

2024

BAD FOR THE GANDER: NEW ZEALAND'S FAILURE TO LEARN FROM THE CAUTIONARY TALES OF UNRELIABLE FORENSIC-SCIENCE EVIDENCE AND WRONGFUL CONVICTIONS ABROAD

Carrie Leonetti

Follow this and additional works at: <https://digitalcommons.onu.edu/ilj>



Part of the [International Law Commons](#)

Recommended Citation

Leonetti, Carrie (2024) "BAD FOR THE GANDER: NEW ZEALAND'S FAILURE TO LEARN FROM THE CAUTIONARY TALES OF UNRELIABLE FORENSIC-SCIENCE EVIDENCE AND WRONGFUL CONVICTIONS ABROAD," *Ohio Northern University International Law Journal*: Vol. 2, Article 4. Available at: <https://digitalcommons.onu.edu/ilj/vol2/iss1/4>

This Article is brought to you for free and open access by the ONU Journals and Publications at DigitalCommons@ONU. It has been accepted for inclusion in Ohio Northern University International Law Journal by an authorized editor of DigitalCommons@ONU. For more information, please contact digitalcommons@onu.edu.

Bad for the Gander: New Zealand's Failure to Learn from the Cautionary Tales of Unreliable Forensic-Science Evidence and Wrongful Convictions Abroad

CARRIE LEONETTI

Abstract

Internationally, there is a robust, critical scientific and academic literature addressing concerns surrounding the reliability of forensic-science evidence in criminal trials and its relationship to wrongful convictions. Nonetheless, little attention has been directed at similar concerns in New Zealand, despite the existence of similar techniques, training, and testimony to those that have faced increasing scrutiny and criticism abroad. This Article argues that New Zealand courts should take more seriously their gatekeeping role in screening forensic-science evidence under section 25 of the Evidence Act 2006. It concludes that the lack of validation studies demonstrating the reliability of the forensic analyses routinely used in New Zealand threatens the accuracy of the criminal-justice system.

I. INTRODUCTION

Internationally, there is a robust, critical scientific and academic literature addressing concerns surrounding the reliability of forensic-science evidence in criminal trials and its relationship to wrongful convictions.¹ Nonetheless, little attention has been directed at similar concerns in New Zealand, despite the existence of similar techniques, training, and testimony to those that have faced increasing scrutiny and criticism abroad. Until now, there has been no

¹ See generally BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG (2011); Lisa Dufrainmont, *New Challenges for the Gatekeeper: The Evolving Law on Expert Evidence in Criminal Cases*, 58 CRIM. L. Q. 531 (2012); Gary Edmond, *(Ad)ministering Justice: Expert Evidence and the Professional Responsibilities of Prosecutors*, 36 UNIV.N.S.W. L. J. 921 (2013); Gary Edmond, *Specialised Knowledge, the Exclusionary Discretions and Reliability: Reassessing Incriminating Expert Opinion Evidence*, 31 UNIV. N.S.W. L. J. 1 (2008); Gary Edmond, *The "Science" of Miscarriages of Justice*, 37 UNIV. N.S.W. L. J. 376 (2014); Keith A. Findley, *Flawed Science and the New Wave of Innocents*, in WRONGFUL CONVICTIONS AND THE DNA REVOLUTION: TWENTY-FIVE YEARS OF FREEING THE INNOCENT 184 (Daniel S. Medwed, ed., 2017); See also Jessica D. Gabel, *Realizing Reliability in Forensic Science from the Ground up*, 104 J. CRIM. L. & CRIMINOLOGY 283, 286 (2014) ("Unreliable science presents itself in a virtual smorgasbord of ways, from the routine (contamination) to the egregious (forensic misconduct) and everything in between (misrepresented or exaggerated results, misinterpretation of results, lack of research for basic assumptions, unqualified analysts, inconsistent lab practices.)"); See Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009); Jon B. Gould & Richard A. Leo, *One Hundred Years Later: Wrongful Convictions after a Century of Research*, 100 J. CRIM. L. & CRIMINOLOGY 825 (2010); Maryam Hamirani, et al., *Examining the Role of Science in the Courtroom: Admissibility and Reliability of Forensic Science in the Courtroom*, 81 ALBANY L. REV. 975 (2017); Jonathan J. Koehler, *Forensics or Fauxrensics? Ascertaining Accuracy in the Forensic Sciences*, 49 ARIZ. ST. L. J. 1369 (2018); Kevin C. McMunigal, *Prosecutors and Corrupt Science*, 36 HOFSTRA L. REV. 437 (2007); Jennifer L. Mnookin, et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725 (2011); Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721 (2007); D. Michael Risinger & Michael J. Saks, *A House with No Foundation*, 20 ISSUES IN SCI. & TECH. 35 (2003); Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892 (2005); Deborah Tuerkheimer, *Science-Dependent Prosecution and the Problem of Epistemic Contingency*, 62 ALA. L. REV. 513 (2011).

