges, the Duke University School of Law's "Judging Science" program, and at the University of Nevada at Reno). I have published over 200 articles and books, including original empirical research as well as legal and scientific commentary, on numerous topics, including forensic science. Among these is the treatise MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (5 volumes, West, updated annually), which I have co-authored and co-edited with David Faigman, David Kaye, Joseph Sanders, and Edward Cheng, in which the chapters on research methodology and forensic science, among others, have been primarily my editorial responsibility. I also co-edited the ANNOTATED REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (the annotated version of the Federal Judicial Center's volume). I have served as the editor of JURIMETRICS (the journal of the ABA Section of Science and Technology). I am a member of the National Conference of Lawyers and Scientists, a joint committee of the American Association for the Advancement of Science and the ABA. I have been invited to speak on forensic science issues to conferences sponsored by the FBI, the International Association for Identification, the American Academy of Forensic Sciences, the International Conference on Forensic Statistics, and the Australia-New Zealand Forensic Science Society, as well as to numerous conferences held by and for academics, lawyers, and judges. My CV accompanies this affidavit.

The Exaggerated Claim of "Identification to the Exclusion of All Others in the World"

2. In his testimony during Petitioner's trial, firearms examiner Robert Sibert made a number of statements which asserted that each and every weapon imparts unique microscopic markings onto the bullets that pass through them, and that as a result of that supposition it is possible to identify the one and only gun in the world that fired any

particular questioned bullet (or cartridge casing). To quote some of those testimonial assertions:

> "When the examiner makes an identification this means it is his opinion that the bullet or the cartridge case was fired by the particular weapon he examined and could have been fired by no other weapon in the world." (Sibert Testimony, 12/20/1976 Transcript, p. 3:5-8)
>
> "Q: Will you tell us what is peculiar about firearms in that particular field that makes this identification possible?
> A: Yes, sir. Well, no two firearms are the same, microscopically." (Sibert Testimony, 12/20/1976 Transcript, p. 3:11-15)
>
> "In the case of cartridge casings, you have individual markings on the tip of the firing pin, on the breach face of the weapon – this is the portion of the slide which is behind the cartridge case. You have the extractor and ejector markings, and these little microscopic imperfections are random. They are unique to each individual weapon and will be reproduced by no other weapon." (Sibert Testimony, 12/20/1976 Transcript, p. 4:6-13)
>
> "These would be different because each of these barrels, although they appear to be the same, have different random microscopic imperfections in them which leaves a unique signature... on the bullets which are fired from them." (Sibert Testimony, 12/20/1976 Transcript, p. 15:10-14)
>
> "[Because] no two weapons will leave the same markings on a fired cartridge case, [the ejector markings on Exhibits 28 and 29] could have been produced by no other ejector in the world." (Sibert Testimony, 12/20/1976 Transcript, p. 41:2-3)
>
> "Q: [I]n your opinion could that [marking] have been produced by... any other ejector... in the world?
> A: No, sir, having identified those ejector marks as having been produced by this ejector, this would be to the exclusion of all other ejectors in the world."
> (Sibert Testimony, 12/20/1976 Transcript, p. 65:14-19)

3. In closing argument, counsel for the government repeatedly emphasized these assertions of unique individualization, for example: "no ejector in the world but this one made those marks on those shells," (Closing Argument, 1/3/77 Transcript, p. 22:21-22) and "no other ejector in the world but that one and that gun made those marks on those shells." (Closing Argument, 1/4/1977 Transcript, p. 44:21-23).

4. Such statements assert as fact something that has never been established as true on the basis of any kind of empirical testing or rigorous theory. These statements are more akin to folklore – never demonstrated to be true, unsupported by any scientific testing, but believed nonetheless.

