



THE NATIONAL REGISTRY
OF EXONERATIONS

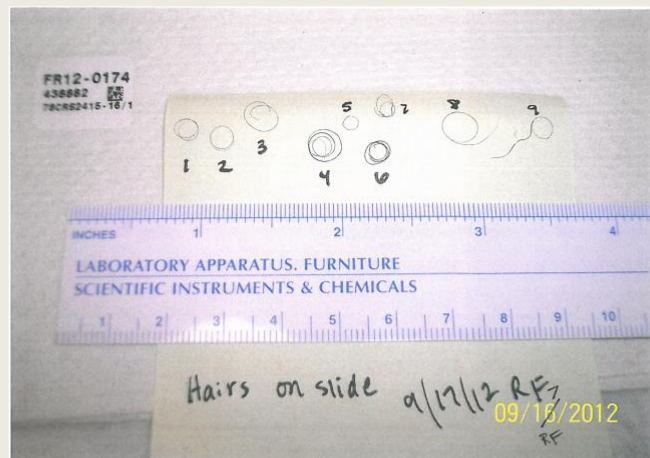
MICROSCOPIC HAIR COMPARISON ANALYSIS AND CONVICTING THE INNOCENT

Simon A. Cole
Jessica Weinstock Paredes
Maurice Possley
Ken Otterbourg

Version 1.1

Persistent Link:

<https://n2t.net/ark:/88112/x2d909>



THE NATIONAL REGISTRY OF EXONERATIONS | DECEMBER, 2023

A PROJECT OF THE UNIVERSITY OF CALIFORNIA IRVINE NEWKIRK
CENTER FOR SCIENCE & SOCIETY, UNIVERSITY OF MICHIGAN LAW
SCHOOL, AND MICHIGAN STATE UNIVERSITY COLLEGE OF LAW



MICROSCOPIC HAIR COMPARISON ANALYSIS AND CONVICTING THE INNOCENT

CONTENTS

I.	EXECUTIVE SUMMARY.....	7
A.	Background.....	7
B.	Objectives.....	7
C.	Major Findings.....	7
1.	Total Contribution of Hair Evidence to False Convictions	7
2.	Years Lost and Compensation	7
3.	Race.....	8
4.	Oral Testimony Compared to Written Reports.....	8
5.	State and Local Testimony Compared to FBI.....	8
6.	... Supposedly “Appropriate” Testimony Was Just as Damaging as “Erroneous” Testimony.....	9
II.	PREFACE.....	10
III.	INTRODUCTION	11
A.	Outside the FBI.....	13
B.	Objectives of the Report.....	14
C.	“Appropriate” v. “Erroneous” and True v. False.....	14
D.	Overview of the Report.....	16
IV.	MICROSCOPIC HAIR COMPARISON ANALYSIS.....	18
V.	THE FBI REVIEW	22
VI.	WHAT WE KNOW ABOUT THE FBI REVIEW.....	23
VII.	DATA AND METHODS.....	28
A.	The NRE Data.....	31

B. Methodological Differences between the Studies 33

VIII. RESULTS 34

A. Erroneous and Appropriate Testimony 34

B. Error Types 36

IX. DISCUSSION 38

A. Case Characteristics 38

 1. Number of People Falsely Convicted on MHCA Evidence of Any Kind 38

 2. Years Lost 39

 3. Number of People Falsely Convicted on MHCA Inclusions 39

 4. Inconsistent but not Excluded 39

 5. Racial Assignments 41

 6. DNA 44

 7. Crime Types 44

 8. The Death Penalty 45

 9. Gender 45

 10. Racial distribution 45

 11. Guilty Pleas 46

 12. Temporal Trends 46

 13. Geography 50

 14. Other Contributing Factors 52

 15. Role of the FBI 55

 16. Exonerations Before 1989 55

 17. Exclusions 55

B. Testimony 56

 1. Erroneous Testimony 56

 a) Error Type 1 56

 b) Error type 2 57

 c) Error Type 3 59

 d) Commonly Encountered Phrases 60

 2. Appropriate Testimony 63

 a) Appropriate Testimony in State and Local Cases 63

 b) Appropriate Testimony in FBI cases 65

C. Defense Experts 70

D. Cross and Re-direct Examination 71

E. Prosecutors and Closing Statements 75

F. Misconduct 76

G. Compensation 77

 1. Statutory Compensation 77

 2. Civil Compensation 78

X. CONCLUSION 79

XI. APPENDIX 1: METHODOLOGICAL APPENDIX ON USING THE NATIONAL REGISTRY OF EXONERATIONS DATA	81
A. Case Selection.....	81
B. Sources	81
XII. APPENDIX 2: METHODOLOGICAL APPENDIX ON THE FBI REVIEW CODING SYSTEM.....	83
A. Erroneous Evidence	83
B. Error Types.....	84
XIII. APPENDIX 3: DATA TABLES	89
XIV. APPENDIX 4: ALL POST-1989 EXONERATION CASES IN WHICH MHCA IMPLICATED DEFENDANT	92
XV. APPENDIX 5. ALL PRE-1989 EXONERATION CASES IN WHICH MHCA IMPLICATED DEFENDANT	100

Tables

Table 1. Summary statistics of FBI Hair Review from two snapshots. Sources: ABS Group (2018), FBI (2015a), (2015b), Innocence Project (2013), Reimer (2013)..... 25

Table 2. Texas Forensic Science Commission audit..... 30

Table 3. Descriptive comparison of all data sets on U.S. MHCA. 30

Table 4. Distribution of testimony types in exoneration cases in which microscopic hair comparison was used (n=184). 31

Table 5. Type of report in NRE inclusion cases. Most of the cases labeled “don’t know” were trials and probably did have testimony. They are labeled “don’t know” because we don’t have evidence that a forensic scientist actually testified, rather than, e.g., a written report being introduced as an exhibit. 32

Table 6. Expert’s employer in non-FBI MHCA inclusion cases. 32

Table 7. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in non-FBI exoneration cases in which inclusionary MHCA was used (n=99). 34

Table 8. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in FBI exoneration cases in which inclusionary MHCA was used (n=19). 35

Table 9. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in all exoneration cases in which inclusionary MHCA was used. These include 117 inclusions and one “race only inclusion” (n=118). 35

Table 10. Comparison of appropriate and erroneous evidence in FBI review and data sources used in this study. 36

Table 11. All post-1989 NRE cases in which MHCA contributed to a false conviction by type of testimony given (n=125). 38

Table 12. MHCA exoneration by crime type. Data as of July 12, 2023. 45

Table 13. Percent of Black Exonerees in MHCA and Non-MHCA Exonerations, by Crime, analysis as of July 12, 2023. 46

Table 14. Contributing factors in MHCA and non-MHCA Exoneration. 52

Table 15. Commonly encountered phrases in inclusion cases. Cases were coded for up to 4 phrases. All phrases derive from the FBI review, the Root Cause Analysis Report, or the TFSC, except those indicated by an asterisk. 61

Table 16. MHCA exonerees’ compensation (n=123). Federal exonerees and exonerees posted in 2023 are excluded. 78

Table 17. Comparison of compensation for MHCA, comparable non-MHCA, and comparable non-MHCA DNA exoneration..... 78

Table 18. Data table for Figure 5. 89

Table 19. Data table for Figure 6. 91

Figures

Figure 1. Conceptual overview of the approach of the FBI review. It began from a set of MHCA cases and proceeded to find cases in which the testimony was erroneous. In some of those cases, we know the defendant was exonerated. 17

Figure 2. Conceptual overview of the approach of the present study. It began from a set of exoneration cases and proceeded to find cases in which MHCA contributed to the underlying false convictions and cases in which testimony was erroneous. Colors are matched to Fig. 1. Numbers in the left 2 boxes refer to Table 9. 17

Figure 3. Review sheet included with notification letter sent to exoneree Richard Beranek showing selection of error types in original evidence in his case. 24

Figure 4. Distribution of error types in cases with erroneous testimony across all data sets (n=82). Error type 1 is individualization, Error Type 2 is an unfounded statistical probability, and Error Type 3 uses experience to vouch for accuracy. 37

Figure 5. MHCA exoneration by year of conviction. Data is shown only through 2009, the latest year of conviction for a MHCA exoneration at the time of publication. The lag between conviction and exoneration in all exoneration cases is approximately 9 years. The full data table for this chart is available in Table 18, Appendix 3. 47

Figure 6. MHCA exoneration by year of exoneration as of July 12, 2023. The data table for this chart is available in Table 19, Appendix 3. 48

Figure 7. Geographic distribution of MHCA exoneration. One federal case is counted for the state (California) in which it occurred. 50

Figure 8. Other evidence placing defendant at the scene or victim at location associated with defendant (n=125)..... 53

Guide to Abbreviations

The following abbreviations are used in this Report:

DOJ: United States Department of Justice

FBI: Federal Bureau of Investigation

IP: The Innocence Project

MHCA: Microscopic Hair Comparison Analysis (sometimes further abbreviated to MHC)

mtDNA: Mitochondrial DNA testing

NACDL: National Association of Criminal Defense Lawyers

NRE: National Registry of Exonerations

POI: Person of interest

TFSC: Texas Forensic Science Commission

WL: Westlaw

Acknowledgements

Vanessa Antoun, of the National Association of Criminal Defense Lawyers, provided information and insight into the FBI review and MHCA. Lynn Garcia, of the Texas Forensic Science Commission, David Kaye, and Chris Mumma, of the North Carolina Center on Actual Innocence, shared documents and data. For comments on drafts of this report, we are grateful to Sarah Chu, Glinda Cooper, Chris Fabricant, and Peter Neufeld of the Innocence Project, and Summer Farrar, Brandon Garrett, Max Houck, Janis Puracal, James Robertson, and Norah Rudin. This report could not have been completed without our colleagues at the National Registry of Exonerations, Samuel Gross, Barbara O'Brien, Catherine Grosso, Meghan Cousino, Jeff Gutman, and Dustin Cabral or without the Registry's supporters. An earlier version of this report was presented at a public conference at the University of Lausanne School of Criminal Justice, May 27, 2021.

On the Cover

Hair evidence from the [Joseph Sledge](https://innocencecommission-nc.gov/wp-content/uploads/state-v-joseph-sledge/commission-hearing-powerpoint-presentation.ppt) case. *Source:* North Carolina Innocence Commission, <https://innocencecommission-nc.gov/wp-content/uploads/state-v-joseph-sledge/commission-hearing-powerpoint-presentation.ppt>.

Suggested Citation

Simon A. Cole, Jessica Weinstock Paredes, Maurice Possley, & Ken Otterbourg, *Microscopic Hair Comparison Analysis and Convicting the Innocent*, National Registry of Exonerations, 2023, <https://n2t.net/ark:/88112/x2d909>.

I. EXECUTIVE SUMMARY

A. BACKGROUND

- In 2015, the Federal Bureau of Investigation (FBI) announced that a review of FBI microscopic hair comparison analysis (MHCA) testimony in 268 cases showed that FBI expert witnesses gave erroneous testimony in 96% of cases in which hair evidence was used to inculcate the defendant. A 2018 update found erroneous testimony in 93% of 484 cases.
- The FBI review covered only testimony by FBI experts, not testimony by state or local MHCA experts, but the FBI noted that its experts had trained hundreds, if not thousands, of state and local MHCA experts over a period of 25-30 years, prompting concerns that the problem at the FBI had been amplified across the entire country.

B. OBJECTIVES

- **Objective 1.** This report used data from the National Registry Exonerations (NRE) to complete the first study that compares the findings of the FBI review to MHCA evidence given by state and local experts.
- **Objective 2.** This report is also the first analysis of MHCA evidence in all known exoneration cases in the United States (as opposed to only DNA exoneration cases).

C. MAJOR FINDINGS

1. TOTAL CONTRIBUTION OF HAIR EVIDENCE TO FALSE CONVICTIONS

- At least 129 people have been falsely convicted based at least in part on MHCA. This is a conservative figure. This report analyzes [125](#) of these cases. (Four cases are not analyzed: [three](#) in which the exoneration occurred before 1989, and one, [Carlton Lewis](#), in which the exoneration occurred as this report was nearing completion.)
- Fifteen of these MHCA exonerees had been [sentenced to death](#), more than double the proportion of all exonerees who were sentenced to death.
- FBI analysts testified in 19 of these 125 exoneration cases. State or local analysts testified in the remainder.
- When asked to make hair comparisons in cases involving defendants who would later be exonerated, MHCA experts reported that unknown hairs from a crime scene and known hairs from a defendant could be from the same source in [7 out of every 10 cases](#) and reported that they were from different sources in only 3 out of 10 cases.

2. YEARS LOST AND COMPENSATION

- Collectively, the 125 MHCA exonerees whose cases are analyzed in this report lost a total of [1,918 years](#) in prison. These MHCA exonerees served an average of 15 years and 4 months.

- Taxpayers paid for the MHCA false convictions. Around two-thirds of the exonerees were compensated for a total of around \$350 million. But 40 of the exonerees—[nearly a third](#) of the total— received no compensation, including 36 who served time in prison.

3. RACE

- The defendant was Black in [more than half](#) of the MHCA exonerations, reflecting the racial skewing among exonerations generally.
- In at least 7 cases, the attribution of a hair to the defendant was accompanied by a supposed racial classification of the hair. In [all 7 cases](#), the defendant was Black.
- In one case, FBI testimony about only the supposed racial classification of the hair contributed to the conviction.

4. ORAL TESTIMONY COMPARED TO WRITTEN REPORTS

- The FBI’s finding of 93% erroneous *testimony* [was not replicated in the FBI witnesses’ written reports](#).
- Erroneous statements by FBI MHCA experts were 86% more common in trial testimony than in written reports.
- In other words, when they testified in court, FBI MHCA experts routinely used erroneous statements to make the evidence sound stronger than it had sounded in their written reports.
- The most plausible explanation for this is that the pressure to convict at trial induced MHCA experts to make erroneous statements at trial that they had not made in writing.
- What does this appalling practice tell us about the state of forensic science. And what does it tell us about our courts’ regulation of expert testimony?

5. STATE AND LOCAL TESTIMONY COMPARED TO FBI

- The FBI’s finding of 93% erroneous testimony [was not replicated among state and local examiners](#).
- State and local MHCA experts testified erroneously in 39% of exoneration cases.
- In many cases, [state and local experts](#), just like FBI experts, made outrageous erroneous statements. They implied that defendants *were* the source of hairs, when they should only have said they *might be* sources of hairs, cited baseless statistics, and misleadingly implied that their experience examining hair was a measure of the accuracy of their conclusions.
- But more often state and local MHCA experts testified “appropriately”—[in 61% of exoneration cases](#).
- This suggests that [whatever culture problem existed at the FBI did not fully carry over](#) to state and local MHCA experts.
- State and local examiners testified more appropriately than the FBI examiners who trained them, but state and local examiners still contributed to [more than their share](#) of false convictions because . . .

6. . . . SUPPOSEDLY “APPROPRIATE” TESTIMONY WAS JUST AS DAMAGING AS “ERRONEOUS” TESTIMONY

- Counterintuitively, “appropriate” testimony (using the FBI’s terms) [contributed to more false convictions than erroneous testimony did](#).
- [Even for FBI experts](#), “appropriate” testimony was more common in exoneration cases than it was in the FBI review.
- The contribution of MHCA evidence to exonerations, then, appears to result as much from the mere use of hair evidence against the defendant as from the erroneous nature of the testimony.
- Seemingly [innocuous statements like “similar” and “consistent with”](#) contributed to [false convictions](#) at least as often as did egregious misstatements.
- Contrary to the implication of the FBI review, the dangerousness of MHCA evidence was not confined to erroneous statements. As appalling as erroneous MHCA evidence was, *any* MHCA evidence, whether erroneous or “appropriate,” appears to have done a good job at convicting defendants who would later be exonerated.
- The kind of “appropriate” and seemingly innocuous, “consistent with. . .” testimony that has proven so damaging when used for MHCA evidence [is widely used in forensic and expert evidence in many other disciplines](#) as well. The example of MHCA warns us to focus on this kind of testimony as a driver of wrongful convictions.

II. PREFACE

This is a Report by the National Registry of Exonerations (NRE), an online archive of all known exoneration cases in the United States. The primary Registry is composed of exonerations that occurred after 1989. In addition to recording and publishing a summary of each exoneration, the Registry undertakes social science coding of each case. The Registry codes “contributing factors” for each exoneration. One of those contributing factors is “false or misleading forensic evidence.” Coding decisions are often extensively discussed by Registry staff and sometimes with external collaborators.

The origins of this Report lie in such discussions. Following discussions with our close collaborators at the Innocence Project (IP), we discovered that we had been coding microscopic hair comparison analysis (MHCA) evidence inconsistently. Although we had no trouble coding well-known cases in which MHCA analysts had testified to made-up statistics,¹ we had not consistently coded cases in which the evidence stated that a hair of questioned origin was “consistent with” a person of interest (POI) or that a POI “could be” the source of a hair. Consider, for example, the wrongful conviction of [Steven Avery](#), a well-known false conviction case because of the television documentary *Making a Murderer*. During the 1985 trial of Avery for rape,

A state forensic serologist, Sherry Culhane, testified that a hair recovered from a shirt of Avery’s was “similar” and “consistent” with Beerntsen’s hair. She conceded that the hairs of many people are consistent with one another, that she could not give a probability that the hairs were from the same source, and that all she could say was “that it’s not impossible” the hairs were from the same source.²

Was this “false or misleading forensic evidence”? On the one hand, these statements were not false. On the other hand, they seem, if not deliberately misleading, at least potentially highly misleading. On the face of it, the evidence seems almost meaningless. But did the jury understand it that way? Or was the evidence destined to be outweighed by the jury and thus contribute to the false conviction of Steven Avery?

As we discussed this issue, it became apparent that it was not limited to MHCA. This sort of “consistent with . . .” evidence appears for many types of expert evidence: medical testimony, pathology, sexual assault examinations, drug analysis, fingerprint, firearms, handwriting, bitemarks, shoeprints, and so on.³

What began as a simple coding discussion mushroomed into a deep dive into the nature of the MHCA evidence in exoneration cases. We already knew, of course, through the work of the IP and its collaborators, that MHCA was a major contributor to wrongful convictions.⁴ We were also well aware of the well-publicized finding by the Federal Bureau of Investigations (FBI) that its

¹ Such as [Jimmy Ray Bromgard](#).

² See Brandon L. Garrett, *Convicting the Innocent: DNA Exonerations Database* (2022), Duke University, available at <https://convictingtheinnocent.com> (<https://perma.cc/FN5F-SKGZ>).

³ E.g., R. Ross et al., *Consistent With: What Doctors Say and Jurors Hear*, 51 *Australian Journal of Forensic Sciences* 109 (2019) <https://doi.org/10.1080/00450618.2017.1324583>.

⁴ E.g., Barry Scheck et al., *Actual Innocence: When Justice Goes Wrong and How to Make It Right* (2003); Brandon L. Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong* (2011).

own examiners had given erroneous testimony in 96% of cases and the FBI and innocence community’s concern that the FBI had trained hundreds or thousands of local forensic analysts in MHCA.⁵ But the rate of erroneous testimony for exoneration cases did not appear to be anywhere close to 96%. Instead, “consistent with . . .” testimony, which the FBI called “appropriate,” appeared to be quite common. Was the FBI’s tendency to testify erroneously *not* replicated by local analysts? Was the testimony in exoneration cases “better” than in non-exoneration cases? Is “consistent with . . .” testimony okay? This Report is an effort to answer these questions.

III. INTRODUCTION

In 2015, the FBI made headlines by announcing the completion of a review of its own forensic analysts’ testimony about MHCA.⁶ These headlines included pronouncements like “FBI Admits Flaw in Hair Analysis over Decades.”⁷

The announcement was the product of a remarkable collaboration between the FBI, the U.S. Department of Justice (DOJ), the National Association of Criminal Defense Lawyers (NACDL), and the Innocence Project (IP).⁸ The four organizations had collaborated after the FBI conceded that its MHCA experts had “exceeded the limits of the science” in a number of cases. Some of these cases were high profile wrongful convictions.⁹

For the collaborative project, the FBI reviewed a sample of transcripts of its own experts’ testimony about MHCA. The researchers reviewed 268 transcripts. They found that FBI experts’ testimony contained “erroneous statements” in 257 (96%) of the 268 cases in which hair evidence was used to inculcate the defendant. It was this finding that led to “breathless”¹⁰ pronouncements in the media, such as “CSI Is a Lie” and “The FBI Faked an Entire Field of Forensic Science.”¹¹

The FBI’s own press release was unsparing. It quoted IP Co-Director Peter Neufeld as saying, “These findings confirm that FBI microscopic hair analysts committed widespread, systematic error, grossly exaggerating the significance of their data under oath with the consequence of unfairly bolstering the prosecutions’ case. . . . [T]his epic miscarriage of justice calls for a rigorous review to determine how this started almost four decades ago and why it took so long to come to light.”¹²

⁵ FBI, FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review, Type, (Apr. 20, 2015), <https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review> (<https://perma.cc/6RKH-KWJD>).

⁶ Id.

⁷ Spencer S. Hsu, *FBI Admits Flaws in Hair Analysis over Decades*, Washington Post, Apr. 18, 2015.

⁸ Norman L. Reimer, *The Hair Microscopy Review Project: An Historic Breakthrough for Law Enforcement and A Daunting Challenge for the Defense Bar*, The Champion 16, July, (2013), <https://www.nacdl.org/champion.aspx?id=29488>.

⁹ Spencer S. Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, Washington Post, Apr. 16, 2012.

¹⁰ Max M. Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, in Forensic Science Reform 25-55, 51 (Koen and Bowers eds., 2017).

¹¹ Conor Friedersdorf, *CSI Is a Lie*, The Atlantic, Apr. 20, (2015); Dahlia Lithwick, *Pseudoscience in the Witness Box: The FBI Faked an Entire Field of Forensic Science*, Slate, Apr. 22, (2015), <https://slate.com/news-and-politics/2015/04/fbis-flawed-forensics-expert-testimony-hair-analysis-bite-marks-fingerprints-arson.html>.

¹² FBI, FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review, Type, (Apr. 20, 2015), <https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review>.

The episode has been cited as evidence of a pathology in forensic science that is broader than the FBI Laboratory and MHCA. Many have interpreted the episode as “a sentinel event for other forensic sciences.”¹³ Some perceive it as indicative of a more general tendency toward overstatement in forensic reporting.¹⁴ Others have suggested that it is an example of the operation of cognitive bias in forensic science¹⁵ or, more generously, cognitive issues in expert judgment.¹⁶ More specifically, some have suggested that it illustrates the excessive orientation of forensic scientists with the prosecution and law enforcement.¹⁷ For Robertson, the problem was that historically FBI examiners were sworn FBI agents, and some lacked scientific training.¹⁸ For others, the episode demonstrated the willingness of forensic scientists to testify in court based on techniques that have not yet been scientifically validated,¹⁹ the judiciary’s failure to demand scientific validation as a condition of admitting the evidence, or both.²⁰ And, for some the episode illustrates the contribution of forensic evidence to wrongful convictions.²¹

Still others have seen the problem as indicative of a “cultural” problem in forensic science generally or in the FBI Laboratory (or the Hair and Fiber Unit) specifically.²² A root cause analysis

¹³ Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, 35.

¹⁴ Christopher J. Plourd & Ron Reinstein, *A Judge Comments*, 42 *Litigation* 33, 553 (2016); Paul Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*, 46 *Crim. L. Bull.* 531 (2010); David H. Kaye, *Ultracrepidarianism in Forensic Science: The Hair Evidence Debacle*, 72 *Washington and Lee Law Review* (2015); Paul Giannelli, *FBI Review of Microscopic Hair*, 29 *Crim. Just.* 31 (2014); Simon A. Cole & Troy Duster, *Microscopic Hair Comparison and the Sociology of Science*, 15 *Contexts* 28-35, Winter, (2016); Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, 52; Valena E. Beety, *Cops in Lab Coats and Forensics in the Courtroom*, 13 *Ohio St. J. Crim. L.* 543 (2016); Jessica Gabel Cino, *Bad Science Begets Bad Convictions: The Need for Postconviction Relief in the Wake of Discredited Forensics*, 7 *U. Denver Crim. L. Rev.* 1, 3 (2017); Jennifer L. Mnookin, *The Uncertain Future of Forensic Science*, 147 *Daedalus* 99, 103 (2018) https://www.mitpressjournals.org/doi/abs/10.1162/daed_a_00523; Brandon L. Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 91 (2021).

¹⁵ Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*; Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, 52; Michael Mansfield, *In the Eye of the Beholder*, 4 *Brit. J. Am. Legal Stud.* 691, 693 (2015); Dana Delger, *What Soda Challenges and Tennis Balls Can Teach us about Cognitive Bias*, *For. Sci. Int.* e1 (2018) <https://doi.org/10.1016/j.forciint.2018.11.009>.

¹⁶ James Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 49 *Australian Journal of Forensic Sciences* 239, 254 (2017) <https://doi.org/10.1080/00450618.2017.1279838>.

¹⁷ Lithwick, *Pseudoscience in the Witness Box: The FBI Faked an Entire Field of Forensic Science*, Friedersdorf, *CSI Is a Lie*.

¹⁸ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 252.

¹⁹ Linda Geddes, *Hanging by a Thread*, 226 *New Scientist* 27, May 2, (2015); Joëlle Vuille & Christophe Champod, *Forensic science and wrongful convictions*, *The Routledge International Handbook of Forensic Intelligence and Criminology* 125 (2017); Aliza B. Kaplan & Janis C. Puracal, *It’s Not a Match: Why the Law Can’t Let Go of Junk Science*, 81 *Alb. L. Rev.* 895, 923 (2017-2018).

²⁰ Samuel D. Hodge, Jr. & Amelia Holjencin, *A Post-Mortem Review of Forensic Hair Analysis - A Technique Whose Current Use in Criminal Investigations is Hanging on by a Hair*, 64 *St. Louis U. L.J.* 219 (2020) <https://scholarship.law.slu.edu/lj/vol64/iss2/5>; Brandon L. Garrett, *Bad Hair: The Legal Response to Mass Forensic Errors*, 42 *Litigation* 32 (2016); Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*; Kaye, *Ultracrepidarianism in Forensic Science: The Hair Evidence Debacle*; Cole & Duster, *Microscopic Hair Comparison and the Sociology of Science*; Jim Norton et al., *Flawed Forensics: Statistical Failings of Microscopic Hair Analysis*, 13 *Significance* 26-29, Apr. 8, (2016), <https://doi.org/10.1111/j.1740-9713.2016.00897.x>; Eric S. Lander, *Fix the Flaws in Forensic Science*, *New York Times*, Apr. 21, 2015.

²¹ Catherine L. Bonventre, *Wrongful Convictions and Forensic Science*, *WIREs Forensic Science* e1406 (2020).

²² Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 253.

commissioned by the FBI attributed the problem primarily to two cultural causes: insufficient leadership and overconfidence on the part of laboratory management.²³

A. OUTSIDE THE FBI

It was widely recognized in 2015 that the FBI review only constituted the tip of the proverbial iceberg of the problem with MHCA. The review covered only testimony by some 28 FBI MHCA analysts, although those analysts were MHCA specialists who may have worked on hundreds of cases annually. But the press release noted that “Over the course of 25 years, the FBI conducted multiple two-week training courses that reached several hundred state and local hair examiners throughout the country and that incorporated some of the same scientifically flawed language that the FBI’s examiners had used in some lab reports and often in trial testimony.”²⁴ Therefore, the FBI promised to “Strongly encourage the states again to conduct their own independent reviews where its examiners were trained by the FBI.”²⁵ And, it noted that “the Texas Forensic Science Commission has already begun a review of cases handled by analysts at state and local crime labs. Similar audits are needed in most other states.”²⁶ Others claimed the FBI understated the issue: that over the course of 30 years *thousands* of state and local examiners from 48 states had attended the “FBI Hair and Fiber School.”²⁷ Commentators noted that “The review focuses on the first few hundred cases, involving FBI examiners, but the same mistakes and faulty testimony were likely presented in any state prosecutions that relied on the between 500 and 1,000 local or state examiners trained by the FBI.”²⁸ Local and state examiners have been estimated to have handled 20-50% of MHCA nationwide. The files of disgraced forensic scientist Joyce Gilchrist, who included MHCA among the disciplines she practiced, show notes from her FBI course that suggest the FBI taught her to elide the fact that one cannot conclude that a specific individual is the source of hair by referring to the rarity of seeing consistent hairs from different people in her experience (what will be called an “Error Type 3” below).²⁹

As far as we know, no more than 17 states have undertaken the audits suggested by the FBI, and the audits are primarily focused—understandably—on litigation: on finding wrongly convicted persons and getting them legal assistance.³⁰ However, there has been far less follow up on the FBI’s suggestion from a research perspective. No results from any state audit have yet been

²³ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 13 (Aug., 2018), <https://vault.fbi.gov/root-cause-analysis-of-microscopic-hair-comparison-analysis/root-cause-analysis-of-microscopic-hair-comparison-analysis-part-01-of-01/view> (<https://perma.cc/C4Y9-GVKN>).

²⁴ FBI, FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review.

²⁵ *Id.*

²⁶ *Id.*

²⁷ Amelia Maxfield et al., Under the Microscope: Strategies for Implementing State & Local Microscopic Hair Comparison Reviews Address at National Association of Criminal Defense Lawyers Webinar (July 20, 2016).

²⁸ Lithwick, *Pseudoscience in the Witness Box: The FBI Faked an Entire Field of Forensic Science*, Spencer S. Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*, Washington Post, Dec. 22, 2011; James R. Robertson & Elizabeth Brooks, *Historical Context to Contemporary Practice in Hair Examinations*, in *A Practical Guide to the Forensic Examination of Hair: From Crime Scene to Court* 1-23 (Robertson and Brooks eds., 2021).

²⁹ Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*.

³⁰ Maxfield et al., Under the Microscope: Strategies for Implementing State & Local Microscopic Hair Comparison Reviews; Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 90. We believe some form of state review exists in at least the following states: Arizona, Arkansas, California, Colorado, Connecticut, Florida, Illinois, Iowa, Kansas, Massachusetts, Missouri, Nebraska, New York, North Carolina, Pennsylvania, Texas, and Virginia.

published in any form. The only institution conducting a state audit that made their data available to us was the Texas Forensic Science Commission (TFSC). Nor is a list of ongoing state audits even publicly available. Therefore, there is no public data available with which to test the proposition that the problem in the FBI MHCA unit was replicated across states and other localities. It would be reasonable to assume that the practice at the FBI was simply propagated across the country, that FBI testimony was “entirely typical.”³¹ But was it?

B. OBJECTIVES OF THE REPORT

The current study has two primary objectives: **Objective 1:** It is the first study that applies the FBI review’s methods to MHCA testimony outside the FBI. In this way, it begins to address the questions raised by both the FBI itself and by outside commentators about whether, and to what extent, the problems identified in the FBI laboratory were disseminated to state and local laboratories, either through the FBI’s massive training program, or through more indirect means. Thus, it is the first study to compare different data sets to the data contained in the FBI review.

In order to carry out that comparison, it was necessary to understand how the FBI review was carried out. That understanding was difficult because the FBI review was never published, and much about it remains obscure, in part because the FBI review was never really intended to be a research project. An ancillary aim of this paper is to try to articulate for the public what is publicly known about the FBI review’s methods and data.