systematic scholarly examination of the use of forensic-science evidence in criminal trials within New Zealand.

This Article argues that New Zealand courts should take more seriously their gatekeeping role in screening forensic-science evidence under section 25 of the Evidence Act 2006. Section II traces the history and development of the growing international consensus around the profound reliability concerns with many common forensic-science disciplines, techniques, and interpretations. Section III outlines the two largest ongoing areas of concern: the way that forensic-science conclusions are communicated to fact finders and the under-validation of many forensic-science claims. Section IV connects those critiques with forensic-science practices in New Zealand, documenting the continued, uncritical acceptance by courts of techniques increasingly coming under scrutiny in other parts of the world. Section V concludes that the lack of any validation studies demonstrating the reliability of the forensic analyses routinely used in New Zealand threatens the accuracy of the criminal-justice system.

II. JUNK SCIENCE IN INTERNATIONAL CONTEXT

A. History

The traditional forensic sciences are forensic but not necessarily scientific. The oldest ones (e.g., fingerprint comparison) are more than one hundred years old, and the forensic sciences generally were developed by law-enforcement agencies solely for the purpose of being used in criminal investigations.² Traditionally, most forensic “scientists” had no formal university training in any scientific discipline, instead learning through on-the-job training from previous generations of forensic examiners through a guild-like indoctrination process.³ While many forensic techniques have a field of science with which they can be affiliated (for example, both forensic firearm examinations and fire-pattern analysis trace some of their fundamentals to physics – ballistics and pyrolysis, respectively), scientific researchers were not involved in the development of either field.⁴ On the contrary, if anything, formally trained scientists tend to be critical of, rather than claim any credit for, the forensic “sciences.”⁵

The primary exception to this history is forensic DNA analysis. Both the testing techniques and the methods for interpreting the probabilistic nature of test results for forensic DNA comparison were designed with significant input from biochemists.⁶ While the older forensic sciences (e.g., fingerprints, fired-bullet comparison) have been admitted as prosecution evidence in criminal trials for decades with little to no critical examination as a “novel” science,

² See Carrie Leonetti, *The History of Forensic-Science Evidence in Criminal Trials and the Role of Early ‘Success’ in Establishing Its Putative Reliability*, 54 ST. MARY’S L.J. 1062(2023).

³ *Id.* at 1063.

⁴ *Id.* at 1062-1063.

⁵ Carrie Leonetti, *Endangered by Junk Science: How the New Zealand Family Court’s Admission of Unreliable Expert Evidence Places Children at Risk*, 43 CHILDREN’S L. RTS. J. 17, 65 (2022) [hereinafter “Leonetti, *Junk Science*”].