5. Assertions require evidence to support their truth. And extraordinary claims require extraordinary evidence to prove their truth. Mere assertion, mere *ipse dixit*, falls well short of supporting the asserted conclusion. Put simply, *Sibert's testimonial claims of unique individualization assert far more than his field could then or can now support with any serious scientific evidence.*

6. Though lacking any scientific justification, for many years firearms examiners accepted the extreme claims passed down to them by their firearms examiner forebears, and passed those along to the next generation of firearms examiners – as well, of course, to judges and juries. Recently, forensic scientists as well as mainstream scientists have made clear the lack of any sound basis for such claims, and the firearms field has begun to back away from them.[1] Perhaps the most authoritative rejection of those claims came in 2009 from the National Academy of Sciences[2] which, pursuant to a Congressional directive, reviewed various forensic sciences and concluded that the claimed ability to identify the one and only source of a crime scene marking (of a fingerprint, a writing, on a bullet, etc.) had not been established for most areas of forensic identification, including firearms:

> "Much forensic evidence including, for example, bite marks and firearm and tool mark identifications is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing...." (NAS Report, p. 108.)
>
> "With the exception of nuclear DNA analysis,...no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source." (NAS Report, p. 7.)
>
> "Forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem." (NAS Report, p. 53.)

[1] In this affidavit I will recount some of that analysis. See Michael J. Saks & Jonathan J. Koehler, The Individualization Fallacy in Forensic Science, 61 Vanderbilt L. Rev. 199 (2008).

[2] National Research Council of the National Academies of Science, Committee on Identifying the Needs of the Forensic Science Community, Strengthening Forensic Science in the United States: A Path Forward (2009) [hereinafter, NAS Report]. For a brief overview of the Report, see Michael J. Saks, The Past and Future of Forensic Science and the Courts, 93 Judicature 94 (2009).

7. An earlier report from a different committee of the same National Research Council, addressed to the question of whether the capability existed to create a national firearms identification database, observed:

> "[E]xaminers tend to cast their assessments in bold absolutes, commonly asserting that a match can be made 'to the exclusion of all other firearms in the world.' Such comments cloak an inherently subjective assessment of a match with an extreme probability statement that has no firm grounding and unrealistically implies an error rate of zero."[3]

8. In his keynote address to a major interdisciplinary conference discussing the NAS Report, held at ASU a few months after its release, Judge Harry Edwards, co-chair of the committee which wrote the report, explained that they looked hard for scientific support for the claims brought to court and found little if any:

> "Not only were we trying to understand how the forensic science disciplines operate, we were also trying to determine the extent to which there is any... scientific research to support the validity and reliability of existing forensic disciplines; in particular, we were looking for scientific studies that address the level of accuracy of forensic disciplines that rely on subjective assessments of matching characteristics. We invited experts in each discipline to refer us to any such research; however, apart from the materials on nuclear and mitochondrial DNA and drug analysis, we received little in the way of compelling scientific research assessing the accuracy of forensic science disciplines."[4]

9. Assertions of pinpoint, unique identification are starting to be abandoned by forensic identification scientists, including firearms examiners. Sometimes they are coming back from the edge of the limb to regions of stronger support as their own appreciation grows of what they can and cannot claim; sometimes pursuant to orders from courts.[5] Some forensic identification subfields are beginning to develop for their own areas the more scientifically defensible approach of empirically based, probabilistic conclusions that are now commonly used in DNA typing. Pursuant to the NAS Report, as well as a developing consensus inside and outside of forensic science, the trend in every field is toward rejecting the "bold absolutes" of Sibert's testimony, and no longer claiming that a match can be made "to the exclusion of all others in the world."

Further Explication of the Limits to Identification

[3] National Research Council of the National Academies of Science, Ballistic Imaging (2008), p. 82.
[4] Harry Edwards, Solving the Problems that Plague the Forensic Science Community, 50 Jurimetrics 5 (2009).
[5] E.g., U.S. v. Glynn, 578 F. Supp. 2d 567 (S.D. N.Y. 2008).