Objective 2 is to use a data set of exonerations to learn what we can about the role of MHCA in false convictions. Exoneration was central to the FBI review. As the FBI itself stated “The FBI and DOJ agreed to conduct a review of criminal cases involving microscopic hair analysis *after the exoneration of three men convicted at least in part because of testimony by three different FBI hair examiners whose testimony was scientifically flawed.*”³² The fact that “74 of the 329 wrongful convictions overturned by DNA evidence involved faulty hair evidence” was cited in the FBI’s press release as a reason for the importance of the review.³³ And, in the same press release, NACDL Director Norman Reimer emphasized the Review’s potential to generate new exonerations. That said, however, the vast majority of the 257 defendants in the cases containing “erroneous” testimony were not exonerated. And, in only a small portion of the 74 DNA exonerations mentioned in the press release was the testimony from an FBI expert.

C. “APPROPRIATE” V. “ERRONEOUS” AND TRUE V. FALSE

In reading further, it is important to understand the meaning of the categories “appropriate” and “erroneous” as used in the FBI review. These terms apply to *testimony* and not to the forensic conclusion reached. The headline of the FBI press release, “FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review,” may have generated some confusion. By “errors,” the FBI meant that the testimony exceeded the limits of science, not that the conclusion about the source of the hair was false or that the convicted

³¹ Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 91.

³² FBI, *FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review.*, emphasis added

³³ *Id.*

defendant was innocent. As Kaye has noted, the FBI review did *not* show that FBI MHCA “produces wrong answers over 90% of the time.”³⁴ *Testimony* could be erroneous, the conclusion (whether two hairs come from the same source) could be true, and the defendant could be guilty. The 96% figure is an “erroneous testimony rate,” not a “false positive error rate,” in the sense that an “error” means that the hair did not originate from the person the analyst said it might have. Thus, a 96% erroneous testimony rate could coexist with any false positive error rate between 0% and 100%.

However, the converse is also true. Just because testimony was “appropriate” does not mean that the conclusion was correct. An expert could testify “appropriately” that a crime scene hair is “consistent with” the defendant’s sample hairs, and that conclusion could be false, and the defendant could be innocent. A rate of appropriate testimony is not an accuracy rate. Any rate of appropriate testimony could coexist with any false positive error rate.

Wrong conclusions in hair comparison can coexist with erroneous testimony. The Director of the Montana State Crime Laboratory, Arnold Melnikoff, testified erroneously in three exoneration cases. However, in addition to its criticism of Melnikoff’s *testimony*, the authors of a Peer Review Report suggested that Melnikoff’s *conclusions* were erroneous as well.³⁵

However, wrong conclusions can also coexist with “appropriate” testimony. In the prosecution of the Exonerated Five for sexual assault and [Steven Lopez](#) for robbery in the famous “Central Park jogger case,” MHCA expert Nicholas Petraco testified “appropriately” about several hair comparisons, including one in which he claimed that hairs found on [Kevin Richardson](#) “could have” originated from the victim. And the prosecutor, Elizabeth Lederer, said, in closing, that Petraco “found on Kevin Richardson’s underpants a hair that matched the head hair” of the survivor. A hair from his T-shirt “matched” the survivor’s pubic hair, she said, and the third hair, from his jeans, “was consistent and similar to” to the survivor’s hair. After DNA evidence cast doubt on the convictions,

Special Agent Douglas Deedrick, the Unit Chief of the FBI’s Information and Evidence section, said he disagreed with Petraco’s trial testimony regarding hairs found on Richardson; the hairs were not suitable for comparison and could not be used to link Richardson to Meili, Deedrick said.

Similarly, in the prosecution of [Thomas Murphy](#) for sexual assault in Wisconsin in 1999, the MHCA expert testified “appropriately” that Murphy “could have” been the source of a hair found on the victim’s bed. But, a year later the court granted Murphy’s motion for a new trial on the ground that the testimony about the hair was erroneous and that the comparison was inconclusive.

In addition, in the [Ronald Carden](#) case, “the FBI . . . discovered that the body hairs used to help convict Carden were not distinctive enough to compare with each other, and this information had been withheld by the prosecution.”

For more examples of possible errors in conclusion (as opposed to erroneous testimony), see the [Anthony Michael Green](#) case discussed [below](#), the [Gary Nelson](#) case discussed [below](#), and the [Omar Saunders](#) case discussed [below](#).

³⁴ Kaye, *Ultracrepidarianism in Forensic Science: The Hair Evidence Debacle*.

³⁵ *Peer Review Report: State v. Bromgard*, (2002),

http://m15080.kaivo.com/LegalDev/NLADA/Defender/forensics/for_lib/Documents/1039040279.64/BROMGARD%20REPORT.doc.

This report, like the FBI review, primarily concerns erroneous *testimony*, rather than erroneous *results*. Thus, it does not address another important question: in how many cases the result was simply wrong—the crime-scene hair was *not* “consistent with” the known hair. In some cases discussed in this report, we know that the POI was *not* the source of the crime scene hair because mitochondrial DNA (mtDNA) testing excluded them. We do not, however, without having an expert examine the evidence, know whether it was incorrect to conclude that the hairs were “consistent.” Even in cases in which mtDNA was not conducted, we can infer, because all of our cases are exoneration cases, that in most (but not all) of them, a MHCA analyst associated two hairs which were not from the same source. (There may be some cases in which the POI *was* the source of the hair, but the defendant was nonetheless innocent.) That false association may have been an error by the analyst or a result of the poor discriminating power of MHCA.³⁶ In short, this report is really about MHCA testimony, but is able to say little about an equally important problem: simply poor MHCA work.

D. OVERVIEW OF THE REPORT

Somewhat paradoxically, then, the FBI review was a study motivated by exonerations but based on a data set of MHCA cases that contained only a small portion of exonerations. This Report can, in some sense, be viewed as the converse: a study about MHCA based on a data set of exonerations that contains only a small portion of MHCA cases (compare Figure 1 with Figure 2). Garrett and Neufeld and the Innocence Project have analyzed and counted the role of MHCA in the narrower set of exonerations that were brought about by post-conviction DNA testing (“DNA exonerations”), finding egregiously inappropriate testimony in many cases, though not nearly at the rate found in the FBI review.³⁷ This study is broader, including all exonerations, not just DNA exonerations. However, this breadth is achieved at the cost of lacking complete transcripts in most (70%) cases (see Appendix 1: [Sources](#)).

This study is not as large as the FBI review, and its data is less detailed: most importantly, the FBI review had access to complete transcripts of nearly every case included in its data. However, as we shall see below, the present study has some advantages in terms of transparency and, paradoxically, in the level of detail it presents.

³⁶ Max M. Houck & Bruce Budowle, *Correlation of Microscopic and Mitochondrial DNA Hair Comparisons*, 47 J. Forensic Sci. 1 (2002).

³⁷ Brandon L. Garrett & Peter Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 Va. L. Rev. 1 (2009); Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*; Daniel Penchina & Katherine Huffman, *Not a Strand of Evidence: DNA Exonerations Involving Unvalidated or Improper Microscopic Hair Comparison Evidence*, The Innocence Project, (Apr., 2015), <https://s3.documentcloud.org/documents/6797315/InnocenceProject-Not-a-Strand-of-Evidence-Updated.pdf> (<https://perma.cc/6RHB-ZZKJ>). These researchers used the Innocence Project’s definition of “DNA exoneration”: a case in which post-conviction DNA testing was central to establishing innocence.

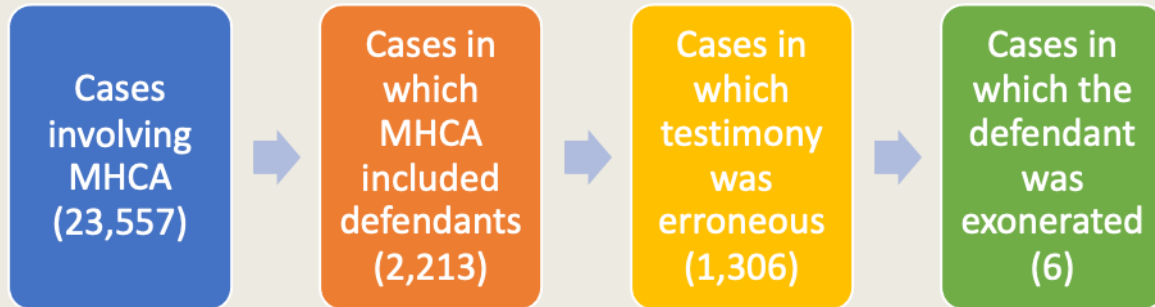


Figure 1. Conceptual overview of the approach of the FBI review. It began from a set of MHCA cases and proceeded to find cases in which the testimony was erroneous. In some of those cases, we know the defendant was exonerated.

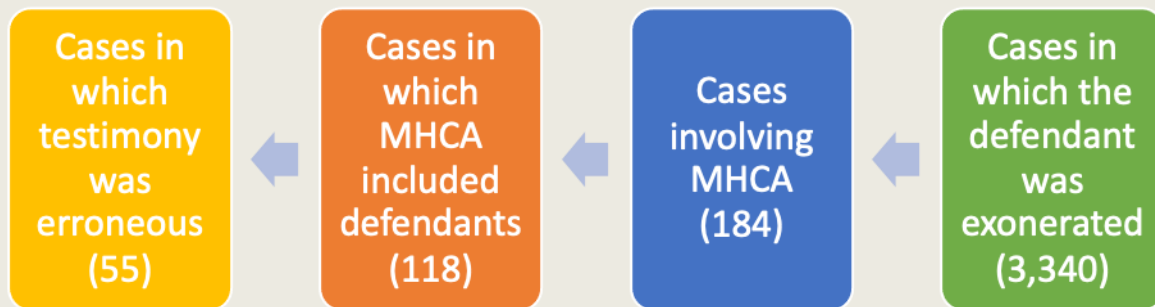


Figure 2. Conceptual overview of the approach of the present study. It began from a set of exoneration cases and proceeded to find cases in which MHCA contributed to the underlying false convictions and cases in which testimony was erroneous. Colors are matched to Fig. 1. Numbers in the left 2 boxes refer to Table 9.

This remainder of this Report proceeds as follows: In [Section IV](#), we provide background information on MHCA. In [Section V](#), we provide background on the FBI review. Information about the FBI review is scattered across two different data releases and other documents, some of which are public and some of which must be requested through the Freedom of Information Act (FOIA). Therefore, in [Section VI](#), we attempt to concisely summarize what is publicly known

about the FBI review. This section discusses a widely overlooked finding of the FBI review: the high rate of erroneous statements in FBI *testimony* was *not* found in FBI *written reports*. We discuss possible explanations for this finding.

[Section VII](#) explains the data and methods used in this Report, and [Section VIII](#) presents the results of our analyses.

In [Section IX](#), we discuss our findings: the total number of exonerees convicted at least in part by MHCA evidence, years lost, different kinds of testimonial statements, the role of DNA and the death penalty, gender, race, temporal patterns, geography, guilty pleas, other contributing factors, and pre-1989 exonerations. This section also contains our most surprising finding: counterintuitively, the kind of “consistent with” testimony that the FBI calls “appropriate” contributed to false convictions at a greater rate than erroneous testimony.

[Section X](#) offers brief concluding thoughts, including the implications of this Report for other expert and forensic disciplines.

Details on the methods used in this report are available in two appendices. [Appendix 1](#) describes our methods for finding MHCA cases in the Registry and for finding and coding the language used by MHCA experts. [Appendix 2](#) summarizes what we are able to learn, including from non-public information, about the coding processes in the FBI review. This information informed our coding used in this Report. This information may be of interest to readers who want to learn more about the FBI review.

[Appendix 3](#) consists of data tables, and [Appendix 4](#) and [Appendix 5](#) list all 128 known exoneration cases to which MHCA contributed.

IV. MICROSCOPIC HAIR COMPARISON ANALYSIS

Occasional uses of MHCA have been reported since the 19th century.³⁸ The earliest U.S. legal opinion on the use of MHCA appeared in 1882.³⁹ The earliest books on the topic appeared in 1884 and 1910.⁴⁰ The FBI Hair and Fibers Unit was established with a single examiner during the 1950s.⁴¹ However, MHCA did not come into widespread use until the 1970s.⁴² MHCA consists of using a microscope to compare a variety of features of hairs of questioned origin with hairs of known origin. The hairs of questioned origin come from a crime scene, which may be a home, a vehicle, clothing, or even a person’s body. A sample of around 50 hairs of known origin may be plucked or combed by consent from a specific body area of a POI or taken from a recently deceased person.⁴³ MHCA examiners then compare more than 20 characteristics between the

³⁸ Hodge & Holjencin, *A Post-Mortem Review of Forensic Hair Analysis - A Technique Whose Current Use in Criminal Investigations is Hanging on by a Hair*, 220; Clive A. Stafford Smith & Patrick D. Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil?*, 27 Colum. Human Rights L. Rev. 227, 234 (1996).

³⁹ Edward J. Imwinkelried, *Forensic Hair Analysis: The Case against the Underemployment of Scientific Evidence*, 39 Washington and Lee Law Review 41, 41 (1982).

⁴⁰ Id.; Robertson & Brooks, *Historical Context to Contemporary Practice in Hair Examinations*, 1.

⁴¹ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 58.

⁴² Hodge & Holjencin, *A Post-Mortem Review of Forensic Hair Analysis - A Technique Whose Current Use in Criminal Investigations is Hanging on by a Hair*, 225.

⁴³ Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, 45.

questioned and known samples to determine whether the questioned hairs are “consistent” with a known sample.⁴⁴

Like many trace evidence techniques, MHCA consists of two steps: (1) a finding of “consistency”; and (2) interpreting the meaning of that finding. For MHCA, both steps are problematic. There is little information about the accuracy of MHCA experts on finding consistencies, but historically, MHCA experts have performed poorly on proficiency tests on this task.⁴⁵

It should be noted that MHCA is not necessarily based on finding all the characteristics observed in the questioned hair in a single known hair. The MHCA literature characterizes a situation in which all characteristics of a questioned hair are found in a single sample hair as an “ideal.”⁴⁶ However, it is acknowledged that this ideal is not always met.⁴⁷ In those cases, MHCA experts can make findings of consistency by comparing the characteristics of the questioned hair to the “range of characteristics” found in the *sample* of known hairs, which, as noted above, might consist of around 50 hairs. In other words, when MHCA experts find “consistency,” it might not be between a single questioned hair and single known hair, but between a single questioned hair and up to 50 known hairs. In such situations, some authorities state that the questioned hair must “also be similar in all *major* characteristics to at least one hair within that known sample.”⁴⁸ Others simply state that “it may sometimes be possible to utilize several hairs from the known sample to locate characteristics which correspond to those of a questioned hair.”⁴⁹

This practice raises obvious “trawling” issues: one must ask how many people might yield samples of 50 hairs that might be found consistent with the characteristics of a single questioned hair. Such associations are obviously much easier to find than a finding of complete consistency between a known and questioned hair.

This raises the second step: interpreting what a finding of consistency means. It has always been well understood that hairs from the same individual are not all the same. Hairs from different parts of the body (head, pubic, and body hair, for example), of course, differ. But, moreover, even hairs from the same body part of the same individual differ. At the same time, it was understood that hairs from different people might share the same microscopically observed features.⁵⁰ But what was not known was *how many* people might share these characteristics.⁵¹ For this reason, it has long

⁴⁴ Norton et al., *Flawed Forensics: Statistical Failings of Microscopic Hair Analysis*, 27.

⁴⁵ Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*, 13; Williamson v. Reynolds, 904 F.Supp. 1529, 1556 (E.D. Okla. 1995); PCAST, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*, President’s Council of Advisors on Science and Technology, 120 (Sept., 2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (<http://perma.cc/EKZ3-FRA6>); Melissa Airlie et al., *Forensic Hair Analysis—Worldwide Survey Results*, 327 *For. Sci. Int.* 110966, 7 (2021) <https://doi.org/10.1016/j.forsciint.2021.110966>.

⁴⁶ Scientific Working Group on Materials Analysis, *Forensic Human Hair Examination Guidelines*, §14.3 (Apr., 2005), <https://www.asteetrace.org/static/images/pdf/01%20Forensic%20Human%20Hair%20Examination%20Guidelines.pdf> (<https://perma.cc/3KB4-4ZRG>); Harold Deadman, *Human Hair Comparisons Based on Microscopic Characteristics* Address at International Symposium on Forensic Hair Comparisons (June 25-27, 1985), 48.

⁴⁷ Cary T. Oien, *Forensic Hair Comparison: Background Information for Interpretation*, 11 *Forensic Science Communications*, Apr., (2009), https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/april2009/review/2009_04_review02.htm (<https://perma.cc/BHB3-MWSZ>).

⁴⁸ Barry D. Gaudette, *Evidential Value of Hair Examination*, in *Forensic Examination of Hair* 243-260, 246 (Robertson ed., 1999). (emphasis added)

⁴⁹ Deadman, *Human Hair Comparisons Based on Microscopic Characteristics*, 49.

⁵⁰ Oien, *Forensic Hair Comparison: Background Information for Interpretation*.

⁵¹ Wendy J. Koen, *Case Studies: Colin Campbell Ross and James Driskell*, in *Forensic Science Reform* 25-33, 29 (Koen and Bowers eds., 2017); Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 85-86.

been generally agreed that a hair examiner should not claim that a particular individual *is* the source of questioned hair, only that the individual could be the source. Prominent MHCA researcher Barry Gaudette took this position as early as 1976.⁵² However, as the FBI review exposed, somehow, despite this scientific caution, MHCA experts began doing just that: stating or implying that an individual was the source of a hair. And at the center of this bad habit was the FBI.

By 1985, even “the more senior FBI representatives accepted that the days of claiming hair examinations could uniquely identify an individual were over but still believed that hair examinations could offer strong support, with discussion centring on how this might be expressed.”⁵³ This left examiners reporting that hairs were “consistent,” that an individual “could be” the source of the hair, and that the individual was included as a possible source of the hair in a donor pool “of unknown size.”

However, a report or testimony that an individual “could be” the source of a hair was maddeningly vague. “Could be” necessarily encompassed a wide range of possibilities from “very likely to be the source” to “very unlikely to be the source.” The jury or judge would inevitably want to know more. Did the donor pool “of unknown size” consist of 10 people or 10 million? The answer, self-evidently, lay in having some information about how common the features that had been found consistent were in the population. But hair examiners lacked this information. Two former FBI unit chiefs explained the issue clearly:

Myron T. “Mike” Scholberg, hair unit chief from 1978 to 1985, and Alan T. “Al” Robillard, chief from 1988 to 1990, said that in hindsight, they were not properly trained to answer a crucial question for jurors: How often might the hairs of different people appear to match? The truth is that there was no scientific way to know.⁵⁴

During the 1970s, Gaudette published widely cited studies attempting to address these questions. However, Gaudette’s studies were subject to methodological criticisms.⁵⁵ Thus, “at the start of the 1980s, there was considerable debate as to the value of hair examination, fueled in large measure by the well-known 1974 paper from Gaudette, and a real sense that perhaps protein analysis could be the way forward to achieving a more objective assessment of individuality. There followed a number of major conferences and meetings that brought together forensic scientists across the world to discuss hair examination,” one in Chicago in 1982 and a workshop in Quantico, Virginia in May 1983, sponsored by the FBI, at which an informal Committee for Hair Examination was established.⁵⁶ A third was the International Symposium on Forensic Hair

⁵² Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 251.

⁵³ *Id.*, 253.

⁵⁴ Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*.

⁵⁵ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 243; PCAST, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*, 118; Walter F. Rowe, *The Current Status of Microscopical Hair Comparisons*, 1 *The Scientific World* 868, 873 (2001); Stafford Smith & Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil*; Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*.

⁵⁶ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 72; Oien, *Forensic Hair Comparison: Background Information for Interpretation*; Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 244.

Comparisons, also sponsored and hosted by the FBI, in June 1985.⁵⁷ At this meeting, there was extensive discussion of the problem of what kind of evidentiary report could be scientifically supported based on MHCA. The Committee for Hair Examination presented some recommended reporting language by its Report Writing, Conclusions, and Court Testimony Subcommittee,⁵⁸ some of which would be considered “exceeding the limits of science” by the FBI today.⁵⁹

The FBI did not formally adopt the guidance of the Report Writing Subcommittee or any other formal guidance following the Symposium. In retrospect, this Symposium would be pointed to as evidence that the FBI was aware of the scientific bankruptcy of its reporting practices for decades prior to the 2015 FBI review. The ABS Group, which conducted a root cause analysis for the FBI after the release of the results of the 2015 Review, suggests that FBI examiners insufficiently consulted the Symposium proceedings, although they were available to them.⁶⁰ The problem, which eventually led to the FBI review, “was that for the following decade or more, evidence presented by FBI hair examiners had an underlying, or implied, view that hair evidence could exclude all other potential sources because a coincidental ‘match’ was a rare event.”⁶¹ Morgan contends that “Right after the 1985 symposium, many analysts limited their testimony to ‘similar physical characteristics.’” However, “Over time,” what Morgan calls “probative value creep” drew “almost all analysts” into using “the strongest possible language—‘consistent with’ the defendant—that was suggested during the 1985 symposium.”⁶²

In 1991, a complaint was filed over the use of the term “perfect match” by an FBI examiner. Laboratory management counseled against the use of this term.⁶³ Nonetheless, in 1995, at the nationally televised trial of O.J. Simpson, an FBI hair examiner attempted to use the term “match.” The trial judge barred the term, and the Hairs and Fibers Unit Chief directed examiners to cease use of it.⁶⁴ It was not, however, until 2009, when the [Donald Gates](#) exoneration, a false conviction to which MHCA made a major contribution,⁶⁵ occurred, that concerns gained traction:

At Gates’ trial in 1982, key testimony came from FBI forensic analyst Michael Malone who said that Gates’ hairs were “microscopically indistinguishable” from hairs found on the victim’s body.

In 2012, the FBI review began, and the initial results were announced in the 2015 press release. The FBI then commissioned a root cause analysis, which was conducted by ABS Group and published in 2018.⁶⁶

⁵⁷ Proceedings of the International Symposium on Forensic Hair Comparisons, International Symposium on Forensic Hair Comparisons, Quantico, Virginia, Laboratory Division, Federal Bureau of Investigation, U.S. Dept. of Justice, (June 25-27); Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 80-81; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 99.

⁵⁸ Norman E. Erickson et al., *Committee on Forensic Hair Comparison - Subcommittee 4 - Report Writing, Conclusions and Court Testimony Address at International Symposium on Forensic Hair Comparisons (June 25-27, 1985)*.

⁵⁹ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 73.

⁶⁰ *Id.*, 76.

⁶¹ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 254.

⁶² John Morgan, *Wrongful Convictions and Forensic Science Errors: Case Studies and Root Causes*, 70 (2023).

⁶³ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 76-77.

⁶⁴ *Id.*, 79.

⁶⁵ Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 76-93.

⁶⁶ For details on the procedures of the Root Cause Analysis, see ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 47-54.

V. THE FBI REVIEW

Despite the volume of press attention it generated, remarkably little is known from a research perspective about the FBI review. The completion of the study has never been announced, and it has never been reported in the conventional manner by which scientific studies are usually reported (i.e., publication in a journal).⁶⁷ This creates numerous obscurities about the Review. To begin with, the transcripts analyzed have not been made public, which would, for example, allow other researchers to examine the statements that the reviewers found “appropriate” or “erroneous.” As one scholar has noted, unless “all the reports and transcripts are made available for more rigorous research” or researchers’ questions about how the data were coded are better answered, “an independent, scientifically credible evaluation of any statistics on the prevalence of errors and their significance is not feasible.”⁶⁸ Another notes, “Until such time as a full report is forthcoming, one can only speculate as to the underlying reasons why the FBI examiners got it so wrong.”⁶⁹ However, scholars despair that such a report will ever be produced: “It appears unlikely that a final report from this review will be forthcoming.”⁷⁰ As Robertson and Brooks summarize the situation: “The final outcomes of the enquiry into former FBI practices in reporting hairs have yet to be made public, if indeed they ever are made public.”⁷¹

The unavailability of the data flies in the face of current trends in science toward “open data.” While it is not surprising that the data is not public, it is not clear that it is necessary or even justifiable not to make it so. The transcripts that comprise the FBI review’s data set are verbatim records of public trials that were conducted and recorded at taxpayer expense—arguably there is no valid reason that they should not be made public.⁷² However, the issue is worse than that: even the *names* of the cases included in the review have never been made public. Knowing the names of the cases would allow researchers to compare other cases to the cases in the FBI review, or to replicate the FBI review itself.

The problem is worse still because important aspects of the *data analysis* have never been made public. Although the breakdown between “appropriate” and “erroneous” testimony has been made public, we know that the FBI review coded three types of erroneous testimony labeled “Error Types” 1, 2, and 3. But no breakdown of the distribution of these error types has even been made public.

Finally, to make matters still worse, the *methods* of the Review have never been made public. By “methods,” we mean such questions as: what rules governed the assignment of cases (or statements) to the “appropriate” and “erroneous” categories or to the Error Type 1, 2, and 3 categories? For this study, we had to try to piece together the Review’s methods from documents obtained by Freedom of Information Act requests and personal communications with people knowledgeable about the study.

⁶⁷ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 252; Robertson & Brooks, *Historical Context to Contemporary Practice in Hair Examinations*.

⁶⁸ Kaye, *Ultrarepidarianism in Forensic Science: The Hair Evidence Debacle*.

⁶⁹ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 252.

⁷⁰ Robertson & Brooks, *Historical Context to Contemporary Practice in Hair Examinations*, 18; See also Garrett, *Autopsy of a Crime Lab: Exposing the Flaws in Forensics*, 92.

⁷¹ Robertson & Brooks, *Historical Context to Contemporary Practice in Hair Examinations*, 22.

⁷² Simon A. Cole et al., *Assessing the Feasibility of Building a Database of Trial Transcripts Containing Scientific Testimony*, Report, Project on Scientific Knowledge and Public Policy, (Oct. 8, 2007); Glenn Koppel, *A Tale of Two Counties: Divergent Responses in Los Angeles and Orange County Superior Courts to the Ban on Electronic Recording in California Court Reporters Association v. Judicial Council*, 37 San Diego Law Review 47 (2000).

All of the items discussed above—details about the data, the analysis, and the methods—would be made known to the public in a conventional scientific publication. While the unavailability of the raw data is not that unusual in contemporary science despite the trend toward open data sharing, the unavailability of the analysis and methods would not be considered acceptable for any conventional scientific research report.

VI. WHAT WE KNOW ABOUT THE FBI REVIEW

Despite the above limitations, in this section we will summarize what we do know about the FBI review. The process began in July 2012 when the four organizations agreed to undertake the review. On November 9, the four organizations agreed to a statement concerning the “limits of science” of MHCA. The statement defined “appropriate” testimony and three “error types” of “erroneous” testimony.⁷³ The FBI Laboratory Division Scientific Review Committee and attorneys from the FBI Office of General Counsel then reviewed a sample of 12 transcripts. Based on this pilot study, the reviewers developed a Review Guidance for coding the transcripts as to appropriateness and, if applicable, error type (e.g., Figure 3).⁷⁴

⁷³ FBI, *Microscopic Hair Comparison Analysis: Standard for Review of Testimony and Laboratory Reports*, U.S. Department of Justice, (Oct. 3, 2013).

⁷⁴ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*, Federal Bureau of Investigation, (May 14, 2013).

It defined the three Error Types as follows:

Error Type 1: The examiner stated or implied that the evidentiary hair could be associated with a specific individual to the exclusion of all others. This type of testimony exceeds the limits of the science.

Error Type 2: The examiner assigned to the positive association a statistical weight or probability or provided a likelihood that the questioned hair originated from a particular source, or an opinion as to the likelihood or rareness of the positive association that could lead the jury to believe that valid statistical weight can be assigned to a microscopic hair association. This type of testimony exceeds the limits of the science.

Error Type 3: The examiner cites the number of hair analyses worked in the lab and the number of samples from different individuals that could not be distinguished from one another as a predictive value to bolster the conclusion that a hair belongs to a specific individual. This type of testimony exceeds the limits of the science.⁷⁵

The public has access to two snapshots of the data: one from the 2015 press release, and one from the ABS Group Root Cause Analysis released in 2018. What we know about the results of the study from these two snapshots is summarized in Table 1.

Table 1. Summary statistics of FBI Hair Review from two snapshots. Sources: ABS Group (2018), FBI (2015a), (2015b), Innocence Project (2013), Reimer (2013).

1	2	3	4	5	6	7
Data release	Cases with MHCA evidence	Cases reviewed	Further review			
FBI Press Release March 2015	21,700	~3,000	~500	Transcripts 268	Erroneous 257 (96%) Appropriate 11 (4%)	
ABS Group Root Cause Analysis June 2018	23,557	3,499	2,213	Transcripts 484 Reports 1,729	Erroneous 450 (93%) Appropriate 34 (7%) Erroneous 856 (50%) Appropriate 873 (50%)	2,213 Erroneous 1306 (59%) Appropriate 907 (41%)

The FBI began with cases in which MHCA was involved in some way. As shown in Column 2, by June 2018, there were 23,557 such cases.⁷⁶ From these cases, the researchers selected only those cases in which the evidence was both inclusionary (as opposed to, say, exclusions or inconclusive results) and probative, the defendant was convicted, no DNA analysis was performed, the analysis was performed prior to 2000, and the FBI provided a report. As

⁷⁵ FBI, *Microscopic Hair Comparison Analysis*, Federal Bureau of Investigation, (Nov. 9, 2012).

⁷⁶ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 17.

shown in Column 3, this yielded 3,499 cases by June 2018.⁷⁷ As shown in Column 4, by 2015 the FBI had reviewed around 500 of these cases; by 2018 they had reviewed 2,213. The criterion for advancing to this further stage of review is not clear, but it appears to have had to do with the availability of a transcript or laboratory report and the elimination of duplication for co-defendants.⁷⁸

As shown in Column 5, the 2015 report covered only trial transcripts and an unspecified number of what it called “transcript substitutes/proxies . . . briefs or other court filings that appear to reflect portions of FBI testimony.”⁷⁹ The 2018 Root Cause Analysis Report included 484 transcripts, but also “1,729 reports were reviewed when transcripts were not available.” The transcripts dated from 1971-2003, and the reports dated from 1973-2000.⁸⁰

As shown in Column 6, in 2015 erroneous testimony (the definition of which will be discussed below) was found in 257 cases—hence the 96% figure that generated so much attention.⁸¹ By 2018, 450 of 484 reviewed transcripts (93%) had been coded as containing erroneous testimony. In 2018, the results of the same analysis of 1,729 laboratory reports were also reported. Laboratory reports split evenly between appropriate and erroneous.

As shown in Column 7, combining transcripts and laboratory reports yielded approximately a 60-40 split in favor of “erroneous” evidence.

The Root Cause Analysis Report offers no explanation for the astonishing 86% increase in erroneous reporting in trial testimony compared to laboratory reports. Nor, as far as we know, has anyone else ventured an explanation. Several possible explanations present themselves. Perhaps the most mundane is that testimony is longer than written reports, which during the period under study were often quite brief. More words offer greater opportunity to make an erroneous statement. In addition, while written reports followed templates,⁸² testimony was not standardized; “there were no explicit criteria regarding statements that exceeded the limits of the science.”⁸³ This explanation would not indicate any behavior differences for MHCA experts between report-writing and testifying. But it would show that having MHCA experts testify introduced a far greater risk that they would communicate an erroneous statement.

The most plausible explanation for why more erroneous statements appeared in testimony than in reports, however, is a desire to win the trial—perhaps due to prosecutorial pressure, adversarial emotions elicited on cross examination, or a simple desire to “be more helpful” (have

⁷⁷ Id.