⁶ David L. Faigman, *Anecdotal Forensics, Phrenology, and Other Abject Lessons from the History of Science*, 59 HASTINGS L.J. 979 (2008).

forensic DNA comparison was put through a screening crucible during the 1990s.⁷ As a result, proponents of expert DNA testimony had to ensure that the testing techniques were reliable and the underlying scientific premises were fundamentally valid.⁸ While the results of forensic DNA analysis are largely admitted in evidence without challenge today, back when DNA was undergoing stringent vetting, prosecutors routinely called expert witnesses with advance degrees in biology and chemistry to lay the foundation for the reliability of the tests and the interpretation of test results to which case examiners would testify.⁹

By comparison, with many other forensic sciences, scientific experts testify for the defense, often engaging in trenchant critiques of the unfounded claims of prosecution witnesses testifying that the defendant was the source of some item of crime-scene evidence. Because the forensic sciences have not been developed by scientists in a culture of science, the prosecution and defense experts tend to talk across one another, with the prosecution focusing on the performance and results of the tests and the defense focusing on the lack of fundamental validation of those claims.¹⁰ Courts have traditionally admitted forensic-science evidence, and left juries to determine what weight to give it, on the theory that the on-the-job training of the forensic examiner is an equally weighty qualification as the academic qualifications of the defense expert. While examining thousands of sets of fingerprints may qualify a fingerprint examiner to testify as to whether two fingerprints “match” one another, scholars have long criticized courts for allowing fingerprint examiners without scientific training and expertise to testify that their match determination is scientifically valid and reliable.

B. Sea Change: The International Movement Challenging Traditional Forensic Sciences

1. The Role of DNA Exonerations

The first 276 DNA exonerations in the United States have been extensively mined for patterns of error and systemic lessons.¹¹ One clear lesson from these exonerations has been the lack of reliability and testing of the forensic sciences and their uncritical acceptance by criminal

⁷ R. v. Pantoja, (1996) 88 A. Crim. R. 554 (N.S.W.); See generally R. v. Lucas, [1992] 2 V.R. 109; R. v. Tran, (1990) 50 A.C.R. 233, 242.

⁸ Executive Office of the President, PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY (“PCAST”), FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS, 2 (2016) [hereinafter “PCAST Report”].

⁹ *Id.*

¹⁰ In a common tableau, prosecutors will often try to undercut the credibility of the academic critic by pointing out that s/he not trained in the forensic technique in question. See Gary Edmond, *et al.*, *Contextual Bias and Cross-Contamination in the Forensic Sciences: the Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals*, 14 L., PROBABILITY & RISK 1, 19 (2014). An apt analogy would be if the prosecution called an astrologist who used tarot cards to identify the defendant as the perpetrator of the charged crime, who claimed that his/her extensive experience with using tarot cards had never failed in the past and that s/he had been to many tarot trainings where other astrologists confirmed claims of infallibility, the defense called a statistician to testify to the faulty mathematics and statistical inferences behind tarot’s claims of infallibility, and the prosecution argued to the court that it should disregard the defense critique because the statistician had no training in tarot.

¹¹ Carrie Leonetti, *The Innocence Checklist*, 58 AMER. CRIM. L. REV. 97, 102 (2021) [hereinafter “Leonetti, *Innocence Checklist*”].

courts. According to the Innocence Project, more than half of the wrongful convictions that they overturned in the United States were caused by faulty forensic-science evidence.¹²

For example, in 1978, Verneal Jimerson and Dennis Williams were convicted of the rape and murder of Carol Schmal and Lawrence Lionberg.¹³ The evidence against them included serology and microscopic hair comparison evidence. At trial, a forensic scientist from the Illinois Bureau of Scientific Services testified that Jimerson, who had type-O blood, could not be excluded as a contributor of biological evidence recovered in Schmal's vaginal swab, although he conceded that 47% of the population had type-O blood.¹⁴ This testimony was misleading because while 47% of the population have type-O blood, more than 80% of the population would have been ineligible for exclusion as the source of the semen because the victim also had type-O blood, so the semen also could have come from a non-secretor with another blood type.¹⁵ The forensic scientist also testified that hairs recovered from the back seat and trunk of Williams's car were microscopically indistinguishable from Lionberg's hair.¹⁶ In 1995, the Illinois Supreme Court quashed Jimerson's conviction after it emerged that the State had offered perjured testimony from a key witness at his trial.¹⁷ In 1996, DNA testing conclusively eliminated Jimerson and Williams as the perpetrators and established the identities of the real perpetrators.¹⁸ Jimerson and Williams's subsequent civil-rights lawsuit against Cook County, Illinois resulted in their being awarded millions of dollars in damages.¹⁹