10. In the paragraphs below I explain *why* the claimed ability to identify the unique source of questioned crime scene evidence is regarded by conventional scientists and the NAS forensic science committees to be an unsupportable exaggeration, and I review some of the difficulties particular to firearms identification.

Probability

11. The most fundamental problem with Sibert's testimony was his insistence that he could identify which "ejector... to the exclusion of all other ejectors in the world" left its mark on a cartridge casing. When pressed to explain how it is that their field is able to perform this remarkable feat, practitioners of forensic identification, including firearms examiners, offer two basic rationales: probability theory and experience.[6]

12. The underlying theory of forensic identification, as expressed by the founders of the toolmark and firearms identification fields, was recited in brief form by Sibert in his testimony: "[T]hese little microscopic imperfections are random. They are unique to each individual weapon and will be reproduced by no other weapon...." This invokes probability theory (i.e., random variations).

13. *To say that probability can support a conclusion of absolute identification is irrational.* By their nature and by definition, probabilistic processes *cannot* lead to absolute conclusions. They lead to statements of probability.

14. How would one obtain these probability estimates? DNA typing illustrates the general approach. One must find out the frequency of various attributes in the relevant population (here, bullets and cartridges fired from different categories of weapons). Using those figures, and with an appropriate statistical model, one would then be able to calculate the probability of a random match (i.e., the probability that a weapon taken at random from the relevant population would be indistinguishable from the weapon involved in the crime, leading to the probability that two or more weapons produce indistinguishable marks) and report that to the factfinder.[7] Firearms examiners do not employ any such calculations – indeed, they have never even collected the data that would be necessary to make such calculations – and so they are unable to inform the factfinder of the relevant probability.

15. Instead, they refer vaguely to the random process that would underlie such a

[6] For more detailed discussion, see Michael J. Saks & Jonathan J. Koehler, The Individualization Fallacy in Forensic Science, 61 Vanderbilt L. Rev. 199 (2008).

[7] A random match is the probability that a weapon chosen at random from the relevant population (all weapons of a particular make, model, caliber) would produce marks which match those of the questioned bullet. The concept is routinely used in DNA typing.

calculation, and then make an illogical leap to assert that the probability of a random match is zero: that no other weapon could make the marks found. The assertion is incoherent, and based on nothing but *ipse dixit.*

Experience

16. A second argument offered to support the claim that no two weapons can produce the same (or confusingly similar) marks is the argument that, in examining many pairs of objects in their casework, examiners have not yet come across two sets of markings produced by different sources that are indistinguishable from each other.

17. One problem is that examiners never go hunting for those potential cases of false inclusion. They typically are presented with a questioned bullet which they are asked to compare to one known weapon. They have no occasion to see whether others match as well or better or not. In an actual case, no one knows which weapon produced the marks on a questioned bullet, so if and when false matches are made, examiners have no way to tell that they have made such errors. They declare an identification and that is the end of the inquiry.[8]

18. But even if examiners were able to make, and were making, systematic efforts to discover how many false matches are possible in a population of weapons, examiners are likely to greatly underestimate the incidence. Here is how we explain this problem in *The Individualization Fallacy*:

> "To illustrate, suppose that exactly 100 pairs of firearms out of an estimated 100,000 guns in a Texas town share indistinguishable gun barrel markings. If each of 100 firearms experts examined 10 pairs of guns from the town's gun population every day for 10 years (n=3,650,000 gun pairs), there is about a 93% chance that none of the indistinguishable pairs will have come under examination. That is, despite 1,000 'collective years' of forensic science experience (100 experts multiplied by 10 years), the failure to find even a single pair of guns with indistinguishable markings would offer little basis for drawing conclusions about whether gun barrel markings, even in this single town, are unique."[9]

19. In short, "experience" is an inadequate guide to what examiners need to know to evaluate the truth of their claims about the uniqueness of such marks, and provides no

[8] Sometimes such errors comes to light. But a study of nearly two dozen cases of false matches of fingerprints makes clear that such error are discovered and become public only under unusual circumstances. See Simon A. Cole, More than Zero: Accounting for Error in Latent Fingerprint Identification, 95 J. Criminal Law & Criminology 985 (2005).
[9] See Michael J. Saks & Jonathan J. Koehler, The Individualization Fallacy in Forensic Science, 61 Vanderbilt L. Rev. 199, 213 (2008).

objective basis for declaring a match.