⁷⁸ The Root Cause Analysis Report notes that “Of these [3,499] cases, some were never brought to trial for a variety of reasons (e.g., a guilty plea, a decision not to prosecute); however, a laboratory *report* would still have been produced,” but it does not say how many cases fell into this category. It also reports that “It is unknown how many of these cases went to trial, but it is certain that only a fraction of them progressed to trial in which *MHCA examiners* actually testified.” Id. (original emphasis). The FBI *Hair Comparison Analysis Review Data*, (Mar., 2015). says that cases were also removed because of duplication due to co-defendants and one case with an incomplete transcript.

⁷⁹ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*, 1.

⁸⁰ The Root Cause Analysis notes that in cases in which it obtained a transcript after reviewing a report in the same cases, it reviewed both documents. Therefore, how many cases the 2,213 documents represent is not known. It is also not known how many of these cases went to trial. Do the 1,729 reports represent cases that did not go to trial, or cases that went to trial in which a transcript could not be obtained? The Root Cause Analysis is not clear enough to answer this question. ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 17-18. (original emphasis).

⁸¹ FBI, FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review.

⁸² *Gates v. District of Columbia, et al.*, 1:11-CV-00040-RWR Deposition of Michael P. Malone, 97 (D.D.C. 2013), on file with the authors.

⁸³ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 66.

more of an effect on the outcome). Some combination of these forces may have induced the examiners to cross over the threshold into erroneous testimony, a threshold they had managed to respect in the dry confines of writing a laboratory report. It has been suggested that prosecutors invited experts to make erroneous statements by communicating to them a desire for statistics in MHCA testimony. MHCA's inability to quantify the size of the pool of potential donors contrasted unfavorably with disciplines which could, such as serology and, later, DNA profiling. This may have prompted a prosecutorial desire for numbers in MHCA testimony and, it is alleged, perhaps instruction to experts to generate statistics based on experience (which would later be called Error Types 2 and 3).⁸⁴

As early as 1985, forensic scientist Peter DeForest seems to have anticipated this strengthening of the evidence between written report and spoken testimony:

I have a problem with the divergence from a laboratory report in which the conclusion is these hairs could have shared a common origin to the presentation of testimony in court when the expert says something to the effect that, "Yes, these hairs were found to be similar and in my experience I have examined thousands of hairs and I have never found two hairs from different sources that were alike." I think that is very misleading and it is not substantiated by any data.⁸⁵

This explanation suggests an analogue to the well-known "trial penalty"⁸⁶ (greater sentences for defendants who decline plea bargains) in the area of expert evidence: evidence that may have been modest during plea negotiations becomes strong if the defendant forces the state to take them to trial.

Still another possibility is the difference between speaking and writing. MHCA experts may have been more careful in writing than in speaking in court. This illustrates one of the undesirable consequences of U.S. law's preference for spoken over written expert evidence.⁸⁷

There are other possible explanations whose plausibility would depend on knowing something we don't know: whether the 1,729 reports came from cases that went to trial, cases that did not go to trial (and for what reasons), or some combination of the two. Around 95% of felony cases end in dismissals or guilty pleas, not trials. Trials with convictions usually generate transcripts of testimony, but cases that end in guilty pleas usually do not. We don't know directly, but it seems likely that the convictions with forensic evidence but without transcripts that the Root Cause Analysis examined in 2018 were primarily based on plea bargains, while the convictions with transcripts examined in 2015 were primarily returned after trials. If so, one explanation for the difference in rates of erroneous testimony may well be due in part to case selection bias: Cases that go to trial are generally those in which the non-MHCA evidence of guilt or innocence is not clear

⁸⁴ Gates v. District of Columbia, et al., 329-330. The reasonable assumption that jurors perceive MHCA evidence whose strength is presented in quantitative terms to be more persuasive than evidence whose strength is presented in qualitative (verbal) terms may not correct. An experimental study of judges and jurors presented with MHCA evidence found that both were "more persuaded by qualitative information than quantitative information." Dawn McQuiston-Surrett & Michael J. Saks, *The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear*, 33 Law & Human Behavior 436, 444 (2009). More generally, see also Dawn McQuiston-Surrett & Michael J. Saks, *Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact*, 59 Hastings Law Journal 1159, 1189 (2008).

⁸⁵ Proceedings of the International Symposium on Forensic Hair Comparisons 204.

⁸⁶ Candace McCoy, *Plea Bargaining as Coercion: The Trial Penalty and Plea Bargaining Reform*, 50 Crim. L.Q. 67 (2005).

⁸⁷ Samuel R. Gross, *Expert Evidence*, 1991 Wis. L. Rev. 1113, 1215 (1991).

one way or the other, and the temptation to put a testimonial thumb on the scale is great because it might make the difference between acquittal and conviction. Forensic analysts may also be exposed to more irrelevant but potentially biasing information about the case during preparation for trial.

To further explore these issues, it would be useful to compare written reports and spoken testimony from the same cases. We do not have sufficient data to undertake such a comparison because we have access to few written reports. But in at least one case we can compare a written report and oral testimony. In the [Emerson Stevens](#) case, former FBI examiner Myron Scholberg's written MHCA report simply said, "A single head hair of Caucasian origin found in this item is microscopically like the head hairs from the victim and, accordingly, could have originated from her"—an "appropriate" statement.⁸⁸ Scholberg gave similar testimony at the trial. But at the trial Scholberg also testified that the hair had "a unique pattern" (Error Type 1), and, according to the prosecutor's summation, Scholberg testified that while "it's a possibility it's not her hair, it is unlikely in the populace there will be two peoples' hair I cannot tell the difference in with my great experience" (Error Type 2). In his trial testimony, Scholberg added erroneous statements that were not contained in his written report. MHCA evidence that would have been considered "appropriate" if the case had been resolved based only on the written report became "erroneous" when the case went to trial.

There remains a great deal about the FBI review that we do not know. We do not know the names of the cases that were reviewed, nor have the transcripts or reports been made publicly available. The descriptions of the FBI study note that three general "types" of erroneous testimonial statements were identified.⁸⁹ However, we do not know the distribution of these error types among the 450 erroneous transcripts and 856 erroneous reports. The Root Cause Analysis Report does make clear that for transcripts, Error Type 2 was the most common and Error Type 3 the least and that Error Type 2 was by far the most common for reports, but it does not provide sufficient data to know the precise distributions.⁹⁰

VII. DATA AND METHODS

The FBI review analyzed only transcripts of testimony from a single forensic service provider, the FBI laboratory. As noted above, MHCA was also used in state and local laboratories and state courts, and the cases the FBI study examined are, therefore, only a sample of MHCA testimony nationwide. To address the questions raised about the extent to which the problems at the FBI extended into state and local laboratories, however, requires data sets of state and local MHCA testimony.

⁸⁸ Stevens v. Bennett, 3:20-cv-00352-MHL Appendix to Petition for Writ of Habeas Corpus, 103 (E.D. Va. 2020).

⁸⁹ Reimer, *The Hair Microscopy Review Project: An Historic Breakthrough for Law Enforcement and A Daunting Challenge for the Defense Bar*; FBI, *Microscopic Hair Comparison Analysis: Standard for Review of Testimony and Laboratory Reports*.

⁹⁰ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 85-100. The Root Cause Analysis Report provides two figures representing data about error type. From Figure 12, it is possible to infer that Error Type 2 is the most prevalent. But it is impossible to derive any numbers from this figure because it is a longitudinal graph intended to represent temporal trends, rather than to provide basic information about the frequency of the 3 error types. Therefore, the Y axis measures errors per transcript (rather than simply the number of errors). The number of transcripts is not represented anywhere in the figure.

Figure 18 does show the raw number of errors of different types. However, since no table is provided and the figure contains no data labels, a researcher would have to guess at the actual number of errors by eyeballing the graph's bars. All a researcher can infer from this figure is that, again, Error Type 2 is most common.

This study uses data from the National Registry of Exonerations (NRE). The NRE is a national archive of exonerations from 1989 through the present. The NRE's definition of exoneration is available [online](#). At the time of the final analysis (July 12, 2023), the NRE housed 3,340 exoneration cases.

We were able to identify two other potential data sources on the use of MHCA in cases involving state and local MHCA. However, neither yielded as large a number of cases as the NRE. In the discussion below, we sometimes use these sources for comparison purposes:

1. **Westlaw.** Westlaw (WL) maintains a small database called Expert Materials, consisting of 3 sub-databases of (1) trial transcripts; (2) expert reports and affidavits; and (3) expert depositions. Like the FBI data, the Westlaw documents are transcripts and reports (or affidavits and deposition transcripts). However, on March 27, 2020, searching Expert Materials for “microscop* & hair & compar*” yielded only 56 documents. After eliminating duplicates, civil cases, and cases already included in the NRE data set, only 32 documents remained. MHCA was only actually used in 9 of these cases, and in three cases the conclusion was an “exclusion.” That left only 6 cases in which conclusions of inclusion were rendered. All 6 represented testimony, not written reports. (Deposition testimony was considered testimony.)

2. **Texas Forensic Science Commission.** As mentioned above, the Texas Forensic Science Commission (TFSC) conducted a statewide audit of MHCA cases. The purpose of the audit was legal, not research; it was designed to identify individuals who might be falsely incarcerated, or even facing execution, because of MHCA evidence. Nonetheless, the data can be used for research purposes. It is the only data from a statewide audit that has been made available for research purposes. However, the data has never been published.

The TFSC used a sub-sampling method to request case files from large crime laboratories. It supplemented these case files with a Lexis-Nexis search for cases. As shown in Table 2, this yielded an initial sample of 813 cases. These cases were then reviewed to determine whether a positive probative association had been made using MHCA and a criminal conviction had resulted. This reduced the sample to 360 cases. These cases were then narrowed to 79 in which it was believed that a trial transcript might be available. Of these 79 cases, a transcript was available in 73.⁹¹ A further 15 cases were eliminated for reasons such as there being no MHCA testimony in the transcript or the MHCA expert being an FBI, not a Texas, examiner, leaving a total of 58 cases for which the researchers were able to determine whether the testimony was appropriate. The TFSC generally followed the FBI guidelines for coding cases, but also slightly deviated from those practices on some points. For example, the TFSC also assigned 3 error types, and it duplicated FBI Error Types 1 and 2. However, while FBI Error Type 3 concerns the use of experience to vouch for accuracy, TFSC Error Type 3 is a more general category of “any other potentially misleading statements or inferences,” which *includes* “things like using the number of cases an examiner worked in his or her career to increase the strength of an association.”⁹²

Like the FBI, the TFSC essentially defined as “appropriate” testimony that did not fall afoul of the three error types. Of the 58 cases analyzed, 33 (56%) were appropriate. All of these reviews were based on trial transcripts, not written laboratory reports.

⁹¹ TFSC, *Update on Review of Hair Microscopy Cases*, Texas Forensic Science Commission, (Jan. 27, 2016).

⁹² TFSC, *Guidelines for Review of Transcripts Texas Hair Microscopy Case Review*, Texas Forensic Science Commission, 1 (n.d.).

Table 2. Texas Forensic Science Commission audit.

	Initial sample	Positive probative association	Trial transcript may be available	Transcript available	Appropriateness determined		Number (percentage of cases reviewed)
August 2017	813	360	79	73	58	Erroneous	25 (44%)
						Appropriate	33 (56%)

While a larger data set than available in Westlaw, the TFSC data set remains smaller than the data set offered by NRE. As explained below, the NRE yielded 99 non-FBI MHCA cases in which testimony of “inclusion” was given. However, the NRE data set differs from the others in several important ways. Most importantly, the NRE data set is a sample of *exonerations*—cases in which a defendant was convicted and then relieved of the consequences of that conviction because of new evidence of innocence. Although the FBI study was, in part, triggered by high-profile exonerations involving MHCA, it was not a data set of exonerations. Rather, it was a data set of *cases*. The researchers sought to pull every transcript in which an FBI expert gave expert evidence about MHCA. While all the defendants in the NRE data set were exonerated, the vast majority of defendants in the FBI data set remain convicted.

Second, the other data sets consisted mostly of trial transcripts. The FBI, TFSC, and Westlaw sought to obtain transcripts of all of the cases in their data sets, and they had the resources and legal authority to do so. In a small number of cases, the FBI researchers were forced to rely on “proxy transcript substitutes,” such as “briefs or other court filings.”⁹³ Transcripts are notoriously difficult and expensive to obtain in the U.S. criminal justice system. The NRE rarely obtains transcripts of the cases it lists. In some cases, we were able to obtain transcripts from other sources, such as [The Innocence Record](#) and [Convicting the Innocent](#). (For more detail, see [Appendix 1](#).)

Table 3 summarizes the attributes of each data set. If we compare the present study to the FBI review, we see that they have different advantages and disadvantages. The FBI review is much larger and had access to complete transcripts in almost all cases. However, it was limited to one forensic service provider, the case names remain anonymized, the data has not been made publicly available, the study is not published, and there is no usable contact to answer questions or respond to requests for clarification.

Table 3. Descriptive comparison of all data sets on U.S. MHCA.

Data set	Sample	% Testimony (as opposed to reports)	Laboratory scope	Unbiased sample of cases (exonerated and not exonerated)	Complete transcripts used	Case names identified	Data available for inspection or sharing	Published	Responds to questions and requests for clarification
FBI Review	2213	22%	FBI	X	X				
NRE	99	≥88%	National			X ⁹⁴	X	X	X
TFSC	58	100%	Texas	X	X	X	X		X
WL	6	100%	National	X	X	X	X	X	

⁹³ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*.

⁹⁴ Exonerees are not human subjects. The information the NRE compiles about exonerations is based on public records. The NRE does not initiate contact with exonerees and does not use private identifiable information. The inclusion of case names is crucial for the credibility of counts of exonerations.

A. THE NRE DATA

In total, we reviewed 471 of the 3,340 exoneration cases housed in the NRE on July 12, 2023, which we suspected may have included MHCA. For details on how and why we targeted those cases, see [Appendix 1](#). Of these 471 cases, MHCA was used in 184. These cases were then coded according to the type of testimony given. Possible types of testimony were:

- **Inclusion**
- **Exclusion**
- **Inconclusive**
- **Inconsistent but not excluded.** Testimony was proffered by the State to the effect that hairs were inconsistent with the POI, but that the POI could not be excluded as the source of the hair. Such cases were not included in the FBI study and therefore are not counted as “inclusions” in our study either. However, they may reasonably be viewed as inclusions of a sort. They will be discussed in further detail in Section IX.A.4.
- **Racial attribution only inclusion.** The expert attributed a racial type, i.e., “Negroid,” “Mongoloid,” “Caucasian,” or “Hispanic,” to the hair, but did not associate the hair with the defendant as an individual. The evidence implicated the defendant.
- **Racial attribution only.** The expert attributed a racial type, i.e., “Negroid,” “Mongoloid,” “Caucasian,” “Hispanic,” to the hair, but did not associate the hair with the defendant as an individual. The evidence did *not* implicate the defendant.
- **Damage.** One case ([Gary Gauger](#)) in which the MHCA evidence concerned only damage to the hair. The forensic scientist testified that hairs near a victim’s body could have been broken and stretched in a manner consistent with the defendant’s confession, but also acknowledged they could have been broken during combing and brushing. The results of that coding are shown in Table 4.

Table 4. Distribution of testimony types in exoneration cases in which microscopic hair comparison was used (n=184).

Testimony Type	Number
inclusion	117
exclusion	49
race only	9
inconsistent but not excluded	6
race only inclusion	1
damage	1
inconclusive	1
Grand Total	184

In 18 inclusion cases, we know that the expert witness was an employee of the FBI. However, there may be other cases in which data limitations prevented us from learning that the expert witness was an FBI employee. Because one of our goals was to compare state and local cases to FBI, we separated these 18 FBI cases. To these 18 inclusion cases, we added one case, [Rudolph Holton](#), in which a “race only inclusion” was made by the FBI (for more detail, see Section IX.A.5), giving us a total set of 19 FBI exoneration cases.

Although the names of the cases in the FBI study have never been made public, we know that at least some of these 19 cases were included in the FBI review, but we don't know how many.⁹⁵ In only 6 of the 19 FBI exoneration cases, do we have sufficient documentary evidence to be confident that they were included in the review, but there may be more.⁹⁶

This left us with 99 non-FBI exoneration cases in which MHCA *inclusions* were made. In most cases the report was given in the form of testimony. In only 2 cases, do we know that it was given in the form of a written report (Table 5).

Table 5. Type of report in NRE inclusion cases. Most of the cases labeled "don't know" were trials and probably did have testimony. They are labeled "don't know" because we don't have evidence that a forensic scientist actually testified, rather than, e.g., a written report being introduced as an exhibit.

Report type	Number
Testimony	88
Don't know	9
Report	2
Grand Total	99

In 15 of the 99 cases, representing 12 different experts,⁹⁷ we had affirmative evidence that the MHCA expert was trained by the FBI, such as testimony about the analyst's FBI training. This probably understates the number of experts who were in fact trained by the FBI because we did not have access to full transcripts in every case and, even when we did, the training may not have been elicited during testimony.

In 44 cases, the MHCA expert was employed by the state. The expert was employed by a city in 14 cases and a county in 6. In 35 cases, we were not able to determine the expert's employer (Table 6).

Table 6. Expert's employer in non-FBI MHCA inclusion cases.

Employer	Number
State	44
Don't know	35
City	14
County	6
Grand Total	99

⁹⁵ Neither the FBI, through Freedom of Information Act requests, nor its partners, the NACDL and IP, were able to provide us with the names of the cases included in the FBI review.

⁹⁶ The six cases are: [John Ausby](#): United States v. Ausby, 916 F.3d 1089, 1092 (App.D.C., 2019); [Richard Beranek](#): Norman Wong to Ismael Ozanne, Re: State of Wisconsin v. Richard E. Beranek, Case No. 1989CF000946 (May 5, 2015); [Dennis Butler](#): United States v. Butler, Crim. No. 1717-70, Motion to Vacate Conviction, Exhibit A (D. D.C., Sept. 16, 2016); [Michael Jones](#): Jones v. United States, No. 15-CO-1104 (D.C., 2019); and [George Perrot](#): Commonwealth v. Perrot, No. 85-5415, Defendant's Post-hearing Memorandum in Support of his Motion for a New Trial (Hampden Super. Ct., 2015); [Joseph Sledge](#), Norman Wong to Jon David, Re: State of North Carolina v. Joseph Sledge, Jr. (Nov. 26, 2014).

⁹⁷ One of these experts (Joyce Gilchrist) testified in 4 different cases.

B. METHODOLOGICAL DIFFERENCES BETWEEN THE STUDIES

Our initial goal was to follow the FBI review's methods as closely as possible. However, the FBI review's methods are far from transparent. One scholar had to file a Freedom of Information Act request in order to get more information about the FBI review's coding methods.⁹⁸ In [Appendix 2](#), we go into detail about how we reconstructed the FBI review's methods in order to have a workable coding system to apply to our own data.

Despite our efforts to follow the FBI review's coding system as closely as possible, a few discrepancies were unavoidable. First, the FBI review had access to complete transcripts, coded the entire transcript for erroneous statements, and counted the number of erroneous statements. Only if there were no erroneous statements, was the case coded "appropriate." Thus, a single case could contain both appropriate and erroneous statements and up to all three error types.

In contrast, for cases coded "erroneous," we assigned each case only a single error type.

If we did have access to a transcript, of course, we reviewed the full testimony of the MHCA expert (or as much as we had). However, because we did not have access to full transcripts in most cases, we often had access to only the "highlights" of the testimony from a legal document or a news report. In the FBI review, many transcripts apparently contained several erroneous statements. According to the 2018 snapshot, there were more than 2,000 testimony errors in 450 transcripts,⁹⁹ around 4 or 5 per transcript. In the FBI review's coding system, erroneous statements are not cured by an appropriate statement or by "limiting language" like "it's not an exact science" or "we can't make absolute identifications."¹⁰⁰ Therefore, we may have missed erroneous statements made that were not prominent in the testimony. Kaye gives an example of a case to which an error was assigned for a statement that was about the fact that the hair in storage was entangled with a thread, an issue seemingly peripheral to the central question of whether the victim was the source of a questioned hair.¹⁰¹

Second, the FBI review apparently consolidated cases involving co-defendants, so that they count as only one case in the study. The unit of analysis for the NRE in all its research has always been, and remains, the individual exoneration. In addition, the NRE counts False or Misleading Forensic Evidence in cases where the evidence implicates a co-defendant if the prosecution's theory of the crime links the co-defendants.¹⁰² We followed both of these conventions in our study. Therefore, the two studies count co-defendant cases differently. There are several multiple defendant cases in our data set. These cases count as the number of cases as there were defendants, but they would have counted as only one case in the FBI review.

The TFSC coded its own cases using a modified version of the FBI review coding system. The TFSC supplemented what little is known about the FBI coding system with its own set of supplementary coding instructions. We used these instructions as a supplemental authority to guide our coding on issues where the FBI coding was obscure. For example, the TFSC coding guidelines were helpful for clarifying that the use of terms like "Same Exact Microscopic

⁹⁸ Kaye, *Ultracrepidarianism in Forensic Science: The Hair Evidence Debacle*.

⁹⁹ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 17.

¹⁰⁰ Vanessa Antoun et al., Hair Microscopy Review Webinar Address at Innocence Network (Dec. 4, 2018).

¹⁰¹ Kaye, *Ultracrepidarianism in Forensic Science: The Hair Evidence Debacle*.

¹⁰² For example, if the State's theory at trial is that the defendants acted together, then forensic evidence implicating a co-defendant implicates the defendant as well.

Characteristics” is erroneous.¹⁰³ The TFSC provided its data to us with its coding. However, we recoded all TFSC cases. We did not have any disagreements with the TFSC as to coding appropriateness. Nor did we have any disagreements with the TFSC coding as to error type; however, like the FBI, the TFSC allowed for more than one error type to be applied to a single case. We recoded the TFSC cases, choosing the error type that best fit the testimony. In addition, where more than one error type clearly occurred, we coded the lower error number, which was deemed more serious (e.g., Error Type 1 “trumped” Error Type 2, but Error Type 2 trumped Error Type 3).

VIII. RESULTS

A. ERRONEOUS AND APPROPRIATE TESTIMONY

Using the FBI review criteria for “appropriate” and “erroneous” testimony, we coded the NRE cases as to whether the evidence was “appropriate” or “erroneous,” yielding the results shown in Table 7.

Table 7. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in non-FBI exoneration cases in which inclusionary MHCA was used (n=99).

Testimony	Number	%
Appropriate	60	61%
Erroneous	39	39%
Grand Total	99	100%

The results differ greatly from the FBI review. Whereas the FBI review found 93% erroneous evidence, we found erroneous evidence in only 39% of cases.

The 19 FBI exoneration cases yielded findings more similar to those found in the FBI review. Eighty-four percent of the evidence was “erroneous,” and in only three cases ([Rudolph Holton](#),¹⁰⁴ [Kevin Martin](#),¹⁰⁵ and [Marvin Thomas](#)) was the evidence “appropriate” (Table 8).¹⁰⁶ The

¹⁰³ TFSC, *Guidelines for Review of Transcripts Texas Hair Microscopy Case Review*.

¹⁰⁴ Holton was an unusual case in that it involved only testimony about racial attribution. See Section IX.A.5.

¹⁰⁵ Martin was also an unusual case because he pled guilty, based on an appropriate written report. The prosecutor then made an erroneous statement, telling the attorneys and judge that “the FBI technician is prepared to say it was Kevin Martin’s hair on Mr. Brown’s sneaker” (Error Type 1) (see Section IX.E). The Martin case demonstrates one avenue through which defendants can be convicted even when the analyst gives appropriate evidence: prosecutors may still mischaracterize the reports, whether during plea negotiations or trial closings.

¹⁰⁶ Antoun et al., *Hair Microscopy Review Webinar*, 7:45., report that the testimony in the [Joseph Sledge](#) case was coded appropriate in the FBI review. However, it appears that the FBI originally coded the Sledge testimony appropriate, but the IP and NACDL dissented from that judgment. The FBI concluded there were no inappropriate statements and noted that the MHCA expert James Frier had used “limiting language” by saying “Hairs do not constitute a basis for positive personal identification as fingerprints.” However, the Review guidelines explicitly state that “limiting language” does not cure inappropriate statements. *Supra* note 100. The IP and NACDL concluded that Frier committed Error Type 2 for saying that the crime scene hairs “were microscopically alike in all respects to some of the hairs comprising the known pubic hair sample from the defendant and could have originated from him or another individual of the same race whose hairs exhibited the exact same microscopic characteristics.” In the end, the DOJ apparently was persuaded the testimony was inappropriate because the attorneys in Sledge’s case were notified that the testimony in his case was erroneous. Norman Wong to Jon David, Re: State of North Carolina v.

rate of erroneous testimony was still much lower than was found in the FBI review. The stark difference between our findings for FBI and non-FBI cases suggests that there is something different about FBI cases, rather than that there is something different about our data.

Table 8. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in FBI exoneration cases in which inclusionary MHCA was used (n=19).

Testimony	Number	%
Appropriate	3	16%
Inappropriate	16	84%
Grand Total	19	100%

The combined results for FBI and non-FBI exoneration cases are shown in Table 9. Evidence was erroneous in 47% of the exoneration cases in which inclusionary MHCA was used. This figure is generally consistent with the findings of Garrett and Neufeld, who found that testimony was “invalid” in 38% of DNA exonerations in which transcripts were located and MHCA was used, and Garrett who found testimony was invalid in 39% of cases. These figures from Garrett and Neufeld, however, include all cases in which MHCA was used at all. If cases are limited to those in which the MHCA testimony “was used to show the guilt of the defendant,” the portion of “invalid” testimony rises to 48%, almost exactly the same as our figure (which is not surprising because all of the cases they analyze are included in our data set).¹⁰⁷

Table 9. Distribution of “appropriate” and “erroneous” evidence, according to criteria used in FBI review, in all exoneration cases in which inclusionary MHCA was used. These include 117 inclusions and one “race only inclusion” (n=118).

Testimony	Count	%
Appropriate	63	53%
Erroneous	55	47%
Grand Total	118	100%

Joseph Sledge, Jr. (Nov. 26, 2014). Therefore, the Sledge case would have been among the 257 cases coded “erroneous” in the 2015 FBI review. Personal communication, Peter Neufeld to Cole (Oct. 13, 2023).

Before we became aware of the FBI, IP, and NACDL’s coding, we independently coded Sledge erroneous for Error Type 2, but for a different statement: “I look at hairs on a day-to-day basis, and I find it *extremely unlikely* when hair samples taken from two different individuals at random cannot be differentiated between. Hairs are quite distinct in their own innate microscopic characteristics” (emphasis added). See <https://convictingtheinnocent.com/exoneree/joseph-sledge/> (<https://perma.cc/F855-3GPQ>). The statement “extremely unlikely” seems to fall squarely within the definition of Error Type 2 which describes, in part, “an opinion as to the likelihood or rareness of the positive association that could lead the jury to believe that valid statistical weight can be assigned to a microscopic hair association.” FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*, 12-13. In this Report, Sledge is coded erroneous for Error Type 2.

The disagreement over the coding of Sledge provides a window into the different perspectives of the FBI and IP/NACDL during the Review. It is unfortunate the full coding decisions of the organizations have not been made public so researchers can better understand the findings of the FBI review. The FBI’s position in the Sledge case might have prompted us to suspect that FBI coders were trying to minimize “erroneous” testimony—were it not for the fact that the FBI review coded 96% of testimony erroneous, which hardly seems like a sign of an effort to minimize.

¹⁰⁷ Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 47; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 95.

In Table 10, the combined results for all non-FBI cases from all data sources are shown and compared with the comparable data from the FBI review.

Table 10. Comparison of appropriate and erroneous evidence in FBI review and data sources used in this study.

Agency	Data Source	Appropriate	Erroneous	Total
Non-FBI	NRE	60 (61%)	39 (39%)	99
	TFSC	33 (57%)	25 (43%)	58
	WL	4 (67%)	2 (33%)	6
	Total	97 (60%)	66 (40%)	163
FBI Review	Testimony	34 (7%)	450 (93%)	484
	Reports	873 (50%)	856 (50%)	1729
	Combined	907 (41%)	1306 (59%)	2213

Considering all these sources suggests yet again that “erroneous” reporting is far rarer outside the FBI than in it. If we compare non-FBI cases to FBI *testimony*, the discrepancy (93% vs. 40%) is startling. The incidence of erroneous testimony is more than double in FBI cases than in non-FBI cases. This is probably the most apt comparison because the vast majority of cases in the non-FBI data sets consist of testimony, rather than written reports (see Table 3). However, even if we add in the much larger number of FBI written reports, the FBI and non-FBI evidence are mirror images: FBI cases are 59% erroneous, whereas non-FBI cases are 60% appropriate.

B. ERROR TYPES

Figure 4 shows the distribution of all error types in all cases studied. Overall, Error Type 2 was the most common error, followed by Error Type 1. This was true across all data sets except TFSC, for which Error Type 1 was most common. However, it was most true of the NRE non-FBI data. This finding is consistent with the Root Cause Analysis finding that Error Type 2 was most common and Error Type 3 the least. As noted above, we do not know the actual distribution of error types in the FBI review.

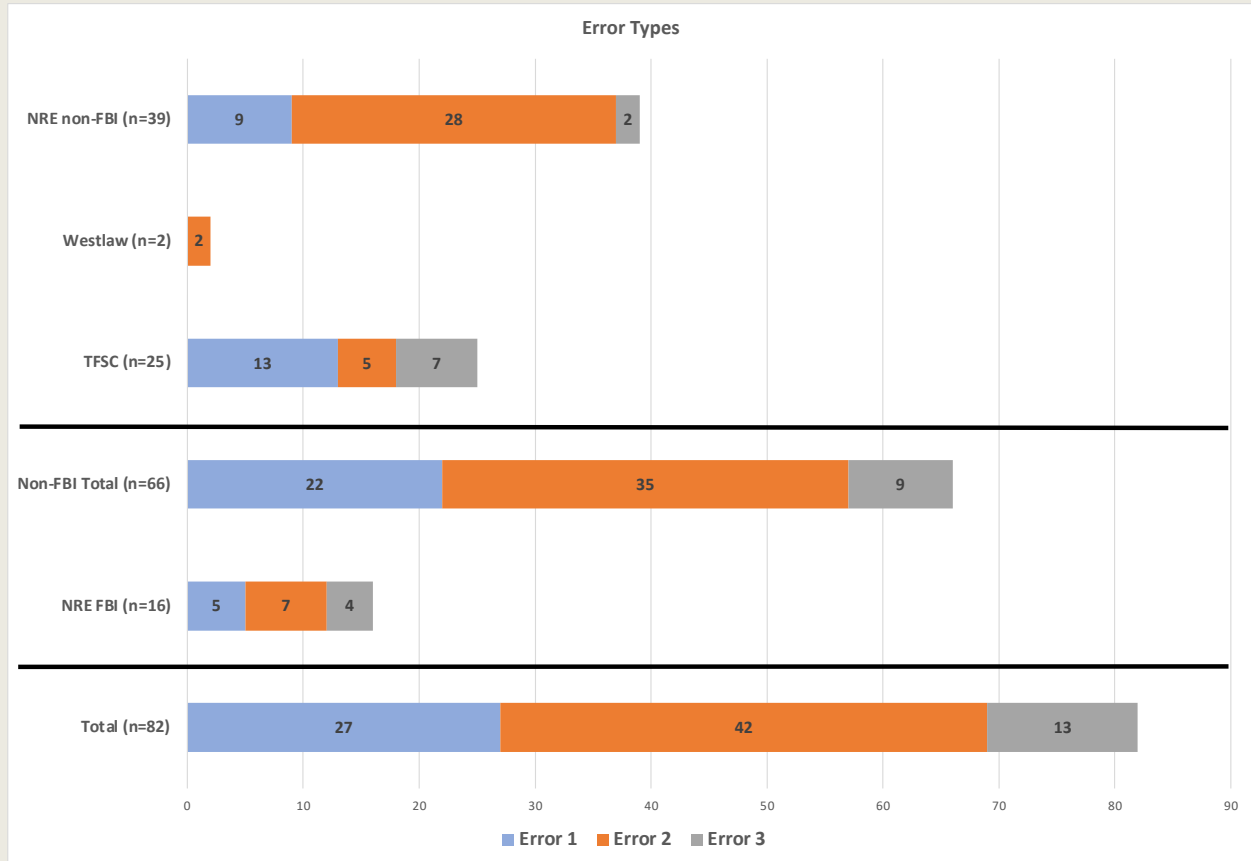


Figure 4. Distribution of error types in cases with erroneous testimony across all data sets (n=82). Error type 1 is individualization, Error Type 2 is an unfounded statistical probability, and Error Type 3 uses experience to vouch for accuracy.