In 1989, John Galvan was convicted of arson and murder after a residential fire killed several of his neighbors in Chicago.²⁰ The State's theory was that Galvan started the fire by throwing a Molotov cocktail at the building.²¹ At Galvan's trial, a detective from the Chicago Police arson section testified that the burn patterns left by the fire indicated the use of a liquid accelerant and an intentional ignition source rather than an accidental fire, even though chemical tests on the debris did not indicate the presence of an accelerant.²² In 2012, the Illinois Court of Appeals granted Galvan a post-conviction review hearing. At the hearing, an expert fire investigator testified that there was no scientific basis to support the conclusion that the fire was intentionally ignited or that an accelerant was used.²³ He explained that the detective's original fire-pattern analysis had been repudiated by modern fire science.²⁴ The trial court nonetheless

¹² INNOCENCE PROJECT, *Misapplication of Forensic Science*, <https://innocenceproject.org/misapplication-of-forensic-science/> (last visited September 30, 2023).

¹³ Maurice Possley, *Verneal Jimerson*, NATIONAL REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3439>.

¹⁴ *Id.*

¹⁵ *Id.*

¹⁶ *Id.*

¹⁷ Possley, *supra* note 14.

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ Maurice Possley, *John Galvan*, NATIONAL REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/pages/casedetail.aspx?caseid=6359>.

²¹ *Id.*

²² *Id.*

²³ *Id.*

²⁴ Possley, *supra* note 21.

upheld Galvan’s conviction.²⁵ In 2019, the Illinois Court of Appeals quashed his conviction and ordered a new trial.²⁶ In 2022, the State ultimately dismissed the charges against Galvan after he had spent thirty-five years in prison.²⁷

In 1994, Patrick Pursley was convicted of a murder that occurred during an armed robbery and sentenced to life imprisonment.²⁸ At Pursley’s trial, a crime lab analyst from the Illinois State Police crime lab testified that the bullets recovered from the victim matched bullets test-fired from a gun associated with Pursley and that Pursley’s gun was the murder weapon “to the exclusion of all other” guns.²⁹ Pursley offered evidence from a firearms expert who testified that he was unable to conclusively match or eliminate bullets fired from Pursley’s gun to the crime-scene bullets because the crime-scene bullets had insufficient markings on them to make a match determination.³⁰ In 2011, the Illinois Court of Appeals granted Pursley’s application to have the crime-scene bullets run through the Integrated Ballistics Identification System (“IBIS”) under a new Illinois statute that authorizes post-conviction IBIS searches for prisoners who have a reasonable likelihood of being exonerated by the searches.³¹ After the IBIS search, the State’s expert conceded that his examination of the bullets fired from Pursley’s gun was inconclusive.³² Two new defense experts also testified that their analysis of Pursley’s gun excluded it as the murder weapon.³³ In 2017, the Winnebago County Circuit Court quashed Pursley’s conviction and ultimately found Pursley not guilty after retrial, concluding that the new firearms evidence eliminated Pursley’s gun as the murder weapon.³⁴ The court granted Pursley a certificate of innocence, and the State awarded him almost \$500,000 in compensation.³⁵

Other countries have also experienced forensic-science failures resulting in wrongful convictions. For example, in England, Simon Hall was convicted of murder on the basis of a fibre “match” between a pair of Hall’s pants and trace evidence from the crime scene and later exonerated.³⁶ In Australia, Raymond Carroll was convicted of murder in 1985 on the basis of the bite-mark comparison testimony of three forensic odontologists, and his conviction was later quashed after bite-mark comparison evidence was discredited.³⁷

2. Scientific Criticism

²⁵ *Id.*

²⁶ *Id.*

²⁷ *Id.*

²⁸ Maurice Possley, Patrick Pursley, NATIONAL REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=5487>.