Subjectivity

20. To this point we have not addressed the question of how examiners come to the conclusion of whether two bullets or shell casings share indistinguishably similar marks. I have discussed only the question of what one can make of such an asserted conclusion, and I have argued that any assertion of identification to the exclusion of all others in the world is necessarily an exaggeration. But how do examiners conclude whether markings were created by the same, or by different, weapons? The process remains a subjective one.

21. Perhaps I can illustrate what that means by describing my first encounter with a firearms examiner. My first visit to a crime laboratory occurred when I was living in the Boston area and was invited by a Massachusetts trial judge, Robert Hallisey, to join him on a tour of the state police crime laboratory. He had become curious about the work of the crime laboratory personnel who testified in his courtroom, arranged a personal tour for himself, and wanted some academic company. Before entering Harvard Law School, Judge Hallisey had earned a degree in physics from Harvard College, so he had both knowledge of and curiosity about science. My most vivid memory of that day remains our visit to the firearms unit. The examiner who was explaining his trade to us began by saying that firearms identification was an "exact science." He explained: That when he would first look at a new pair of questioned and known bullets through his comparison microscope he might not have a clue as to whether or not they were fired by the same weapon. But that after repeatedly pacing back and forth, and staring again through the microscope, and pondering, he would eventually make up his mind. And once he had done so he was absolutely sure that he had reached the correct conclusion. That was what he meant by an "exact science": his subjective confidence that he had intuited the correct conclusion.[10] From the testimony in the Macumber case, it appears that Sibert's examination and conclusions followed the same, or a very similar, approach.

22. Efforts have begun among firearms examiners to move their discipline onto a more solid foundation, but progress is slow. As the NAS Report explains, in 1992 the Association of Firearms and Toolkmark Examiners adopted

> "a theory of identification, but it does not provide a specific protocol. It says that an examiner may offer an opinion that a specific tool or firearm

[10] I was reminded of that experience when I read the opinions of two different federal district judges in Boston describing very similar-sounding testimony by two firearms examiners explaining their methods in two different cases. U.S. v. Monteiro, 407 F. Supp. 2d 351 (D. Mass. 2006); U.S. v. Green, 405 F. Supp. 2d 104 (D. Mass. 2005) (describing in some detail the difficulties of crediting the unsupported claims of firearms examiners). Apparently, not much has changed in the practice, although much has changed in the views mainstream scientists have of the methods and claims of forensic identification scientists, and in the views of some judges.

was the source of a specific set of toolmarks or a bullet striation pattern when 'sufficient agreement' exists in the pattern of two sets of marks. It defines agreement as significant 'when it exceeds the best agreement demonstrated between tool marks known to have been produced by different tools and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool.' The meaning of 'exceeds the best agreement' and 'consistent with' are not specified, and the examiner is expected to draw on his or her own experience. This AFTE document, which is the best guidance available for the field of toolmark identification, does not even consider, let alone address, questions regarding variability, reliability, repeatability, or the number of correlations needed to achieve a given degree of confidence." (NAS Report, p. 155.)