IX. DISCUSSION

A. CASE CHARACTERISTICS

1. NUMBER OF PEOPLE FALSELY CONVICTED ON MHCA EVIDENCE OF ANY KIND

In total, MHCA contributed to the false convictions of 126 people who have been subsequently exonerated since 1989. All but one (125) of those people are described in this Report.¹⁰⁸ The cases are listed in [Appendix 4](#). This represents around 4% of the 3,340 exonerations known on the date of final analysis (July 12, 2023). MHCA ranks among the top four forensic disciplines in contributing to false convictions.¹⁰⁹ This total includes 117 cases in which testimony included the POI, six in which the expert stated that the crime scene hair was inconsistent with the POI but refused to exclude the POI as a source (see Section IX.A.4, one case ([Gary Gauger](#)) in which the defendant was implicated by testimony that damage to the hair was consistent with the defendant’s (false) confession, and one case in which the defendant was implicated only by a racial assignment (see Section IX.A.5) (Table 11).

Table 11. All post-1989 NRE cases in which MHCA contributed to a false conviction by type of testimony given (n=125).

Testimony Type	Count of Evidence Type
inclusion	117
inclusion by failure to exclude	6
race only inclusion	1
damage	1
Grand Total	125

This adds 52 additional cases to the 73 cases listed in the 2015 Innocence Project (IP) report *Not a Strand of Evidence*, a report on all known DNA exonerations to which MHCA contributed.¹¹⁰

¹⁰⁸ One person, [Carlton Lewis](#), was exonerated on August 10, 2023, as this Report was in the final stages of publication. That case is not included in any of the analyses or appendices in this Report. For the record, the testimony in Lewis’s case would have been coded “appropriate.”

¹⁰⁹ This is based on comprehensive Registry study of the contribution of forensic and expert evidence to false convictions that is currently in progress. The highest ranked disciplines are serology, drug analysis, and forensic pathology. This finding is consistent with the findings of John Morgan, *Forensic Testimony Archaeology: Analysis of Exoneration Cases and its Implications for Forensic Science Testimony and Communications*, National Institute of Justice 36 (Mar., 2023), <https://www.ojp.gov/pdffiles1/nij/grants/306259.pdf>.

¹¹⁰ Penchina & Huffman, *Not a Strand of Evidence: DNA Exonerations Involving Unvalidated or Improper Microscopic Hair Comparison Evidence*. *Not a Strand* actually lists 74 cases. However, we believe—and the IP agrees—that one case, [Brandon Moon](#), was listed in error. The hair evidence in Moon was offered by the defense and was an exclusion. These 52 cases were not listed by *Not a Strand* for the following reasons: Nineteen cases were not DNA exonerations. In 16 additional cases, the exoneration occurred after 2015, the publication date of *Not a Strand*. In 9 more cases, DNA was not central to the exoneration and thus not considered a DNA exoneration by the IP. The

This is likely to be an underestimate for several reasons. Despite our efforts, there may yet be exoneration cases in which MHCA contributed to the conviction, but we are not aware of it. In addition, not all false convictions to which MHCA contributed may have resulted in exoneration for a variety of reasons. Some may never have been reviewed. As noted above, state reviews are still ongoing in many states and have not even begun in others. Defendants may have died or been released from prison and not pursued exoneration. Others may have been reviewed but not resulted in exoneration. For example, in some cases, erroneous MHCA evidence contributed to convictions, but courts found this insufficient for relief.¹¹¹

2. YEARS LOST

Collectively, the 125 MHCA exonerees lost a total of 1,918 years in prison. On average, the MHCA exonerees served longer sentences than non-MHCA exonerees convicted of comparable crimes. MHCA exonerees served an average of 15 ½ years, while non-MHCA exonerees for comparable crimes served an average of 12 ¼ years. The MHCA exonerees include [Ledura Watkins](#), who served more than 41 years and was the longest serving exoneree in American history at the time of his release.

3. NUMBER OF PEOPLE FALSELY CONVICTED ON MHCA INCLUSIONS

In total, MHCA inclusions contributed to the false convictions of 117 people who were subsequently exonerated since 1989. Eighteen of those 117 people were convicted in part by evidence from FBI experts. The remaining 99 people were probably convicted by evidence from state or local experts, although we were not able to identify the expert's employer in every case.

4. INCONSISTENT BUT NOT EXCLUDED

In the course of our study, we came across six cases ([Warith Abdal](#), [Richard Alexander](#), [James Waller](#), [David A. Gray](#), [Larry Holdren](#), and [Timothy B. Cole](#)) which we call “inconsistent but not excluded.” These were cases in which the outcome of the MHCA was that hairs were *inconsistent*. However, the hair analyst testified that they were unable to exclude the person of interest (POI) from being the source of the hair. To be sure, there are sometimes legitimate reasons for a MHCA expert to report that they were unable to exclude a POI, such as inadequate

remaining 8 cases—[Bruce Nelson](#), [Dale Mahan](#), [Ronnie Mahan](#), [Larry Holdren](#), [Marcellius Bradford](#), [Richard Alexander](#), [Christopher Ochoa](#), and [Anthony Powell](#)—were DNA exonerations by all standards in which the exoneration occurred before 2015. In these cases, the IP was not aware that MHCA contributed to the conviction at the time of publication of *Not a Strand*.

¹¹¹ A good example is the case of Lonnie Strawhacker. *Strawhacker v. State*, 645 S.W.3d 326, 339 (Ark., 2022) (Wynne, J., dissenting) (“Strawhacker has demonstrated a reasonable probability that, had it been known that Malone’s testimony was erroneous, the outcome of his trial would have been different. The State ‘took great pains’ to introduce the hair-analysis evidence, flying Malone from the FBI lab in Washington, D.C., to give inculpatory expert testimony we now know was riddled with errors. The State argued to the jury that Malone’s testimony was the ‘fatal blow’ to Strawhacker’s case and proved his guilt beyond a reasonable doubt. No other physical evidence was introduced at trial—in fact, fingerprints, blood samples, and saliva samples were not even tested because the State thought the hair analysis was so compelling. The repudiation of Malone’s testimony puts the whole case in such a different light as to undermine confidence in the verdict.”)

samples, very short hairs, or hairs with very little information. However, in the above cases, the MHCA experts transformed the failure to exclude into a conclusion verging on an inclusion. In some cases, they took a position close to that of the extreme philosophical skeptic—that one cannot prove a negative and therefore, can never know anything.¹¹² In the Waller case, for example, the expert testified that the hairs from the scene of the sexual abuse were not consistent with Waller’s sample hairs. But then the analyst said it was impossible to exclude Waller based on hair analysis:

If you wanted to say that this hair did not come from this individual, you would have to check it against every hair to be positive that it did not come from that individual. [One would] practically have to denude a person to make a proper comparison.¹¹³

In the Abdal case, the expert, Michael Krajewski, testified that the hair from the rape scene and Abdal’s hair were distinctively different under microscopic comparison. But then Krajewski testified that Abdal could not be excluded because it was not unusual to observe different-appearing hairs from the same person:

A study shows it would not be unusual to have to look at 4,500 strands of hair from the head in order to get a match with any one particular hair. And, from the pubic hair, one may have to look at as many as 800 hairs.¹¹⁴

For these experts, apparently, MHCA could hurt defendants, but it could never help them. If the hair was “consistent,” they testified that the defendant was included, but if the hair was inconsistent, they refused to exclude them. Such practices suggest the examiners perceived MHCA as a tool to assist only the prosecution.

The statements in these cases would later be contradicted by the FBI’s “limits of science” agreement with the NACDL and IP which stated that exclusion is appropriate using MHCA.¹¹⁵

The FBI review did not count such cases as inclusions. Certainly, it is reasonable to distinguish them from the more typical inclusion in which the expert asserts the *consistency* of two hairs and the likelihood or possibility of their deriving from a common source. But, at the same

¹¹² We do not include the [Calvin Willis](#) case among these 6 cases. In the Willis case, the MHCA expert, Pat Wojtkiewicz, *did* explicitly exclude Willis as the source of hairs found on a bedspread in a child rape case. However, when asked by the prosecutor, “So from that we can deduce he still could have done it but he did not leave any hair there?” Wojtkiewicz answered “That’s correct.” Certainly, we view this as an improper answer for a forensic expert witness to give. However, we do not consider it an example of “inconsistent but not excluded” because Wojtkiewicz *did* exclude Willis as the source of the hair. Instead, in suggesting that the exclusion of Willis as the source of the hair essentially shed no light on the question of whether Willis had been on the bed, Wojtkiewicz applied the same faulty reasoning that underlies “inconsistent but not excluded” evidence to what forensic scientists call the “activity-level proposition,” rather than the “source-level proposition.” R. Cook et al., *A Hierarchy of Propositions: Deciding Which Level to Address in Casework*, 38 *Science & Justice* 231 (1998). The Registry also considers this testimony to be “contextualizing evidence”—“expert testimony designed to provide an explanatory context for negative or missing evidence”—and hence “False or Misleading Forensic Evidence.” William C. Thompson et al., *Evaluating Negative Forensic Evidence: When Do Jurors Treat Absence of Evidence as Evidence of Absence*, 14 *Journal of Empirical Legal Studies* 569, 570 (2017). See also Gwen Jenkins & Regina Schuller, *The Impact of Negative Forensic Evidence on Mock Jurors’ Perceptions of a Trial Drug-Facilitated Sexual Assault*, 31 *Law & Human Behavior* 369 (2007).

¹¹³ Innocence Project, *The Innocence Record*, available at <https://www.innocencerecord.org> (<https://perma.cc/8H75-8URS>).

¹¹⁴ On file with the National Registry of Exonerations and available upon request from id. The Abdal and Waller cases are also discussed in Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 61.

¹¹⁵ FBI, *Microscopic Hair Comparison Analysis: Standard for Review of Testimony and Laboratory Reports*.

time, to call these cases “exclusions” would be to ignore the probative evidence against the defendant that the expert provided.¹¹⁶

These cases show how this “inconsistent but not excluded” testimony might easily contribute to finding an innocent person guilty. This is yet another way, besides inclusions, that MHCA contributed to false convictions. We, therefore, include these 6 exonerees with the 117 against whom inclusions were proffered in the total population of exonerees against whom MHCA contributed to their false conviction.

5. RACIAL ASSIGNMENTS

A racial assignment accompanied a statement including a POI as the source of a hair in at least 7 cases.¹¹⁷ In all 7 cases, the defendant was Black. (As noted [below](#), the defendant was Black in 54% of MHCA exonerations.¹¹⁸) In 6 of those 7 cases, it was a hair that was associated with the defendant that was assigned a race. For example, in the conviction of [William Gregory](#) for sexual assault in 1993 in Kentucky, the MHCA expert testified that the hairs were of “Negroid” origin, they shared “unusual characteristics” with Gregory’s hairs, and they were “more than likely” from Gregory. In the 7th case ([Lee Arthur Hester](#)), a hair was associated with the victim and assigned the race of white.

The use of MHCA for “racial” assignment has been controversial. Historically, the MHCA discipline has endorsed assignment to three “ancestry groups”: “Caucasian, Negroid, and Mongoloid.”¹¹⁹ As recently as 2016, “ancestry group” assignment was described in a draft DOJ document as one of the few things that it *was* appropriate to do with MHCA.¹²⁰ The dermatological community, however, has expressed doubts about the ability to, and propriety of, assigning hair types to “races.”¹²¹ MHCA experts also increasingly view the conventional ancestry groups as “outdated” and have commented that “The ability to differentiate hairs or assign them to a specific ethnic group is questionable, especially in an increasingly multicultural society.”¹²² Cognitive bias concerns have also been raised about a forensic technique that starts by assigning the trace to

¹¹⁶ Paradoxically, courts have noted this problem with “inconsistent but not excluded” evidence when it helps the *defense*. For example, in a case in which footwear analysis failed to exclude an alternate suspect’s shoe as the source of a shoeprint, the California Supreme Court noted that the small size and low quality of the shoeprint “would have made it impossible to exclude any number of shoes as the source of the mark. The criminalist’s inability to exclude [the alternate suspect] as the source of the print was therefore irrelevant and potentially misleading, to the extent the jury might have speculated the result indicated an affirmative identification.” *People v. Turner*, 10 Cal.5th 786, 817 (Cal. 2020).

¹¹⁷ [Johnny Briscoe](#), [Anthony Michael Green](#), [William Gregory](#), [Lee Arthur Hester](#), [Anthony Hicks](#), [Robert Lee Miller, Jr.](#), and [Raymond Towler](#).

¹¹⁸ “MHCA exonerations” are the 125 exoneration cases in which MHCA contributed to the conviction.

¹¹⁹ Robertson & Brooks, *Laboratory Examinations*, 71.

¹²⁰ Department of Justice, *Proposed Uniform Language for Testimony and Reports for the Forensic Hair Examination Discipline*, (2016).

¹²¹ Michael Bigby & Jeffrey D. Bernhard, *Proposed Policy on Identification of Race, Ethnicity or Skin Color in Case Reports and Studies Submitted to the Journal of the American Academy of Dermatology*, 54 *J Am Acad Dermatol* 1077 (2006); Michael Bigby & Diane Thaler, *Describing Patients’ “Race” in Clinical Presentations Should be Abandoned*, id.1074; Jane Unaeze & Michael Bigby, *The Frequency of Reporting of Race/Ethnicity in Case Reports*, id.1067; Nonhlanhla P. Khumalo, *Yes, Let’s Abandon Race—It Does not Accurately Correlate with Hair Form*, 56 *id.*709 (2007).

¹²² Robertson & Brooks, *Laboratory Examinations*, 71; Sandra L. Koch et al., *Microscopical Discrimination of Human Head Hairs Sharing a Mitochondrial Haplogroup*, 66 *J. Forensic Sci.* 56, 67 (2021); Airlie et al., *Forensic Hair Analysis—Worldwide Survey Results*, 6.

crude, contested ethno-racial categories.¹²³ By 2018, the final DOJ document had removed racial assignment from the list of approved statements for MHCA.¹²⁴

Racial assignments were not included in the FBI review.¹²⁵ Nonetheless, racial assignment played an important role in the FBI review through the [Willie Manning](#) case. Willie Manning is the only exoneree currently on death row. Manning had been convicted and sentenced to death in two separate double murders in Mississippi, the 1992 murder of Tiffany Miller and Jon Steckler and the 1993 murder of Alberta Jordan and Emmoline Jimmerson.¹²⁶ During the trial for the Miller-Steckler murder, an FBI hair analyst testified that a hair recovered from their car came from a Black person. Miller and Steckler were white. The analyst did not say anything about Manning individually being the source of the hair. While the FBI review was ongoing, Manning faced an imminent execution date, May 7, 2013. Given Manning's impending execution, the FBI took the unusual step of issuing urgent letters on May 4, 2013, stating that there were errors in the hair evidence in Manning's case.¹²⁷ The letters held that it had been error to report that a Black person was the source of the hair, but, in a curious parsing, still insisted that it was appropriate to report that hair "exhibits traits associated with a particular racial group."¹²⁸

The scientific analysis of hair evidence permits an examiner to offer an opinion that a questioned hair possesses certain traits that are associated with a particular racial group. However, since a statistical probability cannot be determined for classification of hair into a particular racial group, it would be error for an examiner to testify that he can determine that the questioned hairs were from an individual of a particular racial group. Thus, an examiner cannot testify with any statement of probability whether the hair is from a particular racial group, but can testify that a hair exhibits traits associated with a particular racial group.¹²⁹

Four hours before his scheduled execution, Manning was granted a stay of execution in order to allow DNA testing on the hair. In 2015, Manning was exonerated of the Jordan-Jimmerson murder after key witnesses recanted their original testimony and Manning's lawyers discovered that the witnesses had lied about living in the same building as the victims. Manning remains on death row for the Miller-Steckler murder. Manning is an exoneree listed on the National Registry of Exonerations, but he is not included in our study because the MHCA evidence that was used against him was used in the Miller-Steckler prosecution for which he remains convicted.

¹²³ Janis Puracal & Aliza B. Kaplan, *Perpetuating the Presumption of Guilt: The Role of Implicit Racial Bias in Forensic Testimony*, 58 Crim. L. Bull. 317 (2022) <https://forensicjusticeproject.org/wp-content/uploads/2022/07/Perpetuating-the-Presumption-of-Guilt-The-Role-of-Implicit-Racial-Bias-in-Forensic-Testimony.pdf> (<https://perma.cc/4DHD-HT25>).

¹²⁴ Department of Justice, *Approved Uniform Language for Testimony and Reports for the Forensic Hair Discipline*, (Aug. 3, 2018); Department of Justice, *Approved Uniform Language for Testimony and Reports for the Forensic Hair Discipline*, (Mar. 1, 2021), available at https://www.justice.gov/d9/pages/attachments/2021/01/11/finalapproved2020revision_forensichairdisciplineultr_11.25.20.pdf (<https://perma.cc/QC6H-82YQ>).

¹²⁵ Antoun et al., *Hair Microscopy Review Webinar*.

¹²⁶ Andrew Cohen, *Feds Acknowledge Scientific Errors in Testimony in Willie Manning Case*, *The Atlantic*, May 6, (2013).

¹²⁷ *Id.*

¹²⁸ FBI Microscopic Hair Comparison Analysis to Special Counsel (May 4, 2013).

¹²⁹ FBI Laboratory Division *Manning v. Mississippi*, 2013-DR-00491-SCT Microscopic Hair Comparison Analysis to Special Counsel (May 4, 2013).

In addition, even though racial assignment was not included in the FBI review, the review exposed numerous instances of racial assignment by FBI MHCA experts. For example, the experts might say that a “Mexican” person could not be the source of the hair, thus eliminating “Mexicans” as possible perpetrators, which, in turn, narrowed the pool of possible perpetrators around the defendants.¹³⁰

Leaving aside the issue of making racial assignments in the first place,¹³¹ “Mexican” is not even one of the three ancestry groups that the MHCA discipline even *claims* to be able to assign. Mexico is, of course, a nation, not an ancestry group. Nor is “Hispanic” an ancestry group to which the discipline has ever claimed to be able to assign hair. “Hispanic” is considered a cultural-linguistic group that refers to “people who come from Spanish speaking countries,” not a racial group.¹³² MHCA experts have shown particularly poor discrimination when assigning hairs from self-identified Hispanics to the conventionally used ancestry groups.¹³³

Nonetheless, the exoneration cases contained testimony in which hair was described as both “Mexican” and “Hispanic.” For example, during the trial of [Michael Blair](#) for murder in Texas in 1994, the MHCA expert, Charles Linch asserted that “Blair has predominant Mongoloid racial characteristics in his head hair and pubic hair.” Previously, Linch had agreed with the prosecutor that “Hispanics or Mexican-Americans come with this Mongoloid”:

Depending on how much European influence they have. The Hispanics are, by history, Spaniard and American Indian. Now, the predominant influence of American Indian would give them the predominant Mongoloid characteristics.¹³⁴

There was one case in our study in which a racial assignment clearly helped convict the defendant: [Rudolph Holton](#)’s conviction for murder in Florida in 1987. FBI MHCA expert John Quill testified that three hair fragments found in the victim’s mouth had “Negroid characteristics” and therefore could have come from Holton. Although this testimony merely included Holton in one of three groups—“Negroid”—which encompass the entire human population, it nonetheless contributed to the guilty verdict against Holton because there was no independent evidence of the perceived “race” of the perpetrator. We coded this evidence as “appropriate” because it did not fall afoul of the guidance provided in the FBI Laboratory Statement on MHCA and race in the Manning case, discussed above.¹³⁵

We count Holton as the only case in which “race only inclusion” contributed to a wrongful conviction. In the other 9 cases with “race only” testimony, the racial assignment was either

¹³⁰ Antoun et al., Hair Microscopy Review Webinar.

¹³¹ For a critical discussion, see Simon A. Cole, *Individual and Collective Identification in Contemporary Forensics*, 15 *BioSocieties* 350, 359 (2020) <https://doi.org/10.1057/s41292-018-0142-z>.

¹³² <https://www.dictionary.com/browse/hispanic?s=t> (<https://perma.cc/NH4T-4NJ6>); Ann H. Ross, The Concept of ‘Race’: A Forensic Anthropological Perspective on Human Variation Address at Advances in Forensic Anthropology (Aug. 1, 2011); Koch et al., *Microscopical Discrimination of Human Head Hairs Sharing a Mitochondrial Haplogroup*, 67.

¹³³ Koch et al., *Microscopical Discrimination of Human Head Hairs Sharing a Mitochondrial Haplogroup*, 63.

¹³⁴ See transcript at Garrett, Convicting the Innocent: DNA Exonerations Database, available at <https://convictingtheinnocent.com> (<https://perma.cc/FN5F-SKGZ>), 785, 2177. Blair is listed by the National Registry of Exonerations as “White.” The Registry includes “Hispanic” as a possible response to the category “Race/Ethnicity.” The Registry’s designation is not especially meaningful. It may, for example, reflect an attorney’s characterization of how Blair would describe himself. Or it may be an assignment that the Registry inherited from another innocence organization, such as the Innocence Project or the Center on Wrongful Convictions.

¹³⁵ FBI Laboratory Division, Manning v. Mississippi, 2013-DR-00491-SCT Microscopic Hair Comparison Analysis.

exculpatory or not probative of guilt because there was independent evidence of the perpetrator’s perceived “race.”

6. DNA

One hundred and two (82%) of the 125 cases were DNA exonerations.¹³⁶ This is a very high rate of DNA exonerations. DNA exonerations comprise 17% of the 3,340 exoneration cases at the time of analysis (July 12, 2023). The high representation of DNA cases, however, is not surprising. First, the high incidence of MHCA in DNA exonerations has long been known.¹³⁷ Second, DNA exonerations are overwhelmingly murder and rape cases. These crimes involve intimate contact between victim and perpetrator and thus the incidence of hair evidence may well be higher. In addition, the police may be more motivated to search for, and examine, hair evidence because of the seriousness of the crimes. Likewise, police may have turned to hair evidence more often in serious cases to help build the case against the defendant. Finally, it is possible in some cases to perform mtDNA testing on hair. Therefore, cases in which hair evidence was used to convict were more likely to have hair evidence preserved that could later be subjected to mtDNA testing, thus resulting in a DNA exoneration. We did not systematically study in how many cases post-conviction DNA analysis was actually conducted.

7. CRIME TYPES

Only seven crimes, all serious violent crimes, were represented as the top charge in the 125 MHCA exonerations: murder, sexual assault, child sexual abuse, robbery, attempted murder, manslaughter, and kidnapping. MHCA contributed to 6% of the 2,121 total exonerations for those 7 violent crimes. Not surprisingly, the crimes for which the MHCA exonerees were convicted closely mirror the distribution of crimes among DNA exonerations. Compared to non-MHCA exonerations, MHCA exonerations are greatly skewed toward sexual assault—as are DNA exonerations. Sexual assault comprised 42% of MHCA exonerations, whereas only 10% of non-MHCA exonerations are for sexual assault (Table 12). Again, this is not surprising given the utility of hair evidence in both prosecuting and exonerating sexual assaults.

¹³⁶ Here and below we are using the NRE definition of “DNA exoneration”—“cases in which post-conviction DNA testing, analysis, or interpretation was conducted and the results were a factor in establishing the factual basis of the exoneration”—not the narrower IP definition discussed *supra* note 37.

¹³⁷ Edward Connors et al., *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial*, Research Report, National Institute of Justice, (June, 1996); Scheck et al., *Actual Innocence: When Justice Goes Wrong and How to Make It Right*; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*; Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 21.

Table 12. MHCA exonerations by crime type. Data as of July 12, 2023.

Crime	% of MHCA cases (n=125)	% of non-MHCA exoneration cases (n=3,215)	% of non-MHCA DNA exoneration cases (n=478)
Murder	43%	37%	46%
Sexual Assault	42%	10%	38%
Child Sex Abuse	7%	10%	6%
Robbery	2%	5%	3%
Attempted Murder	2%	2%	1%
Manslaughter	2%	2%	1%
Kidnapping	2%	<1%	1%
Grand Total	100%		

8. THE DEATH PENALTY

MHCA occurred frequently in capital cases. Fifteen (12%) of the MHCA exonerees were sentenced to death. This is almost double the proportion of non-MHCA exonerees convicted of the 7 violent crimes included among MHCA exonerations (6%).¹³⁸ It is double the proportion of non-MHCA DNA exonerees sentenced to death (5%).¹³⁹

9. GENDER

The MHCA exonerees are overwhelmingly male. Only one MHCA exoneree, [Paula Gray](#), less than 1% of the total, was female. In contrast, around 9% percent of all exonerees are female.¹⁴⁰ This may follow from the overrepresentation of sexual assault in MHCA exonerations, and, indeed, the one female MHCA exoneree, Gray, was convicted in a sexual assault/murder case.¹⁴¹

10. RACIAL DISTRIBUTION

MHCA exonerees were 54% Black and 41% white. The remaining 7 exonerees (less than 6%) were Hispanic, and there were no Asian, Native American, or “other” race exonerees. Although the proportion of Black exonerees implicated by MHCA (54%) is high compared to the U.S. population, it is reflective of the generally high proportion of Black exonerees for the 7 violent crimes found in MHCA and DNA exonerees (Table 13) and for the Registry as whole, which is 53% Black.¹⁴²

¹³⁸ 121 of 2,121 non-MHCA exonerees convicted of the 7 violent crimes included among MHCA exonerations on the date of final analysis (July 12, 2023) were sentenced to death.

¹³⁹ 24 of 478 non-MHCA DNA exonerations on the date of final analysis (July 12, 2023).

¹⁴⁰ 285/3,340 on the date of final analysis (July 12, 2023).

¹⁴¹ 356/357 (99.7%) of exonerees who were convicted of sexual assault as the top charge are male (as of April 7, 2023), which does not necessarily reflect the gender balance for either prevalence or the commission of that crime.

¹⁴² 1,764/3,340 on the date of final analysis (July 12, 2023).

Table 13. Percent of Black Exonerees in MHCA and Non-MHCA Exonerations, by Crime, analysis as of July 12, 2023.

	Sexual Assault	Murder	Child Sex Abuse	Other Violent Crimes ¹⁴⁸	ALL 7 VIOLENT CRIMES ¹⁴⁴
MHCA Exonerations (125)	60%	48%	44%	60%	54%
Non-MHCA Exonerations for 7 violent crimes (2,121)	60%	56%	27%	58%	52%
<i>Non-MHCA DNA Exonerations for 7 violent crimes (461)</i>	65%	49%	77%	74%	59%

11. GUILTY PLEAS

Only 6 (5%) of the 125 exonerees pled guilty.¹⁴⁵ This is less than a quarter of the overall rate of around 24% guilty pleas in non-MHCA exoneration cases, again probably because MHCA cases are skewed toward more serious crimes.¹⁴⁶

12. TEMPORAL TRENDS

Figure 5 shows the year of conviction for all MHCA exonerations. These are charted against the year of conviction for non-MHCA exonerations for the 7 violent crimes responsible for all MHCA exonerations. Not surprisingly, MHCA false convictions were primarily a product of the late 1970s and 1980s, which is considered the “heyday” of MHCA.¹⁴⁷ MHCA false convictions peaked in 1986, a decade earlier than false convictions in general peaked. And, in 1978 nearly as many (6) exonerees were falsely convicted in part based on MHCA as were falsely convicted of the 7 violent crimes without it (8), a remarkable statistic given that overall only 6% of all convictions for those 7 crimes were based in part on MHCA. The sharp decrease in MHCA exonerations in the 21st century probably reflects the increasing availability of pre-trial DNA profiling and the decreasing use of MHCA alone without, for example, confirmatory use of mtDNA.

¹⁴³ Kidnapping, robbery, attempted murder, & manslaughter.

¹⁴⁴ This refers to the 7 crimes in which MHCA is implicated: Murder, sexual assault, child sex abuse, kidnapping, robbery, attempted murder, & manslaughter.

¹⁴⁵ [Marcellius Bradford](#), [John Huffington](#), [Steven Lopez](#), [Kevin Martin](#), [Christopher Ochoa](#), and [David Vasquez](#).

¹⁴⁶ 808/3,340 cases, as of July 12, 2023. Note that this comparison, unlike most in this report, compares MHCA exonerations to *all* exonerations, not to non-MHCA exonerations. See <https://perma.cc/8SD7-U9FX>.

¹⁴⁷ Morgan, *Wrongful Convictions and Forensic Science Errors: Case Studies and Root Causes*, 55.

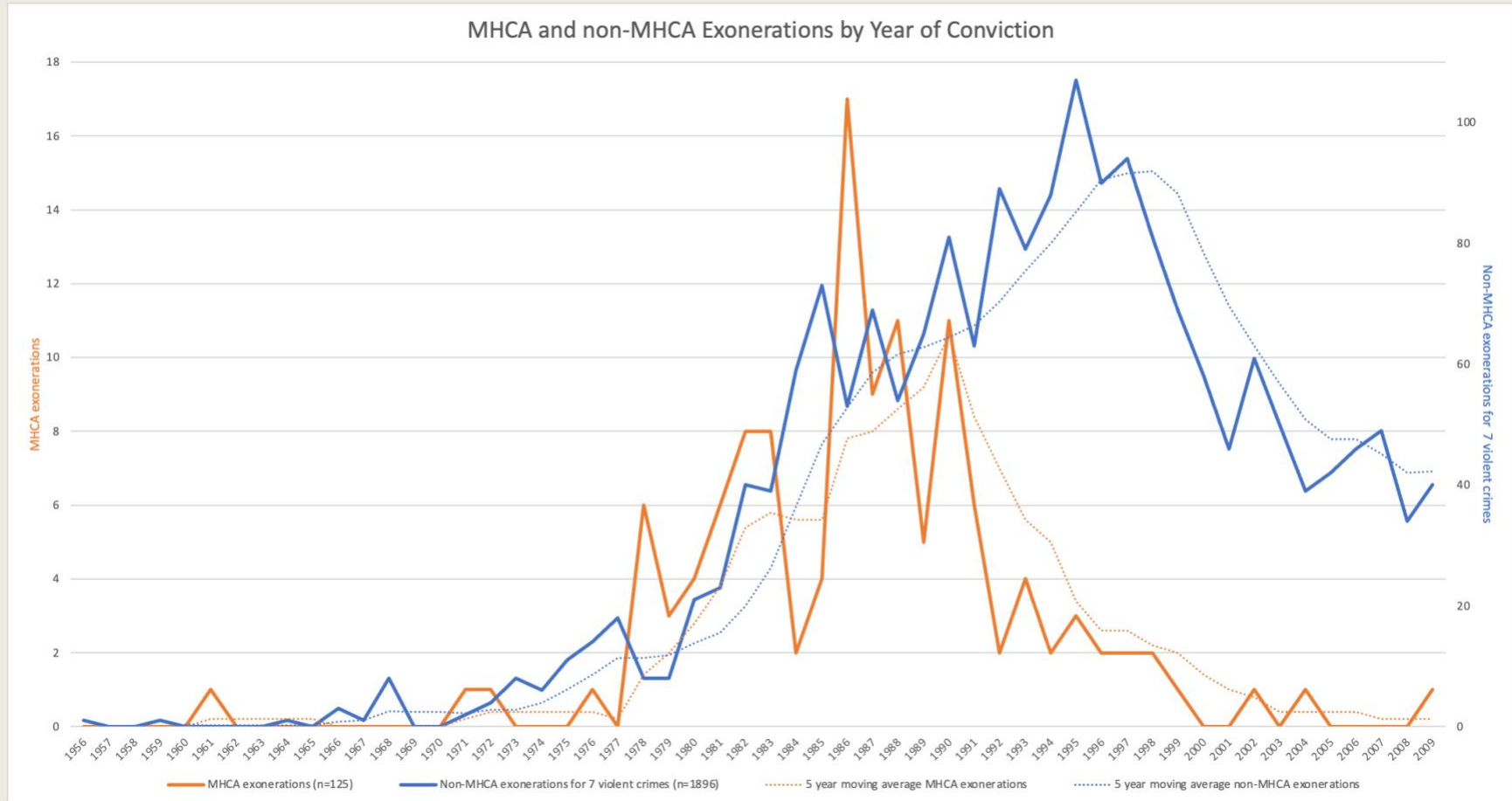


Figure 5. MHCA exonerations by year of conviction. Data is shown only through 2009, the latest year of conviction for a MHCA exonerations at the time of publication. The lag between conviction and exonerations in all exonerations cases is approximately 9 years. The full data table for this chart is available in Table 18, Appendix 3.

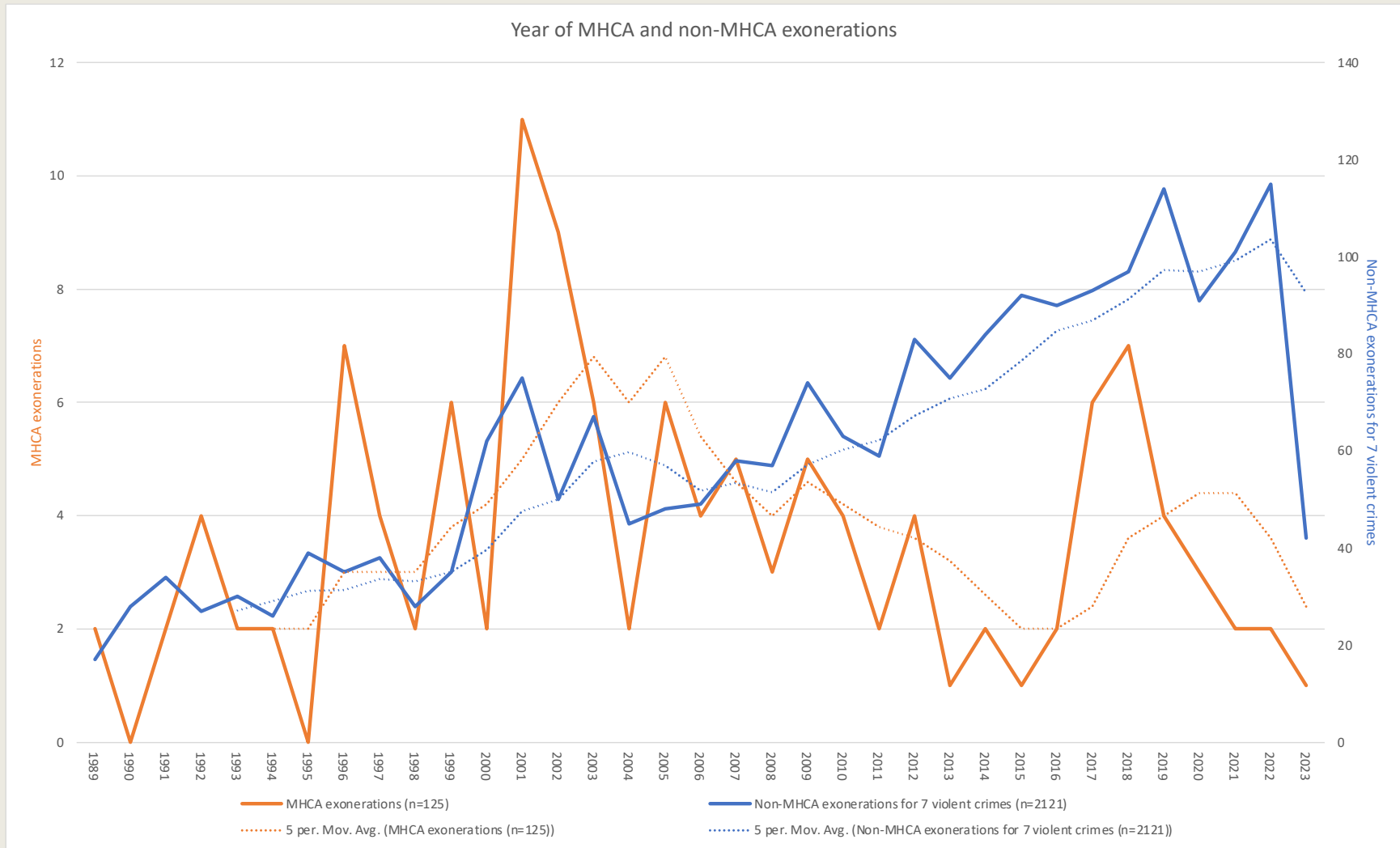


Figure 6. MHCA exonerations by year of exoneration as of July 12, 2023. The data table for this chart is available in Table 19, Appendix 3.

Figure 6 shows the MHCA exonerations by year of exoneration plotted against the same set of non-MHCA exonerations for 7 violent crimes. MHCA exonerations peaked in 2002, but then began to drop off. As noted above, hope was expressed that the 2015 FBI review would stimulate exonerations.¹⁴⁸ Figure 6 suggests that the Review did stimulate renewed focus by exonerators on cases involving MHCA. It also suggests that the burst of attention prompted by the FBI review may be waning as MHCA exonerations began to decrease after 2018. However, exoneration is a slow process—averaging almost 9 years from date of conviction¹⁴⁹—and we should not expect to have seen all the fruits of the statewide audits yet. In total, 27 people convicted based at least in part on MHCA were exonerated after the FBI review in April 2015.

¹⁴⁸ FBI, FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review.

¹⁴⁹ National Registry of Exonerations, Longest Incarcerations (2022), available at <https://www.law.umich.edu/special/exoneration/Pages/longestincarceration.aspx> (<https://n2t.net/ark:/88112/x2x891>).

13. GEOGRAPHY

Figure 7 shows the distribution of MHCA exoneration across different states.

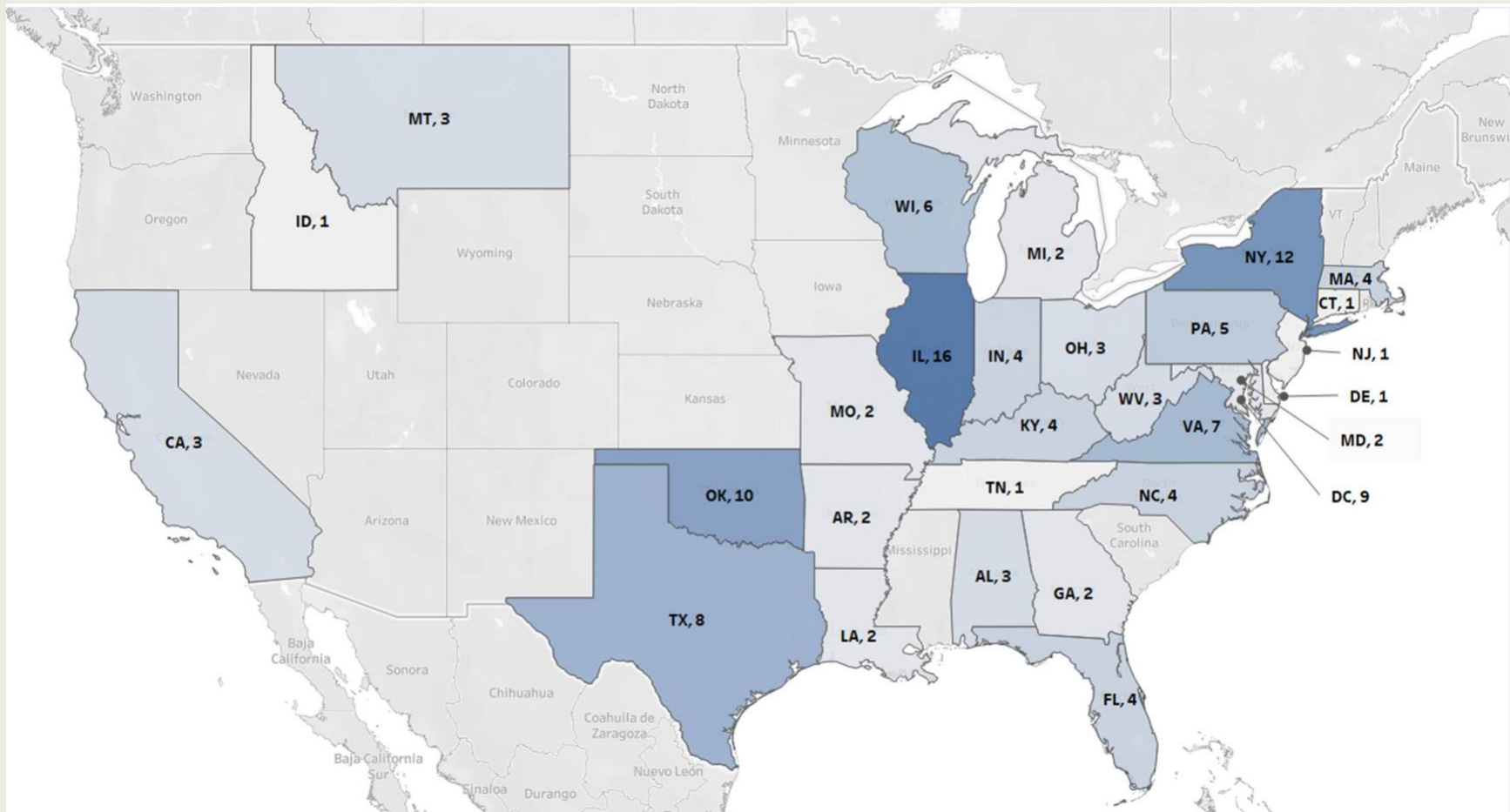


Figure 7. Geographic distribution of MHCA exoneration. One federal case is counted for the state (California) in which it occurred.

Geographic clusters are unlikely to be particularly meaningful because they may reflect differences in law or in available post-conviction resources, but MHCA exoneration exhibit some geographic clustering in particular states, such as Illinois and New York.

The portion of MHCA exonerations in Illinois, however, is consistent with the portion of all exonerations in Illinois, and New York's portion of MHCA exonerations in is actually lower than its portion of all exonerations.

However, some other states' numbers of MHCA exonerations are small, but disproportionately large compared to their share of exonerations in general. Most dramatically, the District of Columbia's (D.C.) 9 MHCA exonerations comprise 7% of all MHCA exonerations, but D.C. has less than 1% of all exonerations. D.C.'s proportion of MHCA exonerations is 10 times greater than its proportion of all exonerations. This is not surprising given the common use of the FBI Laboratory in D.C. criminal prosecutions and D.C.'s role in exposing the problems with MHCA. Another outlier is Oklahoma, which had 8% of all MHCA exonerations but only 1.2% of all exonerations, 2.5% of all DNA exonerations, and 5.2% of all death penalty exonerations.¹⁵⁰ Oklahoma's share of MHCA exonerations is more than 6 times its share of all exonerations. However, Oklahoma was home to a notorious problem forensic scientist, Joyce Gilchrist, who testified often about MHCA. Still, the problem in Oklahoma was not confined to Gilchrist: she testified in only 4 of the 10 Oklahoma MHCA exonerations.

Other states with a disproportionate number of MHCA exonerations are home to other problematic forensic analysts: West Virginia, with a factor of more than 5, was home to disgraced forensic analyst Fred Zain. Montana, also with a factor of more than 5, had a state crime laboratory director, Arnold Melnikoff, whose erroneous testimony prompted the increased scrutiny of MHCA. Two small states, Delaware and Idaho, have a disproportionate number of MHCA exonerations based on single cases. Other states with more than one MHCA exoneration that have a disproportionate share of MHCA exonerations are, in order of the disproportion: Kentucky, Arkansas, Virginia, Alabama, Wisconsin, Indiana, North Carolina, Florida, Massachusetts, Pennsylvania, Georgia, and Maryland.

¹⁵⁰ Thus, Oklahoma contributes to death row exonerations at 4 times the rate it contributes to all exonerations in state cases. By contrast, for Texas, which widely considered the capital of the death penalty in the U.S., the ratio is inverted: Texas contributes to a greater share of all exonerations (13%) than it does to death row exonerations (8%). However, there are 5 states for which the ratio of their contribution to death row exonerations to all exonerations is higher than Oklahoma's: Alabama, which contributes to death row exonerations at 5 times the rate it contributes to all exonerations, Mississippi, with a factor of 5.6, Idaho with a factor of 7.5, Delaware with a factor of 8, and Arizona with a factor of 8.5. On the death penalty in Oklahoma, see Daniel LaChance, *Executing Freedom: The Cultural Life of Capital Punishment in the United States*, 129-154 (2016).

14. OTHER CONTRIBUTING FACTORS

While MHCA contributed to the convictions in 125 cases, this does not tell us *how much* it contributed. The MHCA might have been crucial, or it might have merely bolstered other false evidence, such as other forensic evidence, eyewitness identifications, or false confessions.

We do not have any easy way of measuring the contribution MHCA made to these convictions, and we do not attempt to do so. While we have some (but usually not complete) information about other evidence against the defendants, we, of course, never know what led a jury to decide to convict (or, in some cases, a defendant to decide to plead guilty).

One thing we can do is look at the incidence of other “contributing factors” recorded by the NRE. The occurrence of mistaken witness identification and false confessions was higher in MHCA exonerations than in non-MHCA exonerations for the 7 violent crimes that occurred in MHCA cases and slightly lower for perjury or false accusation (Table 14). In general, MHCA exonerations, not surprisingly, look more like DNA exonerations than like non-DNA exonerations. The exception is for Official Misconduct, for which the incidence in MHCA is higher than for either comparable DNA exonerations or non-DNA exonerations.

Table 14. Contributing factors in MHCA and non-MHCA Exoneration.¹⁵¹

CONTRIBUTING FACTOR	MHCA Exonerations (n=125)	Non-MHCA Exonerations for 7 violent crimes (n=2,121)	<i>Non-MHCA DNA Exonerations for 7 violent crimes (n=461)</i>
Mistaken Witness Identification	50%	37%	<i>57%</i>
False Confession	21%	16%	<i>23%</i>
Perjury or False Accusation	59%	66%	<i>50%</i>
Official Misconduct	64%	60%	<i>53%</i>
Inadequate Legal Defense	15%	33%	<i>20%</i>

Another approach is to ask what, if any, other evidence placed the defendant at the crime scene (or, in some cases, placed the victim at a location associated with the defendant, like their home or car). Of the 125 cases in which MHCA contributed to the conviction, in 85 (68%) of them there was some kind of a false identification.¹⁵² In 26 (21%) of the 125 cases, the exoneree falsely confessed to the crime. (In 14 of those 26 cases there was both a false identification and a false confession; in the remaining 12 cases, there was a false confession alone.) Of the remaining 27 cases, in 17 (14% of the 125 MHCA cases), some forensic or physical evidence, in addition to MHCA, linked the defendant to the crime scene (or the victim to the defendant) (Figure 8).

¹⁵¹ False or Misleading Forensic Evidence is not listed because all MHCA exonerations in these data included False or Misleading Forensic Evidence as a factor that contributed to the defendants’ wrongful convictions.

¹⁵² These cases include the 62 cases with a Mistaken Witness Identification plus 23 in which there was an Other False Identification. These Other False Identifications are cases in which there was a false identification that was not a mistake, but rather a lie. For NRE coding purposes, these cases would be coded Perjury or False Accusation, rather than Mistaken Witness Identification.

Thus, it appears that MHCA rarely convicted alone, but was often used in combination with other evidence. MHCA commonly accompanied eyewitness identification and sometimes accompanied a confession or other forensic or physical evidence. This makes sense. If MHCA is as indiscriminating as the FBI now says it is—including an individual as a possible contributor only in a “pool of unknown size”—then MHCA evidence would be expected to frequently *fail to exclude* innocent people who were suspected by the police for some other reason. If many people can be found “consistent” with any particular hair, then MHCA will often fail to exclude people who are not actually the source of that hair. We expect that MHCA evidence will be found “consistent with” both the true source of the hair and also a “pool of non-contributors of unknown size.” Therefore, when the police suspect an innocent person—because, for example, of a mistaken eyewitness identification or a false confession—we should expect MHCA to find the hair “consistent” with hair fairly often. In short, it may be MHCA’s lack of *discriminating power* that led it to corroborate, rather than disconfirm, so many accusations that later proved to be false.

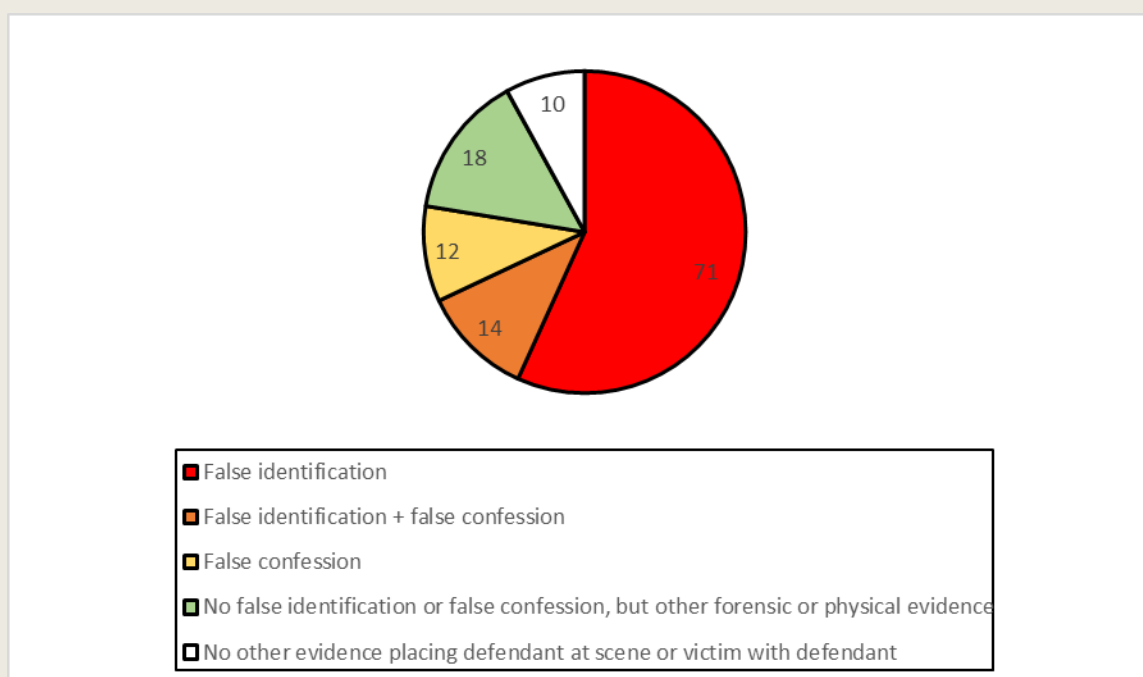


Figure 8. Other evidence placing defendant at the scene or victim at location associated with defendant (n=125)

This leaves 10 (8% of the 125 MHCA exonerations) cases in which no evidence placed the defendant at the crime scene (or placed the victim at a location incriminating to the defendant) other than MHCA (Figure 8).¹⁵³ One such case was the conviction of [Gary Nelson](#) for capital murder and rape of a 6-year-old girl in Georgia in 1980. Nelson was suspected because he had had previous encounters with the police and was the roommate of the man whom the victim was last seen going to visit. At trial, an MHCA expert testified that an arm hair found at the crime scene

¹⁵³ [Timothy Bridges](#), [Jeffrey Clark](#), [Gary Dotson](#), [Donald Gates](#), [Michael Lee McCormick](#), [Thomas Murphy](#), [Gary Nelson](#), [Glenn Payne](#), [Larry Peterson](#), and [Calvin Lee Scott](#). McCormick “confessed” to the crime, but he did so to an undercover police officer posing as a criminal. The Registry does not consider that a false confession. In Murphy and Payne’s cases, there was other forensic evidence—sexual assault examinations—but that evidence spoke to whether there were signs of sexual assault, not placing the defendant at the crime scene. [George Perrot](#) is not counted among the 10, but his false confession was to breaking in to the home, not to the sexual assault for which the MHCA evidence used to build a case.

could have only come from Nelson or one of 120 people in the local area—a baseless statistic. Arm hair analysis itself has long been controversial; as early as the 1985 FBI Symposium, experts argued that arm hairs were of “no value for . . . association to a particular person.”¹⁵⁴ The only other evidence placing Nelson at the scene was the testimony of a police officer who said a witness had identified Nelson in a lineup; that witness then testified that Nelson was not the man she had seen. Nelson was convicted and sentenced to death. The conviction was reversed in 1991 in part because Nelson’s post-conviction attorneys found an FBI report stating that the arm hair which supposedly matched Nelson was “not suitable for significant comparison purposes.”

[Donald Gates](#) was convicted of murder in the District of Columbia in 1982. The only substantial evidence against him was the claim of a paid police informant that Gates had confessed to him and the testimony of FBI MHCA expert Michael Malone, who said that Gates’s hairs were “microscopically indistinguishable” from hairs found on the victim’s body. Gates’s exoneration through DNA testing in 2009 was one of those that prompted the FBI review.

[Timothy Bridges](#) was convicted of sexual assault of an 83-year-old woman in North Carolina in 1991. As with Gates, the only evidence against Bridges were the statements of informants and MHCA evidence offered by an expert who testified that he could make a “strong identification” that the hair at the scene was Bridges’s hair and that there was only a 1 in 1,000 chance that two Caucasian people would have indistinguishable head hair. Bridges was exonerated in 2016.

Perhaps the greatest contribution MHCA played in any one conviction was at the trial of [Calvin Lee Scott](#) in 1983 in Oklahoma for the rape of a 25-year-old woman with her 4-month-old baby next to her. This is one of only two cases with no contributing factor other than the misleading MHCA evidence. Suspicion was cast on Scott by an anonymous phone call to the police. At Scott’s trial, which appears to have been completed in a single morning before lunch, the anonymous call was cursorily mentioned to the jury by the investigating police officer. The survivor testified, but could only describe the race and height of her attacker. The only other evidence implicating Scott was MHCA. The hair analyst, Claud Berry, said hairs found on the sheets and the survivor’s body were “microscopically consistent” with the samples of Scott’s hair. He further gave statistics about the probability of hair from two different people being found microscopically consistent: 1 in 4,500 for head hair and 1 in 800 for pubic hair. As discussed above (Section IV), these statistics were not scientifically supported.¹⁵⁵ Scott was, in essence, convicted and sentenced to 25 years in prison on MHCA alone. Twenty years later, he was exonerated by DNA testing.

Because jury deliberations are secret, we have only glimpses of the effect of MHCA evidence on juries. However, there are anecdotes that attest to the notion that MHCA had a significant impact on juries in some cases. For example, in 2001, when [Michael Blair](#) was appealing his 1994 murder conviction in Texas, which would eventually be overturned, the *Dallas Morning News* interviewed five people who sat on the jury that convicted him. All five “said testimony about hair microscopically similar to Mr. Blair’s and found on” the victim’s “body was most meaningful to them. They said that hair was the most compelling proof against him, although they noted that they considered other factors, such as Mr. Blair’s decision to search for the child, in determining his guilt.”¹⁵⁶

¹⁵⁴ Proceedings of the International Symposium on Forensic Hair Comparisons 203.

¹⁵⁵ Also see Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 57.

¹⁵⁶ Holly Becka, *DNA tests don't link Blair, girl: DA says other evidence ties man to Ashley Estell*, *The Dallas Morning News*, 2001/03/03, 2001 at 1A.

15. ROLE OF THE FBI

Despite the well-documented problems at the FBI, its testimony contributed to fewer than expected known false convictions. FBI testimony contributed to only 15% of known exonerations involving MHCA, but it has been estimated that the FBI performed somewhere between 20-50% of MHCA nationwide.¹⁵⁷ This suggests, surprisingly, that state and local experts contributed a greater than expected share of MHCA exonerations. This is all the more surprising because, as discussed both [above](#) (Section VIII.A) and [below](#) (Section IX.B), FBI MHCA evidence contained far more erroneous statements than state and local MHCA evidence. It is further evidence that the erroneous nature of the testimony is not a crucial factor in MHCA evidence producing false convictions.

16. EXONERATIONS BEFORE 1989

All the above discussion concerns exonerations since 1989. The Registry also maintains a separate registry of [Exonerations Before 1989](#), containing 456 cases at the time of analysis (July 12, 2023). MHCA contributed to three false convictions that resulted in exoneration before 1989: [Ronald Carden](#), [James Hall](#), and [Anthony Ray Peek](#) ([Appendix 5](#)). There were also two cases in which hair evidence excluded the defendant ([Craig Bell](#) and [Juan Ramos](#)).

The testimony in Carden's case was an example of FBI Error Type 1, Peek was a classic case of Error Type 2, and the testimony in Hall's case was "appropriate" by FBI standards. With these three cases included, the total number of exonerations in which MHCA contributed to the false conviction is 128.

17. EXCLUSIONS

MHCA did not always incriminate innocent defendants, although it did not prevent their false convictions. We found 49 cases in which MHCA evidence excluded a defendant who was convicted and later exonerated.

For example, the 1989 conviction of [Kenneth Ireland](#) for murder and sexual assault in Connecticut was based largely on informant statements. The MHCA expert testified that hairs from the scene were "dissimilar" to Ireland's hair. Although deliberations began with the jury split 6-6, after 3 days of deliberations Ireland was convicted and sentenced to 50 years in prison before being exonerated by DNA testing in 2009.

Also see the [Brandon Moon](#) case discussed [below](#).

In short, the number of exclusions (49) in cases in which people who were almost certainly innocent¹⁵⁸ were convicted was less than half the number of inclusions (117). To be sure, the

¹⁵⁷ Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*.

¹⁵⁸ The Registry's definition of exoneration is designed to be a conservative proxy for actual innocence by minimizing, though not eliminating, guilty defendant classification errors (guilty defendants misclassified as innocent) at the cost of a great many innocent defendant classification errors (innocent defendants misclassified as unexonerated). For this reason, we characterize the exonerates in the Registry data set as "almost certainly innocent." We are confident that the number of guilty defendants in the Registry data set is small and that we cannot identify who they are. For further discussion, see Samuel R. Gross & Michael Shaffer, *Exonerations in the United States, 1989-2012*, National Registry of Exonerations, 11-14 (May, 2012),

Registry is more likely to become aware of MHCA evidence if it resulted in inclusion than if it resulted in exclusion. Nonetheless, the discrepancy is disconcerting. If we limit our discussion to just MHCA examinations resulting in inclusion or exclusion, then we can say the following: In 166 cases, MHCA experts were asked to perform hair comparisons, generally between a hair sample from an innocent defendant and a hair in an incriminating location (like the victim’s person, home, or clothing). (In a small number of comparisons, the comparison was between a hair sample from a victim and a hair in an incriminating location, like the defendant’s home, vehicle, clothing, or person.) In those 166 cases, the MHCA expert excluded the hair in only 49 cases (30%). But in 117 cases (70%) the MHCA expert reported that the hairs could be from the same source. Thus, in cases that we now know involved innocence defendants, MHCA failed to exclude in 7 out of every 10 cases. This finding alone casts doubt on the utility of MHCA in criminal investigations.

B. TESTIMONY

1. ERRONEOUS TESTIMONY

Not surprisingly, the same kind of erroneous testimony identified in the FBI review was common in state and local cases. Just like FBI experts, state and local experts made outrageous erroneous statements. They implied that defendants were the source of hairs, when they should only have said they might be sources of hairs, testified about baseless statistics, and misleadingly implied that their experience examining hair was a measure of the probability of a false association.

Figure 4 (above) gives the distribution of error types across all data sets. As noted above, the lack of transparency about the FBI review makes it impossible to do any more than a rough comparison of distribution of error types between the FBI and state and local cases. As noted above, we can discern, however, that Error Type 2 was generally the most common, with the exception of the TFSC cases, for which Error Type 1 was most common. We do not have an explanation for this difference.

Although, as explained in [Appendix 2](#), the boundaries between the three error types sometimes seem arbitrary, the error types are valuable for describing inductively the general patterns of testimony used by MHCA experts. It is, therefore, worth revisiting the error types and using exoneration cases to illustrate how each functioned to turn near-valueless evidence into persuasive evidence that was apparently given non-negligible weight by the fact-finder.¹⁵⁹

a) Error Type 1

Error type 1 describes classic “source attribution” testimony in which a forensic scientist states that a particular individual, body part, or object is *the* source of some trace evidence. The implication, which may be expressed with varying degrees of explicitness, is that that individual, body part, or object is the only possible source of that trace. Such testimony—in any forensic discipline, not just MHCA—has been widely criticized. For statisticians and the statistically-minded, such statements are fundamentally inconsistent with probability theory. “Cromwell’s rule” states

http://www.law.umich.edu/special/exoneration/Documents/exonerations_us_1989_2012_full_report.pdf
(<https://n2t.net/ark:/88112/x2sk6r>).

¹⁵⁹ Helpful examples of erroneous testimony in DNA exoneration cases may also be found in Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 53; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 97-100.

that, given two hypotheses, some probability, no matter how small, must be assigned to each hypothesis.¹⁶⁰ It is, therefore, fundamentally wrong to state or imply, as source attribution statements do, that one hypothesis (e.g., that the POI is *not* the source of the trace) has been ruled out or to state or imply that one hypothesis has been “proven.”

Put another way, for a forensic analyst to make a source attribution is to ignore the other evidence in the case. It is to suggest that there could be no other evidence powerful enough to convince the analyst that their conclusion is wrong, which is plainly absurd. One MHCA expert, Elmer Gist, learned this the hard way when they testified in the sexual assault prosecution of [Edward Honaker](#) in Virginia in 1985 about a hair found on the survivor’s shorts. The expert testified that “it is unlikely that the hair would match anyone other than the defendant; but it is possible.” After Honaker was excluded as the source of a semen stain by DNA testing, he was exonerated in 1994.

When confronted with the fact that Honaker had undergone a vasectomy, the state's forensic expert said that he would not have testified to the definitive hair match that helped convict Honaker at trial.

From a more practical perspective, it is well understood that the microscopic “consistencies” observed in MHCA are present in many different individuals. It would, therefore, be unreasonable to conclude from observations of consistencies that one particular individual is the source of the hair. In contrast to some other forensic disciplines, such as fingerprints and firearms and toolmarks, in which the disciplines claimed until recently that they could narrow the potential donor pool of a trace to a single individual, the MHCA discipline (as noted [above](#)) acknowledged long ago that the reduction of the potential donor pool to a single individual was not possible. Instead, as noted above, what was possible was to state that the POI was included in a donor pool of unknown size. Source attribution testimony, of course, goes far beyond that, improperly exaggerating the probative value of the evidence.