Subclass Features

23. Firearms identification is made more difficult by the fact that weapons impart not only class characteristics (such that all or most of the weapons of a particular make and model will pass on identical markings to the bullets and cartridges they fire), but also subclass characteristics: features that make a bullet or shell casing seem to match a specific weapon being examined, when the features actually are shared by a subgroup of weapons of that type. Even today some examiners are unaware of this problem, or are generally aware of it, but not sufficiently knowledgeable, so that they do not know when they are looking at features that are subclass characteristics and so mistake them for so-called individualizing characteristics.[11] Decades ago, few firearms examiners knew about subclass characteristics and fewer still had the knowledge to take them carefully into account when comparing test and evidence cartridges. As the NAS Report notes, "In earlier years, toolmark examiners relied on their past casework to provide a foundation for distinguishing between individual, class, and subclass characteristics. More recently, extensive training programs using known samples have expanded the knowledge base of examiners." (NAS Report, p. 153.) Sibert's testimony suggests he was unaware of subclass characteristics, since he testifies to the existence of only two kinds of characteristics: class and individual. [12/20/76, p. 4:6-13] If so, Sibert could have easily mistaken a subclass characteristic for an individual characteristic. That is, he could see a mark resulting from a subclass characteristic and erroneously conclude that it individualized the weapon.

Context Effects

[11] According to the forensic firearms examiner who provided the bullet images that appear in Michael J. Saks & Jonathan J. Koehler, The Coming Paradigm Shift in Forensic Identification Science, 309 Science 892 (2005), many examiners would mistake these two bullets as having come from the same gun (although they are known to have been test fired from two different guns), because those erring examiners are insufficiently schooled to recognize subclass characteristics.

24. Firearms examination, like other forensic sciences, must face the problem of human cognition whereby in observing ambiguous stimuli people tend to see what they are expecting or hoping to see. Thus, if a forensic examiner is told that investigators are confident they have identified the perpetrator, and are seeking confirmation by having the suspect's weapon tested, the examiner's conclusions usually will be influenced to some degree by the expectations that a match is going to be found. Most or many fields of science solve this type of problem with blind testing regimes. Most forensic sciences, including firearms examinations, have yet to solve the problem.[12]

25. Relatedly, "confirmation bias" is a phenomenon whereby a person who already believes he knows the correct conclusion, and is searching for evidence to test the validity of that conclusion, tends to look for evidence supporting the conclusion while simultaneously disregarding or giving reduced weight to disconfirming evidence.

26. Perhaps the most dramatic research illustration of the impact of context information is provided by Dror et al.[13] In their study, fingerprint examiners were presented with a pair of prints (both latent and known) that each examiner had examined several years earlier and concluded was a positive identification. On the current occasion, however, they were told that the pair of prints was that of the FBI's most famous erroneous match – of American Brandon Mayfield to a latent print left by a terrorist bomber near Madrid, Spain. Thus, examiners were given the expectation that the prints they were about to compare did not match. Despite the fact that several years earlier all of the examiners had already concluded that the pair they were examining was a positive match, on this occasion only one of five examiners reached the same conclusion that he had previously. An internal investigation by the FBI of its errors in wrongly associating the fingerprints of Brandon Mayfield to the latent print left by the terrorist bomber in Spain concluded that among other failings, the examiners has succumbed to context effects and confirmation bias.[14]

27. For purposes of the present case, to the extent that Sibert was made aware of the wishes or expectations of investigators or prosecuting attorneys when he performed his examinations, his judgments about what he observed would tend to be influenced by such

[12] For an extensive review of such context effects, their cure, and their neglect by forensic science, see D.M. Risinger, M.J. Saks, R. Rosenthal & W. Thompson, The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 U. Cal. L. Rev. 1 (2002). Blind testing is discussed at a number of points in the NRC Report. For indication of some progress in forensic science, at least in DNA typing, see D.E. Krane, S. Ford, J.R. Gilder, K. Inman, A. Jamieson, R. Koppl, I.L. Kornfield, D.M. Risinger, N. Rudin, M.S. Taylor, and W.C Thompson, Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation, 53 J. Forensic Sciences 1006 (2008).

[13] Itiel E. Dror et al., Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications, 156 Forensic Sci. Int'l 74 (2006).