Nonetheless, as the FBI review showed, FBI examiners continued to give such testimony well into the 1990s and may have trained state and local examiners to do so as well. For example, in the 1980 trial of [Santae Tribble](#) for murder in the District of Columbia, an FBI MHCA expert testified that they had compared the hair from a stocking found on the sidewalk a block away from the murder scene with a head hair from Tribble. They said the hairs were identical and that the hair in the stocking came from Tribble.

Our study shows that some state and local examiners similarly exaggerated the probative value of the evidence. For example, in the 1980 trial of [John Jerome White](#) for sexual assault in Georgia, the state’s MHCA expert testified that hairs from a bedsheet at the crime scene were “similar enough” to the sample of hairs taken from White, “to say they have the same origin.” Thus, the expert went beyond merely saying that White was included in the pool (of unknown size) of potential donors to say that the hairs “had the same origin.”

b) Error type 2

Error type 2 concerns testimony that attempts to solve the problem of the near-valuelessness of MHCA evidence by assigning a probability to the association between the questioned hairs and the known hairs. Such testimony is inherently problematic because it has

¹⁶⁰ Dennis V. Lindley, *Understanding Uncertainty* (2006).

been well understood within the field since at least the 1980s that there are no defensible probabilities to assign to an MHCA association. Although Gaudette’s studies from the 1970s did produce some numbers that some MHCA witnesses have used in their testimony—notably a 1 in 40,000 figure—these are not properly understood as the probability that an individual associated with a hair is *not* the source of that hair. Rather, 1 in 40,000 was simply the portion of hairs from different sources which could not be distinguished in Gaudette’s study—whose methodology, as noted [above](#), has been criticized.¹⁶¹ Moreover, Gaudette himself “never suggested that these numbers be used in isolation of a range of other factors that needed to be considered, which would tend to weaken or strengthen hair conclusions.”¹⁶² As the TFSC notes, Gaudette’s “work in the 1970’s was often misunderstood and inappropriately extrapolated to provide probabilities regarding individuals and races in hair microscopy comparisons.”¹⁶³

Perhaps the most notorious instances of Error Type 2 were committed by MHCA expert Arnold Melnikoff, of the Montana State Crime Laboratory. In 3 exonerated cases ([Jimmy Ray Bromgard](#), [Chester Bauer](#), and [Paul Kordonowy](#)), Melnikoff simply conjured up a probability (1 in 100) of the defendant *not* being the source of a questioned head hair. To compound this error, Melnikoff then assigned the same figure to the probability of a coincidental “match” of pubic hairs and multiplied them together to produce a combined probability of 1 in 10,000. Even if the probabilities were not simply made up, head hair characteristics and pubic hair characteristics are not statistically independent and, therefore, cannot simply be multiplied.¹⁶⁴

Melnikoff’s testimony was egregious and scandalous, but the FBI review found numerous instances of FBI MHCA experts assigning probabilities to MHCA association with varying degrees of explicitness. For example, the FBI review reported an unnamed case in which the MHCA expert testified about questioned and known hairs:

Well, they—they could have come from [defendant]. I would say that *it’s only a very remote possibility* that they could have come from some other individual than [defendant].¹⁶⁵

This testimony does not qualify as Error Type 1 because the witness has acknowledged that there is *some* probability that someone other than the defendant is the source of the hair. The witness has not assigned a number to that probability. However, this testimony qualifies as Error Type 2 because the witness has assigned a *verbal* probability that someone other than the defendant is the source of the hair: “very remote.” Error Types 2 include verbal and implied probabilities as well as numerical and explicit ones.

In our study, we saw both of these types. For example, in the 1986 trial of [Anthony Johnson](#) for murder in Louisiana, the MHCA experts testified that there was a “90 percent chance that a hair found in a shower cap at the crime scene belonged to Johnson.” It is unclear where this number came from, and there is no support for it in the MHCA discipline.

¹⁶¹ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 243; PCAST, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*, 118; Rowe, *The Current Status of Microscopical Hair Comparisons*, 873; Stafford Smith & Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil*; Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*.

¹⁶² Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 243.

¹⁶³ TFSC, *Guidelines for Review of Transcripts Texas Hair Microscopy Case Review*, 6.

¹⁶⁴ Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 97-98; Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 53.

¹⁶⁵ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*, 14., FBI’s emphasis

In contrast, in the 1993 trial of [William Gregory](#) for murder in Kentucky, the MHCA expert testified that hairs recovered from a stocking cap left by the assailant were of Negroid origin, shared "unusual characteristics" with Gregory's hairs, and were "more than likely" from Gregory.¹⁶⁶ "More than likely" is not a number, but it is clearly a verbal probability.

c) Error Type 3

Error Type 3 is perhaps the strangest of the three. Like Error Type 2, the expert purports to assign probability to an MHCA association. However, for Error Type 3 the probability is "experience-based." What this means is that the expert purports to generate a fraction by placing in the denominator the number of hair examinations they have performed in their entire career and placing in the numerator the number of hairs from different persons which they were "unable to distinguish." The problem, of course, is that in casework the expert does not know which hairs actually come from different persons and which hairs actually come from the same person. Such testimony has been criticized not only because it generates a sort of fictional probability but also because it uses the (imaginary, undocumented, unverifiable) quantification of the expert's experience (whether measured in years or comparisons) to bolster the testimony.

The FBI review found many errors of this type. Although the case was *not* included in the review, a good example may be found in the 1990 federal trial of [Juan Matta Ballesteros](#) for kidnapping. FBI MHCA expert Michael Malone told the jury that in 10,000 evaluations of whether different hair samples matched, he had only had "two occasions out of those 10,000 where I've had hairs from two different people that I could not tell apart." It is unclear what this statement is supposed to mean. Presumably, Malone was stating that he did 10,000 evaluations and whenever he said that two hairs were from the same source, he was correct. But Malone's 10,000 evaluations were performed during casework, and in casework Malone did not know whether hairs really came from different people or the same person. All he knew was that he had *concluded* that the hair came from different people or the same person. It is not clear what the two exceptions are supposed to be. If Malone could not tell hairs apart in two cases, how did he determine that they in fact came from different sources? These nuances, however, were likely to be lost on the jury which probably assumed that what Malone's statement actually meant was either that the probability that his conclusion was false was 1 in 10,000, or, simply, that almost all hairs are unique.

We found similar testimony by state and local examiners. For example, in the 2009 trial of [William Campbell](#) for manslaughter in Ohio, Michael Trimpe, a forensic analyst at the Hamilton County Coroner's Office Crime Laboratory, said he microscopically examined several hairs from the passenger side window frame of a car that had crashed and found a single hair that Trimpe said was from the victim. Trimpe avoided committing Error Type 1 by saying "Now, that doesn't mean that I can say that that hair had to come from that person and no other, because there can be people with the same microscopic characteristics." However, he then went on to commit Error Type 3 (adding racial stereotyping for good measure) by saying "In all of my years looking at hairs, I've had one case where I couldn't tell the difference between two people, and they were an Asian brother and sister."

Here are the implied "fractions," verbal and numerical, that experts committing Error Type 3 constructed in the NRE FBI and non-FBI cases:

¹⁶⁶ Also see discussion in Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 58.

1

All my years looking at hair[\(William Campbell\)](#)*never*

> 100,000[\(Ledura Watkins\)](#)*never*

his career[\(Michael Jones\)](#)

8 – 10

10 years, while performing thousands of analyses[\(Kirk Odom\)](#)

2

10,000[\(Juan Matta Ballesteros\)](#)

4 – 5

10,000[\(Dennis Butler\)](#)*d) Commonly Encountered Phrases*

We can also do a more specific level of analysis than is accomplished by using the FBI's three error types. The FBI, the TFSC, and the Root Cause Analysis Report generated lists of "commonly encountered phrases" in MHCA testimony.¹⁶⁷ We combined and consolidated those lists, generating a list of 15 commonly encountered phrases. We then coded the 99 non-FBI, and the 18 FBI, inclusion cases. A single case could be coded for more than one phrase (the highest number of phrases in any one case was 4). For cases in which none of the commonly encountered phrases appeared, we added a new phrase to our list, adding a total of three new phrases. In addition, two commonly encountered phrases ("face analogy" and "seldom") identified by the Root Cause Analysis Report did not appear in our data at all and so were dropped from the analysis.¹⁶⁸ Thus, the resulting number of phrases shown in Table 15 is 16.

¹⁶⁷ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*, 3.

¹⁶⁸ In the case of "seldom," this finding was consistent with the finding of the Root Cause Analysis. The word "seldom" was, in fact, seldom used. It was used only three times in testimony exceeding the limits of the science between 1979 and 1999. Why it was called a "commonly encountered phrase" is not clear. ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 287.

Table 15. Commonly encountered phrases in inclusion cases. Cases were coded for up to 4 phrases. All phrases derive from the FBI review, the Root Cause Analysis Report, or the TFSC, except those indicated by an asterisk.

	Non-FBI cases (n=99)	FBI Cases (n=19)	Total (n=118)
(Exact/microscopically) same/similar	54	8	62
consistent with (coming from)	42	5	47
(Perfect) match	12	5	17
Probability statistic	12	3	15
(Completely/microscopically) indistinguishable/identical	6	7	13
(Reasonable degree of) scientific certainty	6	0	6
Experience	1	5	6
Rare, remote	3	2	5
Strong/confident	3	1	4
Unique	3	1	4
Correspond in microscopic characteristics *	1	0	1
Individualization	1	0	1
Unusual	1	0	1
Was in fact there *	1	0	1
Within the range *	1	0	1
Associated back to	1	0	1
TOTAL	148	37	185

The results are shown in Table 15. The most commonly encountered phrases by far were: (1) “same” or “similar” sometimes modified by a term like “exact” or “microscopically”; and (2) “consistent with.” What is notable, of course, is that both of these forms of testimony would usually have been considered “appropriate” in the FBI review. And, yet, they are by far the most common phrases in our data of convictions of innocent people. This, again, demonstrates the way in which seemingly innocuous and careful testimony can nonetheless damage innocent defendants.

For example, at the trial of [Curtis Flowers](#), who was falsely convicted of murder in 2004 in Arkansas, an MHCA expert testified that hair found in a bandana at the crime scene was “microscopically similar” to Flowers’s hair.¹⁶⁹

In the false conviction of [Robert Lee Miller, Jr.](#) for murder in 1988 in Oklahoma, forensic scientist Joyce Gilchrist testified that various hairs from the crime scene were “consistent with” Miller and “consistent with” the victim. She also testified that dog hairs and the crime scene were consistent with hairs from a neighborhood dog Miller was known to care for and that they “could share a common source.”

The problematic term “match,” sometimes modified by “perfect,” and baseless statistics, also appeared in a significant number of cases. For example, in the 1984 conviction of [Edward](#)

¹⁶⁹ This case should not be confused with the better-known exoneration case of Curtis Flowers of Mississippi. On the Mississippi Flowers case, see American Public Media, *In the Dark*, Season 2 (2018), <https://features.apmreports.org/in-the-dark/season-two/> (<https://perma.cc/6AE9-VT5S>); *Flowers v. Mississippi*, 139 S. Ct. 2228 (U.S. 2019).

[Honaker](#) for rape in Virginia, the MHCA expert testified that “it is unlikely that the hair would match anyone other than the defendant; but it is possible.”

In the conviction of [Glenn Payne](#) for the rape of a two-year-old girl in California in 1990, the MHCA expert, Mark Moriyama testified that there was a 1 in 2,700 chance that hair recovered from Payne’s person or clothing came from someone other than the survivor. Moriyama said another hair, recovered from a tablecloth found in a field with the survivor’s underwear, was a pubic hair, was from a Black person, and there was a 1 in 48 chance it came from someone other than Payne. Moriyama then multiplied 1 in 2,700 by 1 in 48 and concluded that the likelihood of those two hairs being present as a random occurrence was 1 in 129,600. Almost 20 years later, Moriyama repudiated his trial testimony, saying he no longer believed “the portion of the expert testimony relating to the statistical weight of the comparisons and [that] the uses of that statistical weight are scientifically valid. The scientific expert testimony that [he] would render now would be different and limited to stating that the hair sample found on the defendant could have come from the victim, and the hair sample found on the tablecloth used to cover the victim could have come from the defendant.” Egregious as Moriyama’s statistical testimony was, given what we learned above about the power of “could have come from” statements, we may legitimately wonder whether even his proposed more modulated testimony would have forestalled Payne’s 15 years in prison.

The phrase “indistinguishable” was also common, although that is strongly driven by FBI uses. The invocation of “experience” was also much more common in FBI testimony than non-FBI.

Of the phrases newly introduced from our data, the phrase “was in fact there” from the conviction of [Curtis McCarty](#) is notable. Forensic analyst Joyce Gilchrist testified that McCarty “was in fact” at the crime scene because a hair from the crime scene was found consistent with his sample hairs.¹⁷⁰ This testimony addressed what forensic scientists call an “activity-level proposition,” in which the analyst claims to be able to determine from the evidence the activity that produced the trace, rather than just the presence of the trace in a particular location.¹⁷¹ Gilchrist’s testimony to this activity-level proposition in McCarty was, of course, unfounded. That such testimony was given about hair evidence is particularly shocking since hair is among the kinds of forensic evidence for which alternative explanations for the appearance of hair in a location are quite plausible (i.e., it is relatively easy to “plant” hair; hair can easily be inadvertently transferred by clothing, people, objects, etc.).

For commonly encountered phrases, the only data from other analyses of MHCA testimony to which we can compare ours is the Root Cause Analysis. However, even comparison with the Root Cause Analysis is difficult because the Root Cause Analysis reports only temporal trends in the uses of commonly encountered phrases, not raw counts of their occurrence. However, from the Root Cause Analysis’s temporal trends reports, we can infer that the most commonly encountered phrases in the FBI review were “match,” which seems to have appeared around 2.5 times per transcript between 1979 and 1999 and a probability statistic (around 2 times). Next most common were “consistent with” (around 1.5 times) and “same” (around 1 time per transcript).¹⁷² Thus, the four most commonly encountered phrases are the same in our data and in the FBI review; however, the position of the top two within the top four is reversed. This is not

¹⁷⁰ Also see Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 98; Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 57.

¹⁷¹ Cook et al., *A Hierarchy of Propositions: Deciding Which Level to Address in Casework*.

¹⁷² ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 275-290.

surprising. “Same” and “consistent with” can be “appropriate,” and the FBI review found almost no “appropriate” testimony.

2. APPROPRIATE TESTIMONY

A reader of the press release and publicity surrounding the FBI review would come away with the impression that the problem with MHCA is erroneous testimony. The FBI review concentrated public attention on the egregiously erroneous testimony that it found in 96% of cases. The handful of cases in which testimony was “appropriate” was so small that appropriate testimony was treated almost an afterthought. A reader might come away with the impression that if testimony were appropriate, rather than erroneous, the problems identified by the FBI review would be solved. Our findings challenge that assumption.

a) *Appropriate Testimony in State and Local Cases*

We expected to find that “appropriate” testimony would be as rare in state/local cases as it was in the FBI review. Far from being rare, appropriate testimony was not only more common in state/local cases than it was in the FBI review—it was more common overall. Appropriate evidence appeared in 60% of non-FBI cases across all data sources (see Table 10). This is far more than the 4% rate reported in the 2015 press release, and it is still far more than the slightly larger 7% figure found in the 2018 snapshot, more than the 41% appropriate rate found by combining FBI testimony with written reports, and even more than the 50% erroneous rate of FBI written reports alone (Table 1).¹⁷³

If we look just at the NRE cases, testimony was appropriate in 61% of state/local exoneration cases and erroneous in only 39%. It might be argued that exoneration cases are different in some way from non-exoneration cases. But compare these results to the TFSC results. Like the NRE cases, the TFSC cases are all convictions. But the NRE cases all resulted in exoneration, and none of the TFSC cases resulted in exoneration. Nonetheless, the proportion of appropriate evidence in TFSC *non-exoneration* cases, 61% (60/99), was almost exactly the same as the proportion of appropriate evidence in NRE *exoneration* cases, 57% (33/58) (Table 10). This suggests, then, that exoneration cases are not especially different from convictions that did not result in exoneration. The distribution of appropriate and erroneous testimony in exoneration cases is similar to its distribution in all state/local convictions.

It appears, then, that the extremely high rate of erroneous FBI testimony was *not* replicated across state and local MHCA evidence. Instead, “appropriate” testimony was *more* common in exoneration cases than erroneous testimony. And, yet, as was widely noted at the time of the FBI review, most state and local MHCA experts were trained, at least in part, by the FBI. And, we did find that erroneous evidence was twice as common (80%) in the 15 cases in which we had affirmative information that the state/local examiner had been trained by the FBI, as it was in the state/local cases where we did not know whether or not the examiner was trained by the FBI. Still, the discrepancy between FBI and state/local case remains dramatic. Why might this be?

¹⁷³ [As noted above](#), however, there is little reason to compare the data reported in this study to FBI written reports because, even though it is not clear whether oral testimony was given in every single case, it is clear that most of the cases in this study involved courtroom testimony not written reports.

(1) Explanations for the Difference Between FBI and State/Local

One possibility is that the difference is explained by different data sources. We have noted [above](#) that there are data differences between the two studies. Most importantly, we did not have transcripts for nearly every case, as the researchers who conducted the FBI review did. One possibility is that by combing through transcripts, the FBI researchers were able to find some error somewhere in nearly every transcript. That same error might not have been apparent to us if it did not appear in the sources—like legal pleadings, judicial opinions, and press reports—that we relied upon in many cases.

Is the lower rate of erroneous testimony by state and local experts explained by transcript availability? This explanation seems unlikely because the lower rate of erroneous testimony holds even across the data sets in which complete transcripts *were* available. In particular, the TFSC review had access to full transcripts in all cases, and the rate of erroneous testimony (43%) is still much lower than the FBI's and only 4 points higher than the 39% rate found in the NRE data set for which we *lacked* transcripts in many cases (Table 10).¹⁷⁴

In addition, we performed a separate analysis of the 28 non-FBI cases for which we had access to trial transcripts (see section XI.B). These 28 cases represented 21 different experts.¹⁷⁵ Erroneous testimony appeared in only 9 (32%) of these 28 cases. In other words, erroneous testimony was *less* common in non-FBI cases for which we had access to transcripts than in cases in which we did not. We then undertook an additional review of the 19 non-FBI cases for which we had access to transcripts and the testimony was coded “appropriate.” In none of these cases did we find “hidden” erroneous testimony. These additional analyses suggest that the much higher prevalence of erroneous evidence in FBI cases reflects the actual practices inside and outside the FBI and not simply transcript availability.

Another possibility is that the FBI dealt with more serious crimes. The 19 FBI exoneration cases consisted of 12 murders, 4 sexual assaults, 1 manslaughter, 1 child sexual abuse, and 1 kidnapping. That is a higher proportion of murders than in the non-FBI cases. But as Table 12 clearly shows, almost all the MHCA exonerations, both FBI and non-FBI, involved very serious crimes. An explanation based on a claim that the FBI deals with more serious crimes would depend on drawing a distinction in terms of “seriousness” between murder and sexual assault, child sexual abuse, and attempted murder that does not seem plausible.

A third possibility is the difference between experts who work for a national, versus a local, agency. FBI MHCA examiners were small in number, but may have done hundreds of cases per year. FBI-trained local MHCA examiners were large in number, but were usually generalist criminalists who only occasionally analyzed hair. Perhaps these generalists testified more cautiously. Another explanation may be the experts' behavioral tendencies in court. In any particular trial, FBI experts are more likely to be “one-shotters,” for whom this may be their only appearance in this particular court with this particular judge and these particular attorneys. Local experts, in contrast, are more likely to be “repeat players” who may expect to spend their careers in this particular court with this particular judge and these particular attorneys.¹⁷⁶ They are, in that sense, part of the “courthouse workgroup.”¹⁷⁷ This may have encouraged FBI experts to be more

¹⁷⁴ As Table 10 shows, 39/99 (39%) of NRE cases were erroneous, and 25/58 (43%) of TFSC cases were erroneous.

¹⁷⁵ Some *experts* were represented in multiple *cases* because they testified in multi-defendant cases.

¹⁷⁶ Marc Galanter, *Why the "Haves" Come out Ahead: Speculations on the Limits of Legal Change*, 9 *Law and Society Review* 95 (1974).

¹⁷⁷ Malcolm M. Feeley, *The Process Is the Punishment: Handling Cases in Lower Criminal Court* (1980).

reckless about exaggerating the probative value of their testimony and local experts, correspondingly, to be more cautious.

On the other hand, the opposite may just as well be true: Local experts may enjoy a rapport in the courtroom that emboldens them to take more risks in their testimony.

It seems the likeliest explanation, however, is one already alluded to by several commentators on the FBI review: the culture in the FBI Hair and Fiber Unit. The Root Cause Analysis Report identified the following cultural problems in the FBI Hair and Fiber Unit:

1. Limited documentation
2. A preference for informal communication
3. Insufficient valuation of accreditation
4. The absence of a culture of “thoughtful-compliance” concerning reporting and testimony
5. Lack of deference to expertise
6. Absence of a questioning and learning environment
7. Excessive autonomy for examiners regarding testimony
8. A culture of identifying as detective rather than impartial scientists
9. Unwelcoming attitude toward non-agent examiners on the part of leadership.¹⁷⁸

Similarly, Robertson, an MHCA expert himself, blames cultural issues at the FBI for the scandal, noting that upon meeting FBI hair examiners in the 1980s, “one could only be struck by the almost robotic responses and belief in what they were doing.”¹⁷⁹

Our findings suggest that the FBI culture was dysfunctional in a way that somehow did not transfer over to the individuals and forensic service providers they trained. It would appear that in exoneration cases state and local experts testified more modestly despite being trained to give erroneous testimony by their FBI tutors.¹⁸⁰

b) Appropriate Testimony in FBI cases

Unlike in state/local cases, appropriate testimony did not contribute to a majority of exonerations in FBI cases. But appropriate testimony still did contribute to 3 of the 19 FBI exonerations (Table 8). That’s a rate of 16%, quadruple the 4% rate of appropriate testimony found in the FBI review.¹⁸¹ Appropriate testimony was much more common in exoneration cases than it was in the FBI review.

We cannot draw conclusions from this surprising finding, which is based on only nineteen cases. It may have arisen by chance, and we are limited in what we can conclude solely from data on exonerations without sets of undisturbed convictions and acquittals to which to compare them.¹⁸² Prosecutions that eventually ended in exoneration may have been weaker cases. This

¹⁷⁸ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 229-40.

¹⁷⁹ Robertson, *Managing the forensic examination of human hairs in contemporary forensic practice*, 253.

¹⁸⁰ M. Chris Fabricant & Tucker Carrington, *The Shifted Paradigm: Forensic Science’s Overdue Evolution from Magic to Law*, 4 Va. J. Crim. L. 1, 88 (2016).

¹⁸¹ Perhaps the FBI exoneration cases are different in some way from the cases in the FBI review? We don’t know because we don’t know how many of the 19 FBI exonerates, other than the 6 noted in note 96, were included in the review because the FBI has never released the names of the cases included in the review.

¹⁸² It would be ideal to have a comparison group of acquittals to compare the relationship between appropriate and erroneous testimony and acquittal or conviction. No such comparison data exists.

might have induced experts to be more cautious, but the opposite might also be true: that weaker cases induced experts to be more reckless.

But at the very least, the surprisingly high occurrence of exoneration cases involving appropriate testimony across all data sets and experts' employers challenges the intuitive assumption at the foundation of the FBI review—that erroneous testimony is the main, or even only, problem and that defendants will be safe from false conviction as long as testimony is appropriate. To the contrary, our results suggest that there may well, in fact, be little or no difference between “appropriate” and “erroneous” MHCA evidence in the degree to which they help win both true and false convictions.

(1) Why “Appropriate” Testimony May Be as Dangerous as “Erroneous” Testimony

Because we lack data on acquittals in cases with MHCA evidence, we do not know how many false convictions may have been averted by an analyst giving “appropriate,” rather than erroneous, testimony. What we know from exoneration cases, however, suggests, counterintuitively, that “appropriate” evidence may convict the innocent no less efficiently than “erroneous” evidence, and that there may be little or no difference between appropriate and erroneous evidence in terms of contributing to false convictions. This is consistent with McQuiston-Surrett and Saks's experimental study of the impact of MHCA on jurors, which found that erroneous “match” testimony and appropriate “similar-in-all-microscopic-characteristics” were equally “damaging to the defense, while communicating a comfortingly simple and easily grasped (though not very informative and presumably misleading) understanding of the basis for the identification opinion.”¹⁸³

The FBI review focused the public attention on erroneous statements. Egregious though those statements were, that focus may have been somewhat misplaced. What matters seems to be more the mere existence of the evidence—that there was forensic evidence of some kind against the defendant—than what was said about the evidence—even when what was said about the evidence is egregiously misleading. What, precisely, was said about that evidence—whether it was vague and almost meaningless or an outrageous overstatement—appears to matter less than we might have thought. This supports the findings of psychologists that jurors have difficulty assigning proper weight to evidence.¹⁸⁴ Juries, it would appear, believe forensic evidence implicates the defendant, no matter what is said about it.

Why is appropriate testimony so damaging and so prone to contribute to false convictions?

Essentially, “appropriate” testimony is testimony in which the expert resists the temptation to overstate the probative value of the evidence and instead reports what the FBI claims the science can support, which is:

that hair comparison could not be used to make a positive identification, but that it could indicate, at the broad class level, that a contributor of a known sample could be included in a pool of people of unknown size, as a possible source of the hair evidence.¹⁸⁵

¹⁸³ McQuiston-Surrett & Saks, *The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear*, 450.

¹⁸⁴ E.g., William C Thompson et al., *Perceived strength of forensic scientists' reporting statements about source conclusions*, 17 *Law, Probability and Risk* 133 (2018).

¹⁸⁵ FBI, *Microscopic Hair Comparison Analysis*.

We are left with a statement that is pernicious in its ambiguity. On the one hand, the statement, if read carefully, means almost nothing. The expert is told to testify that the POI is “included in a pool of people of unknown size” with no further guidance as to the size of that pool. For all we know, the pool could consist of millions, or even billions, of people. In some of the testimony in our data set, witnesses candidly acknowledged this near valuelessness. As noted above, during the 1985 trial of [Steven Avery](#) for rape,

A state forensic serologist, Sherry Culhane, testified that a hair recovered from a shirt of Avery’s was “similar” and “consistent” with Beerntsen’s hair. She conceded that the hairs of many people are consistent with one another, that she could not give a probability that the hairs were from the same source, and that all she could say was “that it’s not impossible” the hairs were from the same source.

It would be difficult to call Culhane’s testimony “false.” For all we know, the hair she looked at *was* “similar” and “consistent” with Beerntsen’s hair. Moreover, the testimony Culhane offered was as modest as any evidence imaginable: she claimed only “that it’s not impossible” that Beerntsen was the source of the hair. This is an extremely modest claim that is certainly not refuted by the fact that post-conviction DNA testing excluded Avery as the rapist and that his conviction was dismissed.

But, at the same time, we must ask: did the jury *interpret* the evidence to be as weak as a literal reading of Culhane’s words suggests? If the jury took literally Culhane’s statement that she had only established that it was not impossible that Beerntsen was the source of the hair, then the jury should have treated the MHCA evidence as insignificant; they should have afforded it almost no weight at all. But did the jury perceive the hair evidence as more significant than that? Did Culhane’s testimony that the hair was “similar” and “consistent” with Beerntsen’s seem to the jury like it must mean *something* and thus help contribute to Avery’s conviction? As the Canadian Driskell Report [on a miscarriage of justice involving MHCA](#) noted, “The main danger associated with the language of ‘consistency’ is that a jury or judge may misinterpret a statement meant by a witness to suggest only a weak association, as indicating a much closer connection than the witness intends.”¹⁸⁶

On the one hand, one can commend Culhane for her candor. She indeed transparently conveyed the limits of the science of MHCA. But, at the same time, one must wonder about the impact of such testimony on fact-finders. To say that a defendant is included in a pool of potential contributors that could consist of up to billions of people means that the evidence should have almost no weight. Do fact-finders assign it that little weight? Or, does the fact that the hair was “consistent” lull them into assigning the evidence far more weight than the expert intended? This potential for misunderstanding suggests that the testimony’s prejudicial value may outweigh its probative value which may be grounds for exclusion under Federal Rule of Evidence 403.

Perhaps the best way to illustrate this point is to draw on our own data and offer some examples of testimony that we coded “appropriate.” The reader is asked to consider whether a fact-finder would in fact have assigned the evidence as little weight as the FBI thinks it warranted.

Consider, for example, the trial of [Dale Brison](#) for sexual assault in Pennsylvania in 1991. On direct examination, the MHCA expert, Debra Fertal, testified:

¹⁸⁶ Patrick J. LeSage, *Report of the Commission of Inquiry into Certain Aspects of the Trial and Conviction of James Driskell*, 165 (Jan., 2007), http://www.driskellinquiry.ca/pdf/final_report_jan2007.pdf (<https://perma.cc/8VT5-TKR7>).

A: . . . The hair I found in the underwear exhibited characteristics that were similar to the characteristics in Dale Brison's standard head hair (144).

On cross-examination, Fertal was asked:

Q: When you say similar, you are not able to make an exact match of hair?

A: No. It's subjective.

Q: What do you mean by subjective?

A: You're looking at something side by side and you are visually deciding whether it looks similar to you.

Q: Certainly a hair comparison is no where [sic] near exact as a fingerprint comparison, is it?

A: No. . . .

Q: And would you agree with me that certainly other individuals can have hair consistent with the hair that you're examining?

A: Yes. There's no way to say that this hair comes from one particular person to the exclusion of all others.

Q: Therefore, if you examined another black male's hair, it could have consistent patterns to what you found here?

A: Yes, the possibility does exist.

Q: At least some percentage of the population would have similar type hair?

A: I would have no way of knowing the population.

On re-direct, Fertal was asked:

Q: Mr. DiFabio asked you if you had done a comparison or a study on the population with this specific hair?

A: Well, statistics really doesn't hold itself to hair comparison, because it is so subjective.

Q: But you don't do any population studies on any hair comparisons you do?

A: Oh, no, never.

Q: Miss [sic] Fertal, are you confident in saying that the hair that was found in that underwear is similar and consistent with the hair that you received from Dale Brison's head?

A: Yes, I am.¹⁸⁷

This testimony was coded appropriate. The witness admirably resisted crossing the line into erroneous testimony. Yet consider whether a fact-finder would come away with the impression that Brison was merely included in a pool of millions, or perhaps billions, of people whose hair might be found consistent with the hair from the underwear. Or, would the fact-finder come away with the impression that the similarity between Brison's hair and the hair in the underwear was somehow significant? The prosecutor's elicitation of the expert's expression of confidence on re-direct, also illustrates how the adversarial exchanges of cross and re-direct examination can serve to bolster seemingly "appropriate" testimony, an issue discussed at greater length [below](#).

As another example, consider the testimony of MHCA expert Melvin Hett at the well-known trial of [Ron Williamson](#) and [Dennis Fritz](#) for the murder of Debra Sue Carter in Oklahoma in 1988.¹⁸⁸ Hett testified that hairs on a washcloth found in the victim's mouth and known pubic hairs taken from Fritz were "consistent microscopically and could have the same source." He then clarified:

[T]he hairs either did originate from that [known] source, or there could be or might be another individual in the world somewhere that might have the same microscopic characteristics. In other words, hairs are not an absolute identification, but they either came from this individual or there is—could be another individual somewhere in the world that would have the same characteristics to their hair.¹⁸⁹

This testimony seemed to us very close to the boundary between "appropriate" and "erroneous" testimony. On the one hand, it acknowledged the possibility that another person is the source of the hair. On the other hand, rather than stating that alternative contributors constitute "a pool of unknown size," it appeared to limit that pool to a single individual. During our coding process, we believed that the FBI would have coded this testimony appropriate, and that opinion was confirmed by someone knowledgeable about the FBI review.¹⁹⁰ However, one has to question whether the statement "could be another individual somewhere in the world" really conveys the spirit of inclusion in a pool of unknown size which could consist of millions or even billions of people.¹⁹¹

Indeed, paradoxically, although Hett's testimony would be considered "appropriate" by the guidelines of the FBI review, it was not even considered *admissible* by the Eastern District Court of

¹⁸⁷ On file with the National Registry of Exonerations and available upon request from Innocence Project, The Innocence Record, available at <https://www.innocencerecord.org> (<https://perma.cc/8H75-8URS>), BRI-000629-000642. (pp. 144-157)

¹⁸⁸ John Grisham, *The Innocent Man: Murder and Injustice in a Small Town* (2006); Dennis Fritz, *Journey Toward Justice* (2006); Scheck et al., *Actual Innocence: When Justice Goes Wrong and How to Make It Right*; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 96.

¹⁸⁹ *Williamson v. Reynolds*, 1554.

¹⁹⁰ Vanessa Antoun Personal communication to Cole (Mar. 10, 2020).

¹⁹¹ Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*.

Oklahoma.¹⁹² In making this finding, the court noted the potential for the testimony to mislead the fact-finder:

Other forms of expert examination and testimony have been criticized because jurors may be awed by an “aura of special reliability and trustworthiness” which may cause undue prejudice, confuse the issues or mislead the jury. . . . In the case of hair expert testimony the jurors do not have the opportunity for direct evaluation. Instead, they hear an abbreviated summary of the characteristics of hair and testimony of the expert's overall conclusions. . . .

The clear implication from the expert's testimony in Petitioner's trial was that 4 of the hairs found at the victim's apartment belonged to Petitioner. As witness to the incorrect conclusion that could result from this testimony, the prosecutor said in his closing argument, “[T]here's a match”. . . . Even the Court of Criminal Appeals misinterpreted and overstated the hair evidence by writing, “Hair evidence *placed [Petitioner] at the decedent's apartment.*”¹⁹³

The District Court's ruling was reversed by the Tenth Circuit Court of Appeals.¹⁹⁴ The Tenth Circuit did not rule on the District Court's findings, but rather ruled that the District Court had applied the wrong standard. The circuit court ordered the state trial court to hold an evidentiary hearing on MHCA, a hearing that was never held because Williamson and Fritz were exonerated by post-conviction DNA testing.

Despite a federal district court's finding as early as 1995 in “[t]he most significant post-*Daubert* challenge to microscopic hair analysis” that even “*appropriate*” MHCA testimony was inadmissible, “[s]urprisingly, *Williamson* did not have an impact. Many cases continued to admit testimony about microscopic hair analysis,” both “appropriate” and “erroneous,” for years afterwards.¹⁹⁵

C. DEFENSE EXPERTS

Although we were not able to systematically search all 125 cases to determine in how many defense experts testified, defense experts appeared to be rare. In at least two cases, however, defense experts contradicted the state's MHCA experts' conclusions.

In the 1988 prosecutions of [Omar Saunders](#), [Calvin](#) and [Larry Ollins](#), and [Marcellius Bradford](#), for the 1986 murder of Lori Roscetti in Chicago, the State's MHCA expert, Raymond Lenz, testified that a hair from Roscetti's car was “similar” to a known sample of Saunders's hair and a hair from her clothing was “similar” to a sample of Larry Ollins's hair.¹⁹⁶ At Saunders's trial, a

¹⁹² *Williamson v. Reynolds*, 1558.

¹⁹³ *Id.*, 1557., citations omitted; District Court's emphasis

¹⁹⁴ *Williamson v. Ward*, 110 F.3d 1508, 1522 (10th Cir. 1997).

¹⁹⁵ Giannelli, *Microscopic Hair Comparisons: A Cautionary Tale*, 21; Fabricant & Carrington, *The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law*, 84; Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, 96. “*Daubert*” refers to *Daubert v. Merrell Dow Pharmaceuticals*, a landmark U.S. Supreme Court case that places a “gatekeeping” responsibility on trial judges to halt the admission of expert testimony that is not both “relevant and reliable.” *Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (U.S. 1993).

¹⁹⁶ *People v. Larry Ollins*, Tr, trans., Cir. Ct. Cook Cty. Ill., June 16, 1988 at 34-40, on file with [The Innocence Record](#), OLL-000680-000686.

well-known criminalist, Skip Palenik, testified for the defense that Saunders was “unlikely” to be the source of the hair.¹⁹⁷

In 1982, law Professor Edward Imwinkelried suggested that defense attorneys rebut MHCA evidence by calling experts who use “more reliable” techniques than conventional microscopy.¹⁹⁸ Coincidentally or not, that same year, at the trial of [Steven Linscott](#) for murder (also in Chicago), after the State’s expert, Mohammed Tahir, testified that hair from the crime scene was “consistent with” and “similar to” the standards taken from Linscott, the defense called an expert witness, Kenneth Siegesmund, an Associate Professor of Anatomy at Medical College in Milwaukee with expertise in “microevidence,” including hair. In addition to the microscopic observations used in conventional MHCA, Siegesmund did what he claimed were higher-tech analyses using a scanning electron microscope and x-ray elemental micro-analysis. Based on all three analyses, he excluded Linscott as the source of the hair.¹⁹⁹

These rebuttals of the MHCA evidence were to no avail. All the above defendants were convicted.

It is important to note that there are also some exoneration cases in which an MHCA expert testified on behalf of the defense. An example is the [Brandon Moon](#) conviction for sexual assault in 1988 in Texas. MHCA expert David Mayham testified that Moon was excluded as the source of a hair. Moon was nonetheless convicted based on eyewitness and serology evidence.

D. CROSS AND RE-DIRECT EXAMINATION

How effective was cross-examination of MHCA experts? Were “erroneous” statements elicited during direct testimony, cross examination, or re-direct testimony? Because we do not have full transcripts of all our cases, counting these variables would not be particularly meaningful. However, based on the transcripts and direct and indirect quotations that we do have, we can make some comments about this issue and give some examples.

Many defense attorneys were able to raise the “pool of unknown size” issue on cross and elicit some sort of admission that another person, or many other people, might be the source of the hair. Often, on re-direct, the prosecutor might try to compensate for this admission by soliciting some sort of stronger affirmation that the POI was really the true source of the hair.

The transcript of the 1994 murder conviction of [Michael Blair](#) for murder in Texas²⁰⁰ illustrates how cross and re-direct examination impact MHCA testimony. On direct examination, the MHCA expert, Charles Linch, initially made the following statement:

The microscopic appearance of [the victim’s] head hairs, the internal structures, are the same in all microscopic characteristics as the head hair fragment found in the passenger front floor of Mr. Blair’s vehicle.

With regard to other hairs, he testified that the victim:

¹⁹⁷ Matt O’Connor, *2nd Convicted in Student’s Death*, Chicago Tribune, May 11, 1988.

¹⁹⁸ Imwinkelried, *Forensic Hair Analysis: The Case against the Underemployment of Scientific Evidence*, 58.

¹⁹⁹ Document in [The Innocence Record](#), LIN-000052-000064. Other experts dispute that these techniques are appropriate for hair analysis.

²⁰⁰ Also see discussion in Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 59.

cannot be excluded as the donor of the Caucasian hairs recovered from the blanket in Mr. Blair's car.

And, with regard to a third hair, he testified:

The opaque head hair from the strand of hairs from Jack Carter Park has strong microscopic indications that they had origin with Mr. Blair.

On cross, defense counsel elicited a number of qualifications:

My testimony is such that the three head hairs from the car had the same microscopic characteristics as [the victim]. I'm not here to tell you that they came from her to the exclusion of all other people in the world.

And:

Q: Okay. Because there is no such thing or is there such a thing as a—a situation where you can say this hair came from this person?

A: You can't do that. No, sir.

Q: It's—like you can with a fingerprint?

A: I'm not a fingerprint expert, but I understand a fingerprint from a person is that person. Right.

And:

Q: . . . do those, do any of those exhibits at all, tell you who the unknown hair belonged to?

A: With absolute certainty of all persons, no, sir. They do not.

After defense counsel elicited these concessions, the prosecutor elicited the following on re-direct:

Q: Okay. There were two hairs that you found in that clump that had, did they not, the same fine microscopic characteristics as that of the Defendant, Michael Blair; is that correct?

A: Yes, sir. They did.

Q: It's either Michael Blair's hair or someone with those exact same characteristics?

A: That's right.²⁰¹

Thus, we see that defense counsel may have weakened the evidence on cross by establishing that hair associations do not reduce the potential donor pool to a single person. The prosecutor then compensated for that on re-direct by eliciting that any other members of the donor pool would have to have the “exact same characteristics” as Blair. The jury was given no evidence as to how rare it is for people in general to share “exact same characteristics,” let alone the specific

²⁰¹ See transcript at Garrett, Convicting the Innocent: DNA Exonerations Database, available at <https://convictingtheinnocent.com> (<https://perma.cc/FN5F-SKGZ>), 737-807. Morgan, Wrongful Convictions and Forensic Science Errors: Case Studies and Root Causes, 57-58. reports that Linch, unbeknownst to the defense, was hospitalized in a psychiatric unit at the time of his testimony in the Blair case and later stated that he should not have been permitted to testify.

characteristics of Blair’s hair, but they may well have assumed that it was very rare or even non-existent.

For a particularly convoluted example, consider the conviction of [Drew Whitley](#) for murder in 1989 in Pennsylvania.²⁰² The evidence appears to have been weak, even by MHCA standards. The expert, Dorothy Menges was not even willing to use what we have seen above is the most commonly used language: “consistent.” On direct examination, she testified:

Because these hair fragments were so small, I could not make the statement that they were microscopically consistent, but I did see so many overlapping characteristics within the questioned hairs and the standard hairs that I want to make some kind of statement as to their similarities.

However, having said she could not use the term “consistent,” she almost immediately proceeded to do so:

What was present of these questioned hairs had very similar consistent microscopic characteristics of the known standards of the facial hair of Drew Whitley.

Defense counsel correctly picked up on this contradiction and attempted to force the witness to commit on the issue of “consistency”:

Q: Your term, the most positive you can get is “microscopically consistent”?

A: That is correct.

Q: That is not what you said in this case. You said they overlap.

A: I say there are very many overlapping characteristics due to the lack of characteristics in such small hairs. That is all I would like to say about them.

The judge, perhaps sensing the lack of clarity, intervened:

Q: Can you say whether or not they are consistent?

A: I have no reason to believe that those questioned hairs I examined and compared to the facial hairs of Drew Whitley, I have no reason to believe these hairs could not have come from Drew Whitley.

The court had now elicited a statement that may well have sounded to the jury much stronger than the expert’s original statement made on direct examination. Both the judge and defense counsel immediately stated that the expert’s answer was not responding to the question they had asked, and the court made another attempt at clarification:

Q: I will rephrase the question. Can you say whether or not the samples of hairs you analyzed, the ones you analyzed from the stocking cap and the samples of the defendant, Mr. Drew Whitley, can you say those two samples you analyzed, are they consistent? Can you say whether or not they are consistent?

A: I believe they have many, many consistencies microscopically.

Q: That is not an answer to the question. . . .

²⁰² Also see discussion in Garrett & Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 61.

A: All I can say, they are not inconsistent.

Q: Can you say they are consistent?

A: I found no inconsistencies. Based on what I am basing my comparison on, yes, they are consistent.

Q: You are saying they are consistent?

A: I found no inconsistencies in the hairs, and if there's no inconsistencies based on what I am seeing of the hairs, there is a consistency.

Q: You are saying they are consistent?

A: Yes.

The expert had begun her direct testimony saying she did not want to use the term “consistent.” During reasonably competent cross examination, the judge intervened and through persistent questioning, managed to get the expert to say precisely what she originally said she wanted to avoid. Again, defense counsel picked up on the contradiction:

Q: On direct examination you said, and I wrote it down, these hairs were not microscopically consistent.

The expert then reversed course again:

A: I wouldn't go that far to say they were microscopically consistent . . . there were so many microscopic characteristics that were similar and some and some overlapping characteristics that were similar to the facial hair standards of Drew Whitley.

Defense counsel then again elicited that MHCA cannot provide certain identification:

Q: You can't say any hair submitted to you in this case besides the standards of Drew Whitley you received came from him?

A: That is correct.

The judge then intervened again:

Q: You can't say every hair belonged to a certain individual? You can't say it belongs to the defendant?

A: That is correct.

Q: . . . well, can you limit the number of people whom these samples could be associated with?

A: They were also very similar to each other in the nature of the way they were cut, the way the colors were consistent among themselves, the diameter, pigment distribution, the type of surface that the hairs had, I would believe they would come from one individual.

Q: Pardon me?

A: It was my opinion they would have come from one individual.²⁰³

²⁰³ See transcript at Garrett, Convicting the Innocent: DNA Exonerations Database, available at <https://convictingtheinnocent.com> (<https://perma.cc/FN5F-SKGZ>), 898-951.

At this point, the witness was excused. In this final exchange, it is unclear what the judge and expert are talking about or even whether they are both talking about the same thing. When the expert said “one individual,” were they merely saying that all the questioned hairs came from the same person? That would tell the jury absolutely nothing about who that person was. Or, was the expert saying that the hairs do not come from “a pool of unknown size”—that the potential donor pool of the hairs is “one individual,” the defendant? While it is impossible to tell from this transcript what the two speakers actually meant, it is quite possible that the jury understood the expert to mean the latter. Thus, we see that through the back-and-forth of adversarial questioning what began as evidence that was weak even by the standards of MHCA ended up transformed into a statement that could be interpreted as a statement of individualization (what the FBI review called Error Type 1).

We can see that part of the reason for this transformation may be the nature of adversarial questioning itself: as the expert, who was clearly uncomfortable with the evidence, sought to defend their rather vague statement, the evidence strengthened. But we can also see that the MHCA discipline left its experts ill-equipped with language to describe their findings, leaving them with language that—as the lawyers quickly realized—didn’t make a lot of sense.

E. PROSECUTORS AND CLOSING STATEMENTS

In cases in which MHCA testimony was both “appropriate” and “erroneous,” prosecutors sometimes made statements indicating that the strength of the evidence was greater even than the MHCA expert said it was.

For example, in the trial of [Michael Tillman](#) for murder in 1986 in Chicago, the MHCA expert testified that hair from the victim’s bathroom floor was similar to Tillman’s hair (“appropriate” testimony). State’s Attorney Lawrence M. Lykowski, however, told the jury that Tillman’s pubic hair had been found at the crime scene.

In some cases, prosecutors engaged in still worse behavior. In the prosecution of [Kevin Martin](#) for manslaughter in the District of Columbia in 1984, the MHCA report stated that a pubic hair found on one of the victim’s shoes was “like” Martin’s hair, but also “like” the victim’s hair. The hair then fell out of the evidence bag and was lost. The prosecutor told Martin’s lawyer and the judge that “the FBI technician is prepared to say it was Kevin Martin’s hair on Mr. Brown’s sneaker” (Error Type 1), and Martin pled no contest to manslaughter.

Even when the expert testimony was erroneous, prosecutors sometimes made it even more so. In the federal trial of [Juan Matta Ballesteros](#) for kidnapping in 1990, FBI MHCA expert Michael Malone testified about hairs that “matched”—an improper term—the victim and Ballesteros. He inappropriately used the number of analyses performed during his career to suggest an accuracy rate for MHCA.

The prosecutor made this “erroneous” testimony even stronger:

in the opening statement to the jury, the prosecution said that in “a bathroom adjoining a bedroom of the main house at Lope de Vega was found a hair which matches in every comparable respect” the hair of [Matta Ballesteros]. In addition, the prosecution said that “a hair was found in the adjoining bedroom, ladies and gentlemen, which matches in every comparable respect the known hair of agent Camarena.”

In the closing argument to the jury, the prosecution declared, “[Malone] went on at great

length – and I won’t repeat it – but he has incredible training, has examined thousands and thousands of hairs. He takes final exams, and every time he passes exams with flying colors. He has taught, he has published. This man knows his stuff. That is unequivocal. And his conclusion to you after perhaps a full day of testimony, was that Camarena was at the Lope de Vega house and particularly in the guest house. And Camarena had hair forcibly removed from him in the guest house.”

Finally, the prosecutor told the jury that a hair found in the guest house was “absolutely indistinguishable” from Matta Ballesteros’s hair.

F. MISCONDUCT

As discussed above, the primary problems with MHCA evidence are overstated testimony, vague testimony that masks the technique’s poor discriminating power, and wrong conclusions. However, in some cases we know about outright misconduct involving MHCA evidence. There are probably more such cases that we don’t know about.

In the 1988 conviction of [Anthony Michael Green](#) for sexual assault in Cleveland, forensic examiner Joseph Serowik

testified about his analysis of a hair found on the washcloth. He found the hair inconsistent with Green’s pubic hair. Trace evidence expert Max Houck would later state in a 2004 affidavit that “this should have precluded further examinations” of Green’s hairs. Determining the somatic origin of a hair (which part of the body it came from) is considered an easy task. By comparing the washcloth hair to Green’s pubic hair, Serowik should have already determined that the washcloth hair was a pubic hair, according to Houck.

Shockingly, however,

instead of ceasing examination, Serowik requested samples of chest and head hair from Green and compared the washcloth hair to those samples. Serowik then reported that the washcloth hair was “found to be consistent with the known hair sample” from Green “with respect to all of the characteristics considered.”

But this was false. Serowik’s notes stated that the medullae of the washcloth hair and Green’s hair “did not match” and that the washcloth hair and Green’s head hair samples “appear to be similar in all respects except root, length, and perhaps color.” Houck later stated, “Color is one of the most discriminating characteristics for hair comparisons and for [Serowik] not to be able to compare the hairs’ color invalidates his analysis and conclusions. . . . [Serowik] failed to exclude a questioned hair that exhibited, by his own casework notes, significant differences from a known hair sample.” Houck also stated that “to report that the hairs were consistent in ‘all of the characteristics considered’ when the case notes declared the hairs ‘appear’ to be similar ‘except for’ certain traits is professional misconduct.”

In the 1981 capital conviction of [John Huffington](#) in Maryland, FBI MHCA examiner Michael Malone

testified that the hair from [the victim’s] trailer “microscopically matched the head hairs of Mr. Huffington—that is, they were indistinguishable from Mr. Huffington’s head hairs. You could not tell them apart.” Malone did acknowledge that microscopic hair comparison was “not a positive personal identification.”

In 1999, as part of a DOJ Office of the Inspector General audit of Malone’s work,

Steve Robertson, a hair and fiber analyst, reviewed Malone’s bench notes, 80 evidence specimens, and Malone’s testimony at Huffington’s [trial].

Robertson said that while Malone’s testimony was consistent with his laboratory report, it was not consistent with his bench notes.

In March 2013, the FBI reported that the DNA testing had excluded Huffington as the source of the hairs that Malone had said were microscopically similar to Huffington’s.

G. COMPENSATION

There are two primary means of compensation for exonerees. The first is statutory compensation by some states which allocate set amounts for time served for qualifying exonerees. The second is civil lawsuits. The amount awarded in these lawsuits can vary.²⁰⁴

MHCA exonerees were better compensated than non-MHCA exonerees convicted of the 7 comparable violent crimes prior to 2010.²⁰⁵

1. STATUTORY COMPENSATION

Seventy-four (60%) of 123²⁰⁶ MHCA exonerees received statutory compensation for a total of more than \$95 million, an average of around \$1.2 million per compensated exoneree. This amounted to around \$81,000 per year lost for compensated exonerees.

By contrast, 734 (39%) of the 1,874²⁰⁷ comparable non-MHCA exonerees received statutory compensation, an average of around \$1 million per compensated exoneree. This amounted to around \$68,000 per years lost for compensated exonerees.

²⁰⁴ For more discussion of compensation of exonerees, see the Registry’s [Issues page on Compensation](#).

²⁰⁵ No MHCA exoneree was convicted after 2009, so we limit the comparison set of non-MHCA cases to convictions prior to 2010.

²⁰⁶ Two MHCA exonerees are excluded from these calculations: Juan Matta Ballesteros because the Registry’s compensation data excludes federal exonerees; and John Ausby because he was added to the Registry after the most recent update of the compensation data. Neither is likely to receive compensation. Ausby died before being exonerated, and Matta Ballesteros is in prison for another crime.

²⁰⁷ The number of exonerees convicted of comparable crimes is slightly different for the compensation data than in the rest of this Report. This is because the Registry’s compensation data excludes federal cases, and the compensation data is maintained separately and is slightly older than the data used elsewhere in this Report.

MHCA exonerees are mostly DNA exonerees, and, not surprisingly, their compensation was similar to that of DNA exonerees. Two hundred fifty-eight (58%) of 447 non-MHCA DNA exonerees convicted prior to 2010 received statutory compensation, an average of around \$1 million per compensated exoneree. This amounted to around \$67,000 per year lost.

2. CIVIL COMPENSATION

Fifty-one (41%) of the 123 received civil compensation. They received a total of more than \$250 million and an average of around \$5 million per compensated exoneree. This amounts to around \$300,000 per year lost for compensated exonerees.

In contrast, 484 (26%) of the 1,874 comparable non-MHCA exonerees received civil compensation. The non-MHCA compensated exonerees received a similar average of around \$5 million per compensated exoneree. However, they served, on average, one less year in prison, so this amounted to around \$320,000 per year lost for compensated exonerees.

One hundred sixty-one of the 447 comparable DNA exonerees (36%) received civil compensation, an average of around \$5.6 million per compensated exoneree. This amounted to around \$350,000 per year lost for compensated exonerees.

However, MHCA exonerees' compensation was concentrated in a subset of individuals. Thirty-eight of the 123 MHCA exonerees, nearly a third, received no compensation at all. Only two of those 38 served no time in prison, leaving 36 who served time in prison and received no compensation. Forty received both statutory and civil compensation. Thirty-four received statutory compensation but no civil compensation, and 11 received civil compensation, but no statutory compensation (Table 16).

Table 16. MHCA exonerees' compensation (n=123). Federal exonerees and exonerees posted in 2023 are excluded.

Statutory	Civil		Total
	Yes	No	
Yes	40 (32%)	34 (28%)	74
No	11 (9%)	38 (31%)	49
Total	51	72	123

Still, MHCA exonerees fared better than comparable non-MHCA exonerees, almost half of whom received no compensation. In addition, many fewer non-MHCA exonerees received both forms of compensation. Again, MHCA exonerees' compensation was more similar to non-MHCA DNA exonerees' compensation (Table 17).

Table 17. Comparison of compensation for MHCA, comparable non-MHCA, and comparable non-MHCA DNA exoneration.

MHCA (n=123)	Civil		Non-MHCA (n=1,874)	Civil		Non-MHCA DNA (n=447)	Civil	
	Yes	No		Statutory	Yes		No	Statutory
Yes	32%	28%	Yes	13%	26%	Yes	22%	36%
No	9%	31%	No	13%	48%	No	14%	28%

X. CONCLUSION

While this Report has documented many gross injustices to which MHCA contributed, the use of MHCA is declining in the face of competition from a newer technology, mtDNA. Some may wonder whether MHCA is consigned to the past and whether it matters anymore.

MHCA still does matter for several reasons. First, MHCA is still used. As the Root Cause Analysis Report noted in 2018, “The justice system relies on microscopic comparison analysis of hair. Law enforcement and investigators continue to find human hairs at crime scenes and request that the FBI microscopically analyze them and testify on the results. MHCA remains a source of data for the justice system.”²⁰⁸ Other laboratories continue to use MHCA, and experts continue to advocate using it either standalone or as a screening tool.²⁰⁹ In addition, people convicted based on MHCA are still being exonerated today, and most of the promised state audits have still not been completed.

Perhaps most importantly, though, the kinds of testimonial misstatements documented in this Report—both the egregious misstatements and the anodyne “consistent with” statements—transcend the MHCA discipline. We find source attribution statements and baseless probabilities in other forensic disciplines. We find “consistent with” testimony in numerous other disciplines, such as forensic pathology, other medical testimony, and sexual assault examination.²¹⁰ These statements are appealing to expert disciplines that lack foundational data because they mean almost nothing while appearing to mean at least something and perhaps anything. The problems exposed with MHCA should be a warning about these kinds of statements in any expert domain.

Our findings about the perniciousness of supposedly appropriate “consistent with” testimony may also be relevant to legal debates about the admissibility of expert evidence. Socio-legal scholars, including an author of this Report, have tended to focus on trying to control what expert witnesses say to fact-finders, preventing the kinds of egregious misstatements documented above and requiring experts to make statements that are scientifically defensible, logical, and statistically coherent.²¹¹

This Report challenges that approach. For MHCA evidence, in exoneration cases at least, what was said about the evidence seems to have mattered little. Surprisingly, erroneous MHCA

²⁰⁸ ABS Group, *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners*, 18.

²⁰⁹ Houck, *Is Microscopic Hair Comparison a Legitimate Science?*, 52; Laura Wilkinson & Claire Gwinnett, *An International Survey into the Analysis and Interpretation of Microscopic Hair Evidence by Forensic Hair Examiners*, 308 *For. Sci. Int.* 110158 (2020) <http://dx.doi.org/10.1016/j.forsciint.2020.110158>; Robertson & Brooks, *Historical Context to Contemporary Practice in Hair Examinations*, 22-23; Airlie et al., *Forensic Hair Analysis—Worldwide Survey Results*, 7; FBI, *FBI Testimony on Microscopic Hair Analysis Contained Error in at Least 90 Percent of Cases in Ongoing Review*. NRC, *Strengthening Forensic Science in the United States: A Path Forward*, National Research Council, 160 (Feb., 2009), <https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf> (<http://perma.cc/DVV2-6CTZ>); Driskell Inquiry, *Transcript of Proceedings before the Commission*, Commission of Inquiry into Certain Aspects of the Trial and Conviction of James Driskell, 5559 (Sept. 21, 2006).

²¹⁰ Ross et al., *Consistent With: What Doctors Say and Jurors Hear*.

²¹¹ E.g., Edward J. Imwinkelried, *The Relativity of Reliability*, 34 *Seton Hall Law Review* 269, 277 (2004); Bert Black, *Focus on Science, Not Checklists*, 39 *Trial* 26 (2003); D. A. Nance, *Reliability and the Admissibility of Experts*, 34 *Seton Hall Law Review* 191, 253 (2003); Margaret A. Berger, *Expert Testimony in Criminal Proceedings: Questions Daubert Does Not Answer*, 33 *id.* 1125, 1140; Erica Beecher-Monas, *A Ray of Light for Judges Blinded by Science: Triers of Science and Intellectual Due Process*, 33 *Ga. L. Rev.* 1047, 1062 (1999); David L. Faigman, *Expert Evidence in Flatland: The Geometry of a World Without Scientific Culture*, 33 *Seton Hall Law Review* 255, 258 (2004); Samuel Gross & Jennifer L. Mnookin, *Expert Information and Expert Evidence: A Preliminary Taxonomy*, 34 *id.* 143 (2003); Simon A. Cole, *Where the Rubber Meets the Road: Thinking About Expert Evidence as Expert Testimony*, 52 *Villanova Law Review* 803, 116 (2007).

evidence, appalling though it may have been, appears not to have been a crucial factor in MHCA's contribution to false convictions. Instead, any MHCA evidence, whether erroneous or "appropriate," appears to have done an equally good job at contributing to false convictions. As Garrett and Mitchell suggested in 2013, "while the legal and forensics community may be rightly concerned about the manner in which forensic conclusions are expressed, modest testimonial conclusions may be just as problematic as overstated conclusions."²¹² It is difficult to avoid the conclusion that what actually mattered to juries was the mere fact that the evidence was called "forensic" and that it appeared to tend to implicate the defendant in some way. (Consider once again the Avery case discussed in the Preface, in which Avery was convicted in part based on MHCA evidence in which the expert witness merely said it was "not impossible" that Avery was the source of the hair found on the survivor.) If this is true, then control of egregious misstatements will not be enough; we will need more robust gatekeeping, not just for MHCA but for all expert evidence.

As we discussed in the Preface, this Report originated from a coding discussion about "appropriate" MHCA evidence. In the end, we decided that even MHCA testimony about mere "consistency" was "False or Misleading Forensic Evidence." While these statements were not "false," they were potentially highly misleading.

We think this Report vindicates our coding decision. The Report exposes the sheer number of innocent people who were convicted, at least in part, based on these supposedly true and innocuous statements. This is not to minimize the egregious exaggerations of the value of MHCA evidence which *also* contributed to many false convictions. As this Report documents, MHCA testimony spanned the range from the most egregious overstatements to the most anodyne statements. What these two extremes had in common, however, is that they both contributed to the conviction of innocent people.

²¹² Brandon L. Garrett & Greg Mitchell, *How Jurors Evaluate Fingerprint Evidence: The Relative Importance of Match Language, Method Information, and Error Acknowledgement*, 10 *Journal of Empirical Legal Studies* 484, 506 (2013).

XI. APPENDIX 1: METHODOLOGICAL APPENDIX ON USING THE NATIONAL REGISTRY OF EXONERATIONS DATA

A. CASE SELECTION

Although the NRE has an elaborate system of social science coding, it does not have any code which would readily allow one to obtain all cases involving MHCA. It does, however, code whether False or Misleading Forensic Evidence (F/MFE) contributed to the wrongful conviction. However, some NRE coders apparently considered testimony about mere “consistency” or “similarity” to not necessarily be “false or misleading”—presumably for the same reason the FBI review considered this testimony “appropriate.” Therefore, restricting our search to cases coded for F/MFE would have missed many cases.

Finding MHCA cases in the Registry, therefore, required keyword searches for “hair” in all Registry case summaries. Unfortunately, the word “hair” is quite common in NRE narrative summaries, and there are many reasons why hair might be relevant to a criminal case, other than MHCA. Most commonly, a witness might describe a perpetrator’s hair. In addition, because of limitations in the NRE’s software’s search functionality, it was not possible to constrict the search to eliminate false alarm words like “chair.”²¹³ In addition, in some cases, mitochondrial DNA analysis of hair, rather than MHCA, was done.

There are also some cases in which MHCA was used, but it was not described anywhere in the Registry’s coding. These cases escaped capture through keyword searching. We discovered a number of such cases (examples are [Omar Saunders](#) and co-defendants, [Clarence Richard Dexter, Jr.](#), and [Michael Morton](#)) adventitiously through other research projects while our MHCA research project was ongoing. Two additional cases ([Anthony Michael Green](#) and [Bennie Starks](#)) were discovered by cross-checking our data against the Innocence Project’s compilation of DNA exonerations involving MHCA.²¹⁴ There may be more such cases. Therefore, the numbers of cases in this Report may be undercounts and are unlikely to be overcounts.

B. SOURCES

As noted above, the FBI review used transcripts. The transcript therefore was our ideal; our goal was to determine what the expert witness actually said in court. The NRE, however, usually lacks transcripts. We used a variety of sources to try to get enough information to know or infer the exact words the expert used. We followed this procedure until we had enough information to code our variables of interest (testimony type, appropriate/erroneous, error type, commonly encountered phrases). In the interest of efficiency, once we felt confident that we could accurately

²¹³ It might have been possible to solve this problem using Excel.