[14] R.B. Stacey, Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case, 54 J. Forensic Identification 706 (2004).

information. Influence frequently is conveyed through cover letters requesting examinations of evidence, such as the August 29, 1974 letter from the MCSD to the FBI crime laboratory, stating that "[w]e have currently developed a new suspect and have new evidence for comparison" and urging that the FBI "rush [the] examination." An additional example of such influence in Sibert's examination of the firearms evidence in this case is that the second trial required him to revisit the firearms evidence to which he had testified in the first trial. Asking a person to take a fresh look at something he already had reached a conclusion about is perhaps the least "independent" an examination can get, from the perspective of its being free of influences extraneous to the current (second) examination. (The study described in paragraph 26 provides a contrast – imagine those examiners had been told what their prior decisions had been and were asked to repeat their examinations.)

28. Additionally, an example of "confirmation bias" is provided by Sibert's preparation of photomicrographs only of the apparently matching marks and none of the non-matching marks – suggesting a selective search for marks that would support a conclusions of inclusion and a disregard or lowered weighting of marks that would suggest exclusion.

29. Perhaps the most noteworthy indication that Sibert's examination suffered from these cognitive biases is his treatment of the finding that the firing pin and breach face marks on the four recovered cartridges did not match Macumber's weapon. Rather than giving appropriate weight to this evidence (which tended to exclude the suspect's weapon), Sibert proposed a number of hypotheses about why the weapon might, nevertheless, have been the murder weapon. Those hypotheses allowed him to explain away the exculpatory evidence. This is not to say that Sibert's theories were necessarily incorrect; no one can have any idea whether he was right or wrong. What is noteworthy is that without doing anything to test his new hypotheses against hard evidence, he gave those far greater weight than he gave the fact that the cartridges did not match the weapon. Had Sibert given all the evidence equal weight, it seems highly likely that rather than thinking up reasons why a non-suspect might actually be the perpetrator, Sibert would have treated exactly the same evidence as excluding the weapon. This illustrates how assuming a conclusion to be the correct one can subtly influence the process by which an expert's conclusion is reached.

Conclusion

30. The central emphasis of Sibert's testimony – that he was able to identify the specific weapon that fired a specific bullet "to the exclusion of all other firearms in the world" – is an empirically unsupportable exaggeration. It is, in essence, a statement of 100% certainty, which numerous recent scientific inquiries have revealed to be baseless. As the National Academy of Sciences report concluded in 2009: "The validity of the

fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.... A significant amount of research would be needed to scientifically determine the degree to which firearms-related toolmarks are unique or even to quantitatively characterize the probability of uniqueness." (NAS Report, p. 154.)

31. The firearms identification process is provided with little actual scientific guidance, and as a result the process of examination and of drawing conclusions that is employed by most examiners, including Sibert, is overwhelmingly subjective. Context effects and the risk of mistaking subclass features for individual characteristics can further infect the decision process, and increase the risk of examiner error. Thus, claims so extreme as those made by Sibert (and which continue to be made by some or many examiners today) cannot be supported by the limited evidence on which they rest. They are little more than *ipse dixit*.

32. In short, Sibert's testimony substantially exaggerated both his field's abilities and his own abilities, while seeming to be unaware of factors that contribute to erroneous conclusions. Most importantly, there was not then, and there is not now, a scientifically supportable basis for concluding that a particular mark was made by a particular weapon "to the exclusion of all others in the world," as Sibert did repeatedly.

SWORN this 4 day of May, 2011.

Michael J. Saks

SUBSCRIBED AND SWORN to before me this 4th day of May, 2011.

Tammy Vavra
Notary Public – Arizona

My Commission Expires: 3-9-2015

OFFICIAL SEAL
TAMMY VAVRA
Notary Public - State of Arizona
MARICOPA COUNTY
My Comm. Expires Mar. 9, 2015