²¹⁴ Penchina & Huffman, *Not a Strand of Evidence: DNA Exonerations Involving Unvalidated or Improper Microscopic Hair Comparison Evidence*.

code our variables of interest, we stopped pursuing sources that were more time-consuming to obtain and review. Therefore, the type of sources consulted varied from case to case.

We first consulted a text box used by Registry coders which asks the coder to “Describe Forensic Evidence.” If this did not give enough information to accurately code the case, we then consulted the NRE narrative summary, which is published online. If this did not give enough information to accurately code the case, we then consulted the NRE source documents. The NRE maintains a repository of Source Documents for every case it lists. As noted above, transcripts are rarely included in Source Documents. They primarily consist of media reports and legal documents, such as judicial rulings, briefs, and post-conviction pleadings and exhibits. Sometimes these documents helped explicate the words that the expert used. Legal documents, in particular, sometimes quoted the exact words that the expert used, thus constituting the sort of “transcript by proxy” [that the FBI review used in a small number of cases](#).

If these sources did not give enough information to accurately code the case, there were two additional sources of information we could consult for DNA exoneration cases only. [The Innocence Record](#) is a repository of legal materials pertaining to all DNA exonerations in the U.S. maintained by the Innocence Project and the law firm Winston and Strawn. Very often these materials contain transcripts. In addition, the materials are very well indexed, which often allowed us to go directly to the pages in which the MHCA evidence was discussed. This was an invaluable resource

Second, [Convicting the Innocent](#)²¹⁵ is another repository of materials pertaining to DNA exonerations maintained by Garrett to accompany his book.²¹⁶ This source sometimes contained legal rulings, transcripts, or summaries of forensic expert testimony. This was another invaluable resource.

For sources other than verbatim transcripts, where the source did not directly quote the expert, we sometimes faced the dilemma of determining whether the words used in our sources were the words used by the expert or a paraphrase. For example, if one of our sources said the expert said a hair from the crime scene “matched” the defendant’s hair, did that mean that the expert used the word “match,” or was that an attorney, judge, journalist, or coder’s shorthand for what the expert actually said? As we read more cases, some paraphrasing became more obvious. For example, while MHCA experts sometimes do use the word “match,” they rarely, if ever, use the word “link,” a word that is often used by attorneys and journalists. In cases in which such uncertainties made a difference to the coding, we consulted all available sources to try to resolve the issues. In some cases, though, we were unable to be certain whether the language was the expert’s language or a paraphrase. In such cases, we simply coded the language that we had. In the end, we were able to look at direct quotations from the expert witness in 68 (two thirds) of the 99 NRE non-FBI inclusion cases. In 28 of those cases, we had access to a transcript; in 40, we read another source, such as a judicial opinion, a legal brief, or a news story that provided us with a direct quotation.

²¹⁵ Garrett, *Convicting the Innocent: DNA Exonerations Database*, available at <https://convictingtheinnocent.com> (<https://perma.cc/FN5F-SKGZ>),

²¹⁶ Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*.

XII. APPENDIX 2: METHODOLOGICAL APPENDIX ON THE FBI REVIEW CODING SYSTEM

A. ERRONEOUS EVIDENCE

Our goal was to follow the FBI review’s coding methods as closely as possible. As mentioned above, the FBI review is not a conventional research study, and so a clear discussion of its methods has not been made publicly available. Just as we had to piece the findings together from multiple sources, so too did we have to piece the coding methods together from multiple sources. Below we try to discuss as clearly as possible our understanding of those coding methods.

The distinction between appropriate and erroneous testimony in the FBI study was established by an agreement between the FBI, the Innocence Project, and the NACDL on “the limits of science.” The agreed-upon definition of “appropriate” testimony was as follows:

The examiner’s testimony appropriately reflected the fact that hair comparison could not be used to make a positive identification, but that it could indicate, at the broad class level, that a contributor of a known sample could be included in a pool of people of unknown size, as a possible source of the hair evidence (without in any way giving probabilities, an opinion as to the likelihood or rareness of the positive association, or the size of the class) or that the contributor of a known sample could be excluded as a possible source of the hair evidence based on the known sample provided. An opinion as to the likelihood or rareness of a positive association may be appropriate in certain cases in which the examined hair samples display unusual or distinct characteristics, *e.g.*, repeated artificial treatments resulting in color variations along the length of the hair, hairs that have been crushed, broken, burned or damaged in some distinctive manner, or hairs that display specific characteristics associated with certain diseases such as pili annulate, monilethrix, or trichorrhesis nodosa.²¹⁷

The FBI Review Guidance provides examples of “appropriate” and “erroneous” testimony.²¹⁸ However, this document illustrates the difficulty of applying this definition in practice. Why, for example, is the first testimony below considered “erroneous” and second considered “appropriate”?

1. Now these hairs matched in every observable microscopic characteristic to that known hair sample of [Victim] and consistent with having originated from her.
2. Well, I go stronger than it was similar to it. It exhibited all the same microscopic characteristics as the known sample for [Mr. X.].

²¹⁷ FBI, *Microscopic Hair Comparison Analysis*.

²¹⁸ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*.

Is it the word “match”? The Review Guidance does not explain the difference between these two statements. Only information from individuals with inside knowledge of the review process provides the answer: The review process drew a distinction between statements of consistency with a hair sample and statements of consistency with a person.²¹⁹ The first statement is “erroneous” because of the words “with having originated from her,” whereas the second, “appropriate,” statement asserts consistency solely with the sample.

As we already know, “appropriate” testimony was uncommon in the FBI review. Therefore, statements of consistency with a sample were necessarily uncommon as well. However, the same was not true of our data. A great number of cases contained statements of consistency. In order to code consistently with the FBI Review Guidance, it was necessary for us to determine whether the statement of consistency was to a sample or a person. As noted above, we often began with a NRE coding note, a NRE narrative summary, a media account, or an appellate brief. For example, in the case of [Steven Linscott](#), the Describe Forensic Evidence box said “An expert testified that hair found at the scene was consistent with” the defendant’s. If the expert said “consistent with coming from the defendant,” the case would be considered “erroneous.” Only by examining the trial transcript, were we able to determine that the expert said “consistent with” the hair standards obtained from Linscott, rendering the testimony “appropriate.” The coding of many cases in our data set depended upon this fine distinction.

Certainly, one might question the salience of this distinction. From a forensic scientific perspective, what is the difference between testifying that a hair is consistent with a sample known to come from an individual and testifying that the same hair is consistent with coming from that individual? From a legal-psychological perspective, one might question whether the distinction would make any difference to a fact-finder. One might argue, therefore, that the FBI’s notion of “appropriate” testimony is overly broad. And yet, remarkably, FBI testimony was only 4% “appropriate” even with this possibly overbroad definition!

Kaye, however, makes the opposite argument: that the FBI’s definition of “*erroneous*” testimony is overly broad. He argues that statements of consistency—presumably regardless of whether they are to a sample or a person—should be considered scientifically appropriate. He acknowledges that these statements might be considered legally problematic “as creating an unacceptable risk that (in the absence of clarification on direct examination, cross-examination, or by judicial instruction) jurors will think the words connote a source attribution.” However, “arguments of this sort stray from determinations that an examiner has made statements that ‘exceed the limits of science’ to judgments that an examiner has made statements that are scientifically acceptable but prone to being misunderstood.” Therefore, he argues, “it is misleading to include scientifically acceptable but psychologically dangerous phrasing in the counts of scientifically erroneous statements.”²²⁰

B. ERROR TYPES

The FBI review reported that the erroneous testimony it found consisted of three specific types of error. It defined these three types as follows:

²¹⁹ Antoun et al., Hair Microscopy Review Webinar.

²²⁰ Kaye, *Ultrarepidarianism in Forensic Science: The Hair Evidence Debacle*.

Error Type 1: The examiner stated or implied that the evidentiary hair could be associated with a specific individual to the exclusion of all others. This type of testimony exceeds the limits of the science.

Error Type 2: The examiner assigned to the positive association a statistical weight or probability or provided a likelihood that the questioned hair originated from a particular source, or an opinion as to the likelihood or rareness of the positive association that could lead the jury to believe that valid statistical weight can be assigned to a microscopic hair association. This type of testimony exceeds the limits of the science.

Error Type 3: The examiner cites the number of hair analyses worked in the lab and the number of samples from different individuals that could not be distinguished from one another as a predictive value to bolster the conclusion that a hair belongs to a specific individual. This type of testimony exceeds the limits of the science.²²¹

Perhaps one reason for the great public interest in the FBI review is the way that these three error types illustrate in detail how expert witnesses exaggerate and distort the probative value of forensic evidence in actual criminal trials. The three error types offer clearly discrete ways of overstating the probative value of the evidence. Error Type 1 is essentially a form of “individualization”—the forensic claim to be able to reduce the potential donor pool of a trace to a single source—that has been widely criticized—and yet still used—in forensic science. Error Type 2 consists of concocting a faux probability. Defensible probabilities for MHCA do not exist,²²² but, as was already known from well-known exoneration cases, such as that of [Jimmy Ray Bromgard](#), hair analysts have been known to fabricate impressive-sounding statistics from thin air, and, in some cases even combine these baseless probabilities inappropriately (e.g., by multiplying probabilities that are not statistically independent, such as the supposed probability of coincidental “matches” of head and pubic hairs). Error Type 3 is, if possible, even more remarkable. Here the expert witness purports to construct a probability or error rate. The denominator is supposedly the number of hair analyses the expert has ever done. This number is most likely an estimate. The numerator is the number of hairs from different sources which could not be distinguished. But this number presupposes that the analyst knows ground truth in casework. In casework, the analyst does not know which hairs come from different sources. When an expert cannot distinguish two hairs, the expert does not know whether that is: (a) because the hairs come from the same source; or (b) because the hairs come from different sources but coincidentally appear consistent to the analyst. Put another way, the testimony assumes the conclusion that the expert witness *can* accurately determine when hairs come from the same source.

Reimer’s discussion of the FBI review offered manufactured examples of each type of testimony:

Error Type 1 Example:

A: I found brown, Caucasian head hairs on two items of clothing, the sports coat, and a pair of slacks that were reported to me as belonging to [the defendant]. Now, these hairs matched in every observable microscopic characteristic to that known hair sample of DEC

²²¹ FBI, *Microscopic Hair Comparison Analysis*.

²²² FBI, *Microscopic Hair Comparison Analysis: Standard for Review of Testimony and Laboratory Reports*.

[the decedent] and consistent with having originated from her. In my opinion, based on my experience in the laboratory and having done 16,000 hair examinations, my opinion is that those hairs came from DEC.

Error Type 2 Example:

Q: Now, based on your training and experience and your expertise in the field, and based on your knowledge of hair transfer and hair comparison, and based on the work done in this case, do you have an opinion, within the degree of scientific certainty, as to whether or not the pubic hair found in the underpants of [victim] came from [defendant]?

A: I would say that it would be a very high degree of probability that it does. Or to reverse it, I would say the chances of it being from somebody else, other than Mr. XX, would be highly unlikely at best.

Error Type 3 Example:

A: Now over the last 12 years, I personally have looked at hairs from about 10,000 different people, and over that time, I've only had two occasions out of the 10,000 people where I had hairs from two different people that I could not separate them.²²³

When we turn to the FBI Review Guidance, which was derived from actual, not manufactured, transcripts, however, things get murkier. There is a clear example of Error Type 1:

“Q: Do you have any doubt, Mr. ____, those two hairs we've been referring to in Item 34, in fact came from Defendant ____ head?

A: I have absolutely no doubt they're consistent with coming from his head.²²⁴

Less clear, however, is the following example of Error Type 1:

And the whole purpose of hair examinations is—or the whole key to it is that hairs from different people look different when all of these characteristics are examined and compared.

It would appear that the researchers interpreted this as a (very) implicit statement of individualization: “hairs from different people look different” could be read to mean “*all* hairs from different people look different” and thus imply that individualization is achievable for MHCA, in addition to being simply false. But the statement doesn't seem to directly implicate the defendant at all.

When it comes to Error Type 2, however, more questions are raised. Error Type 2, recall, was about faux probabilities, and one might expect to see bogus probability figures of the kind that were proffered in the notorious Bromgard case, for example. But the Review Guidance contains *no* statements of numerical probability.²²⁵ What probabilistic statements there are are verbal and

²²³ Reimer, *The Hair Microscopy Review Project: An Historic Breakthrough for Law Enforcement and A Daunting Challenge for the Defense Bar*.

²²⁴ FBI, *Microscopic Hair Comparison Analysis (MHCA) Review Lab Report/Transcript Review Guidance*.

²²⁵ *Id.*

vague: “very remote possibility,” “strong possibility,” “strong association.” Instead, the examples of Error Type 2 given in the Review Guidance are mostly “consistent with” statements, such as this:

The hair removed from the towel . . . exhibited the same microscopic characteristics as the known pubic hair sample from [redacted] and I concluded it was consistent with having originated from him.

One wonders why this testimony was considered Error Type 2 (which concerns the assignment of a probability), rather than Error Type 1 (which concerns implying that the hair comes from a specific individual)?

The answer—which is not publicly available, but can only be obtained through personal communications with people knowledgeable about the review—is that the FBI considers “consistent with” to connote a probability, whereas Error Type 1 is reserved for testimony that implies certainty or individualization.²²⁶ At first impression, it is surprising to see “consistent with” treated as a statement of probability. It is perhaps reasonable to interpret “consistent with” as connoting something less than certainty (which would be Error Type 1)—and hence a probability. However, it seems odd to place a “consistent with” statement in a category defined as “The examiner assigned to the positive association a statistical weight or probability or provided a likelihood that the questioned hair originated from a particular source, or an opinion as to the likelihood or rareness of the positive association . . .” As Kaye, who was also perplexed by the coding of “consistent with” statements as Error Type 2, noted about another “consistent with” statement, “The examiner had not ‘assigned to the positive association a statistical weight or probability.’ He had not ‘provided a likelihood that the questioned hair originated from a particular source.’ He had not expressed an ‘opinion as to the likelihood or rareness of the positive association.’”²²⁷

Error type 3 offered additional challenges. Through personal communication with people knowledgeable about the Review, we learned that Error type 3 required that both a denominator (e.g., number of cases worked by in the expert’s lifetime) and a numerator (e.g., purported number of cases in which hairs from different sources were indistinguishable) be stated—and stated numerically.²²⁸ If a numerator was not stated numerically, the statement would be coded as Error type 2. “Rare” was not considered a number, and so the following statement from an exoneration case ([Willie Grimes](#)) in our data set was coded as Error type 2:

in seven years of doing hair examination, it is rare that I see two individuals in the general population whose hair is the same under the microscope and has the same characteristics.

However, “never” was considered a number (zero), and so a statement using that word would be coded as Error type 3, as for example, in the evidence in the [Michael Jones](#) case, in which FBI MHCA expert Robert Fram

said he had never in his career been unable to distinguish between hair samples from two different people.

²²⁶ Antoun, Personal communication.

²²⁷ Kaye, *Ultrarepidarianism in Forensic Science: The Hair Evidence Debacle*.

²²⁸ Antoun, Personal communication.

For difficult coding decisions on both the above questions, we sought guidance from informants knowledgeable about the FBI review. These informants helped us understand “unwritten” coding rules that we could not get from publicly available documents. Where possible we followed these informants’ guidance on the coding specific cases or specific language.

XIII. APPENDIX 3: DATA TABLES

Table 18. Data table for Figure 5.

Year Convicted	Non-MHCA exonerations for 7 violent crimes (n=1896)	MHCA exonerations (n=125)
1956	1	0
1957	0	0
1958	0	0
1959	1	0
1960	0	0
1961	0	1
1962	0	0
1963	0	0
1964	1	0
1965	0	0
1966	3	0
1967	1	0
1968	8	0
1969	0	0
1970	0	0
1971	2	1
1972	4	1
1973	8	0
1974	6	0
1975	11	0
1976	14	1
1977	18	0
1978	8	6
1979	8	3
1980	21	4
1981	23	6
1982	40	8
1983	39	8
1984	59	2
1985	73	4
1986	53	17
1987	69	9
1988	54	11

Year Convicted	Non-MHCA exonerations for 7 violent crimes (n=1896)	MHCA exonerations (n=125)
1989	65	5
1990	81	11
1991	63	6
1992	89	2
1993	79	4
1994	88	2
1995	107	3
1996	90	2
1997	94	2
1998	81	2
1999	69	1
2000	58	0
2001	46	0
2002	61	1
2003	50	0
2004	39	1
2005	42	0
2006	46	0
2007	49	0
2008	34	0
2009	40	1
Total	1896	125

Table 19. Data table for Figure 6.

Year exonerated	Non-MHCA exonerations for 7 violent crimes	MHCA exonerations
1989	17	2
1990	28	0
1991	34	2
1992	27	4
1993	30	2
1994	26	2
1995	39	0
1996	35	7
1997	38	4
1998	28	2
1999	35	6
2000	62	2
2001	75	11
2002	50	9
2003	67	6
2004	45	2
2005	48	6
2006	49	4
2007	58	5
2008	57	3
2009	74	5
2010	63	4
2011	59	2
2012	83	4
2013	75	1
2014	84	2
2015	92	1
2016	90	2
2017	93	6
2018	97	7
2019	114	4
2020	91	3
2021	101	2
2022	115	2
2023	42	1
Total	2121	125

XIV. APPENDIX 4: ALL POST-1989 EXONERATION CASES IN WHICH MHCA IMPLICATED DEFENDANT

Cases are sorted first by Agency, with State/local listed above FBI; then by Evidence Type; then by State; County; Year of Exoneration; Year of Conviction; Last Name; and First Name. Multiple defendants prosecuted for the same crimes are shaded and color-coded.

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
1	Gauger	Gary	State/local	Damage	1996	1993	IL	McHenry	Murder
2	White	George	State/local	Inclusion	1992	1987	AL	Coffee	Murder
3	Mahan	Dale	State/local	Inclusion	1998	1986	AL	Jefferson	Sexual Assault
4	Mahan	Ronnie	State/local	Inclusion	1998	1986	AL	Jefferson	Sexual Assault
5	Flowers	Curtis	State/local	Inclusion	2010	2004	AR	Columbia	Murder
6	Newman	Rickey	State/local	Inclusion	2017	2002	AR	Crawford	Murder
7	Rose	Peter	State/local	Inclusion	2005	1995	CA	San Joaquin	Child Sex Abuse
8	Payne	Glenn	State/local	Inclusion	2018	1990	CA	Santa Clara	Child Sex Abuse
9	Reid	Mark	State/local	Inclusion	2003	1997	CT	Hartford	Sexual Assault
10	Dedge	Wilton	State/local	Inclusion	2004	1982	FL	Brevard	Sexual Assault
11	Crotzer	Alan	State/local	Inclusion	2006	1982	FL	Hillsborough	Sexual Assault
12	Baker	Dontrell	State/local	Inclusion	1997	1997	FL	Pinellas	Robbery
13	Nelson	Gary	State/local	Inclusion	1991	1980	GA	Chatham	Murder
14	White	John Jerome	State/local	Inclusion	2007	1980	GA	Meriwether	Sexual

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
									Assault
15	Dotson	Gary	State/local	Inclusion	1989	1979	IL	Cook	Sexual Assault
16	Linscott	Steven	State/local	Inclusion	1992	1982	IL	Cook	Murder
17	Adams	Kenneth	State/local	Inclusion	1996	1978	IL	Cook	Murder
18	Raingé	Willie	State/local	Inclusion	1996	1978	IL	Cook	Murder
19	Williams	Dennis	State/local	Inclusion	1996	1978	IL	Cook	Murder
20	Jimerson	Verneal	State/local	Inclusion	1996	1985	IL	Cook	Murder
21	Bradford	Marcellius	State/local	Inclusion	2001	1988	IL	Cook	Kidnapping
22	Ollins	Calvin	State/local	Inclusion	2001	1988	IL	Cook	Murder
23	Ollins	Larry	State/local	Inclusion	2001	1988	IL	Cook	Murder
24	Saunders	Omar	State/local	Inclusion	2001	1988	IL	Cook	Murder
25	Gray	Paula	State/local	Inclusion	2002	1978	IL	Cook	Murder
26	Tillman	Michael	State/local	Inclusion	2010	1986	IL	Cook	Murder
27	Hester	Lee Arthur	State/local	Inclusion	2019	1961	IL	Cook	Murder
28	Starks	Bennie	State/local	Inclusion	2013	1986	IL	Lake	Sexual Assault
29	Barnhouse	William	State/local	Inclusion	2017	1992	IN	Delaware	Sexual Assault
30	Pinkins	Darryl	State/local	Inclusion	2016	1991	IN	Lake	Sexual Assault
31	Glenn	Roosevelt	State/local	Inclusion	2017	1993	IN	Lake	Sexual Assault
32	Gregory	William	State/local	Inclusion	2000	1993	KY	Jefferson	Sexual Assault
33	Partin	Delmar	State/local	Inclusion	2019	1994	KY	Knox	Murder
34	Clark	Jeffrey	State/local	Inclusion	2018	1995	KY	Meade	Murder

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
35	Hardin	Garr Keith	State/local	Inclusion	2018	1995	KY	Meade	Murder
36	Charles	Clyde	State/local	Inclusion	1999	1982	LA	Terrebonne	Sexual Assault
37	Johnson	Anthony	State/local	Inclusion	2010	1986	LA	Washington	Murder
38	Velasquez	Eduardo	State/local	Inclusion	2001	1988	MA	Hampden	Sexual Assault
39	Sullivan	Michael	State/local	Inclusion	2019	1987	MA	Middlesex	Murder
40	Powell	Anthony	State/local	Inclusion	2004	1992	MA	Suffolk	Sexual Assault
41	Owens	James L.	State/local	Inclusion	2008	1988	MD	Baltimore City	Murder
42	Hatchett	Nathaniel	State/local	Inclusion	2008	1998	MI	Macomb	Sexual Assault
43	Watkins	Ledura	State/local	Inclusion	2017	1976	MI	Wayne	Murder
44	Dexter, Jr.	Clarence Richard	State/local	Inclusion	1999	1991	MO	Clay	Murder
45	Briscoe	Johnny	State/local	Inclusion	2006	1983	MO	St. Louis	Sexual Assault
46	Kordonowy	Paul D.	State/local	Inclusion	2003	1990	MT	Richland	Sexual Assault
47	Bauer	Chester	State/local	Inclusion	1997	1983	MT	Silver Bow	Sexual Assault
48	Bromgard	Jimmy Ray	State/local	Inclusion	2002	1987	MT	Yellowstone	Child Sex Abuse
49	Grimes	Willie	State/local	Inclusion	2012	1988	NC	Catawba	Sexual Assault
50	Bridges	Timothy	State/local	Inclusion	2016	1991	NC	Mecklenburg	Sexual Assault
51	Dail	Dwayne Allen	State/local	Inclusion	2007	1989	NC	Wayne	Child Sex Abuse

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
52	Peterson	Larry	State/local	Inclusion	2006	1989	NJ	Burlington	Murder
53	Halstead	Dennis	State/local	Inclusion	2005	1986	NY	Nassau	Murder
54	Kogut	John	State/local	Inclusion	2005	1986	NY	Nassau	Murder
55	Restivo	John	State/local	Inclusion	2005	1986	NY	Nassau	Murder
56	McCray	Antron	State/local	Inclusion	2002	1990	NY	New York	Sexual Assault
57	Richardson	Kevin	State/local	Inclusion	2002	1990	NY	New York	Sexual Assault
58	Salaam	Yusef	State/local	Inclusion	2002	1990	NY	New York	Sexual Assault
59	Santana	Raymond	State/local	Inclusion	2002	1990	NY	New York	Sexual Assault
60	Wise	Korey	State/local	Inclusion	2002	1990	NY	New York	Sexual Assault
61	Lopez	Steven	State/local	Inclusion	2022	1991	NY	New York	Robbery
62	Barnes	Steven	State/local	Inclusion	2009	1989	NY	Oneida	Murder
63	Broadwater	Anthony	State/local	Inclusion	2021	1982	NY	Onondaga	Sexual Assault
64	Green	Anthony Michael	State/local	Inclusion	2001	1988	OH	Cuyahoga	Sexual Assault
65	Towler	Raymond	State/local	Inclusion	2010	1981	OH	Cuyahoga	Child Sex Abuse
66	Campbell	William	State/local	Inclusion	2020	2009	OH	Hamilton	Manslaughter
67	Webb, III	Thomas	State/local	Inclusion	1996	1983	OK	Cleveland	Sexual Assault
68	Pierce	Jeffrey Todd	State/local	Inclusion	2001	1986	OK	Oklahoma	Sexual Assault
69	Bryson	David	State/local	Inclusion	2003	1983	OK	Oklahoma	Sexual Assault

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
70	McCarty	Curtis	State/local	Inclusion	2007	1986	OK	Oklahoma	Murder
71	Miller, Jr.	Robert Lee	State/local	Inclusion	2017	1988	OK	Oklahoma	Murder
72	Fritz	Dennis	State/local	Inclusion	1999	1988	OK	Pontotoc	Murder
73	Williamson	Ronald Keith	State/local	Inclusion	1999	1988	OK	Pontotoc	Murder
74	Scott	Calvin Lee	State/local	Inclusion	2003	1983	OK	Pontotoc	Sexual Assault
75	Durham	Timothy	State/local	Inclusion	1997	1993	OK	Tulsa	Child Sex Abuse
76	Courtney	Sedrick	State/local	Inclusion	2012	1996	OK	Tulsa	Robbery
77	Nelson	Bruce	State/local	Inclusion	1991	1982	PA	Allegheny	Murder
78	Whitley	Drew	State/local	Inclusion	2006	1989	PA	Allegheny	Murder
79	Brisson	Dale	State/local	Inclusion	1994	1991	PA	Chester	Sexual Assault
80	Nesmith	Willie James	State/local	Inclusion	2001	1982	PA	Cumberland	Sexual Assault
81	Smith	Jay C.	State/local	Inclusion	1992	1986	PA	Dauphin	Murder
82	Blair	Michael	State/local	Inclusion	2008	1994	TX	Collin	Murder
83	Rodriguez	George	State/local	Inclusion	2005	1987	TX	Harris	Child Sex Abuse
84	Sonnier	Ernest	State/local	Inclusion	2018	1986	TX	Harris	Kidnapping
85	Ochoa	Christopher	State/local	Inclusion	2002	1989	TX	Travis	Murder
86	Danziger	Richard	State/local	Inclusion	2002	1990	TX	Travis	Sexual Assault
87	Morton	Michael	State/local	Inclusion	2011	1987	TX	Williamson	Murder
88	Vasquez	David	State/local	Inclusion	1989	1985	VA	Arlington	Murder
89	Snyder	Walter	State/local	Inclusion	1993	1986	VA	Arlington	Sexual Assault

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
90	Stevens	Emerson	State/local	Inclusion	2021	1986	VA	Lancaster	Murder
91	Honaker	Edward	State/local	Inclusion	1994	1985	VA	Nelson	Sexual Assault
92	Cunningham	Calvin Wayne	State/local	Inclusion	2011	1981	VA	Newport News City	Sexual Assault
93	Davidson	Willie	State/local	Inclusion	2005	1981	VA	Norfolk City	Sexual Assault
94	Burnette	Victor	State/local	Inclusion	2009	1979	VA	Richmond City	Sexual Assault
95	Saecker	Frederic	State/local	Inclusion	1996	1990	WI	Buffalo	Sexual Assault
96	Hicks	Anthony	State/local	Inclusion	1997	1991	WI	Dane	Sexual Assault
97	Armstrong	Ralph	State/local	Inclusion	2009	1981	WI	Dane	Murder
98	Avery	Steven	State/local	Inclusion	2003	1985	WI	Manitowoc	Attempted Murder
99	Murphy	Thomas	State/local	Inclusion	2001	1999	WI	Wood	Sexual Assault
100	Woodall	Glen	State/local	Inclusion	1992	1987	WV	Cabell	Sexual Assault
101	Gray	David A.	State/local	Inconsistent but not excluded	1999	1978	IL	Madison	Attempted Murder
102	Alexander	Richard	State/local	Inconsistent but not excluded	2001	1998	IN	St. Joseph	Sexual Assault

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
103	Abdal	Warith Habib	State/local	Inconsistent but not excluded	1999	1983	NY	Erie	Sexual Assault
104	Waller	James	State/local	Inconsistent but not excluded	2007	1983	TX	Dallas	Child Sex Abuse
105	Cole	Timothy B.	State/local	Inconsistent but not excluded	2009	1986	TX	Lubbock	Sexual Assault
106	Holdren	Larry	State/local	Inconsistent but not excluded	2000	1984	WV	Kanawha	Sexual Assault
107	Gates	Donald Eugene	FBI	Inclusion	2009	1982	DC		Murder
108	Tribble	Santae	FBI	Inclusion	2012	1980	DC		Murder
109	Odom	Kirk	FBI	Inclusion	2012	1981	DC		Sexual Assault
110	Wright	Cleveland	FBI	Inclusion	2014	1979	DC		Murder
111	Martin	Kevin	FBI	Inclusion	2014	1984	DC		Manslaughter
112	Jones	Michael	FBI	Inclusion	2019	1996	DC		Sexual Assault
113	Butler	Dennis	FBI	Inclusion	2020	1971	DC		Murder
114	Ausby	John	FBI	Inclusion	2020	1972	DC		Murder
115	Nelson	Derrie	FBI	Inclusion	2022	1986	DC		Murder
116	Daniels	Elmer	FBI	Inclusion	2018	1980	DE	New Castle	Child Sex Abuse
117	Matta Ballesteros	Juan	FBI	Inclusion	2018	1990	F-	(Central)	Kidnapping

HAIR ANALYSIS AND WRONGFUL CONVICTION

No.	Last Name	First Name	Agency	Evidence Type	Year Exonerated	Year Convicted	ST	County	Crime
							CA		
118	Fain	Charles	FBI	Inclusion	2001	1983	ID	Canyon	Murder
119	Perrot	George	FBI	Inclusion	2017	1987	MA	Hampden	Sexual Assault
120	Huffington	John	FBI	Inclusion	2023	1981	MD	Harford	Murder
121	Sledge	Joseph	FBI	Inclusion	2015	1978	NC	Bladen	Murder
122	McCormick	Michael Lee	FBI	Inclusion	2007	1987	TN	Jefferson	Murder
123	Beranek	Richard	FBI	Inclusion	2018	1990	WI	Dane	Sexual Assault
124	Thomas	Marvin	FBI	Inclusion	1993	1987	WV	Wood	Murder
125	Holton	Rudolph	FBI	Race only inclusion	2003	1986	FL	Hillsborough	Murder
126 ²²⁹	Lewis	Carlton	State/local	Inclusion	2023	1990	NY	Onondaga	Murder

²²⁹ The Carlton Lewis case is not analyzed in this Report and is listed at the bottom of this Table for informational purposes. *See supra* note 108.

XV. APPENDIX 5. ALL PRE-1989 EXONERATION CASES IN WHICH MHCA IMPLICATED DEFENDANT

No.	Last name	First name	Agency	Evidence type	Year Exonerated	Year Convicted	County	ST	Crime
1	Peek	Anthony Ray	State/local	Inclusion	1987	1978	Polk	FL	Murder
2	Carden	Ronald	State/local	Inclusion	1982	1982	Pulaski	AR	Murder
3	Hall	James	Don't know	Inclusion	1984	1974	Johnson	IA	Murder