²⁹ *Id.*

³⁰ *Id.*

³¹ *Id.*

³² Possley, *supra* note 29.

³³ *Id.*

³⁴ *Id.*

³⁵ *Id.*

³⁶ R. v. Simon John Hall, [2011] E.W.C.A. Crim. 4. (Eng.).

³⁷ Liam Mannix, *Top Judge Worried Forensic Evidence Putting Innocent People Behind Bars*, SYDNEY MORNING HERALD, (Sept. 1, 2019, 11:59 PM), <https://www.smh.com.au/national/top-judge-worried-forensic-evidence-putting-innocent-people-behind-bars-20190823-p52k3l.html>.

As DNA analysis began to demonstrate the flaws in older forensic sciences, forensic laboratories became a sudden area of interest for governmental commissions. The watershed event giving rise to the international explosion of scientifically and scholarly critique of traditional forensic sciences was the 2009 publication of the scathing report, *Strengthening Forensic Science in the United States: a Path Forward*, by the National Research Council of the American National Academy of Sciences (“the NAS Report”).³⁸ The NAS Report was directed at courts and criminal-justice practitioners and was critical of all of the traditional forensic sciences whose evidence is routinely in use in criminal trials in the United States and around the world, other than single-source forensic DNA analysis.³⁹ It was also critical of the absence of controlling standards, validation studies, and controls to eliminate the contamination of forensic analyst’s subjective decision making by contextually biasing information.⁴⁰ It was also critical of the way that forensic analysts communicated their results in testimony, employing terminology that was inconsistent and misleading (“inclusion,” “match,” “identification,” “individualization,” “to the exclusion of all others”) and inflating interpretive claims beyond the limits of what scientific validation did exist.⁴¹

Despite its revolutionary criticism of several previously unexamined and uncritically accepted forensic sciences, little legal reform followed the issuance of the NAS Report. The NAS Report was aimed at the practitioners of forensic science, who reacted to the report defensively rather than proactively. As a result, then-President of the United States Barack Obama asked his Presidential Council of Advisors of Science and Technology (“PCAST”) to explore what contributions the legal system could make in regulating the forensic sciences. Seven years later, PCAST issued its report (the “PCAST Report”), which focused specifically on the pattern-matching forensic disciplines and reiterated many of the same criticisms as the NAS Report.⁴² The PCAST report found that bite mark, footwear, and hair comparisons lacked foundational validity and were unable ever to achieve applied validity.⁴³ It found that firearms analysis and latent fingerprint analysis might have foundational validity but had not yet established applied validity, particularly in light of their high rates of false positives and vulnerability to cognitive biases.⁴⁴ The PCAST Report concluded: “Where there are not adequate empirical studies and/or statistical models to provide meaningful information about the accuracy

³⁸ NATIONAL RESEARCH COUNCIL, *STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD* (2009) [hereinafter “NAS Report”].

³⁹ *Id.* at 7 (“With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.”).

⁴⁰ *Id.* at 6 (“Often there are no standard protocols governing forensic practice in a given discipline. And, even when protocols are in place . . . they often are vague and not enforced in any meaningful way.”); *Id.* at 189 (“Little rigorous systematic research has been done to validate the basic premises and techniques in a number of forensic science disciplines.”).

⁴¹ *Id.* at 185-86.

⁴² See PCAST Report, *supra* note 9.

⁴³ See *id.* at 7.

⁴⁴ See *infra* Section II (C).

of a forensic feature-comparison method, [the prosecution and its experts] should not offer testimony based on the method.”⁴⁵

In 2021, the UK Parliament enacted the Forensic Science Regulator Act, which codified the post of the Forensic Science Regulator for England and Wales and obligated the Regulator to promulgate a Code of Practice for forensic scientists there.⁴⁶ In March 2023, the Regulator issued the Code, which came into effect in October 2023.⁴⁷ The Code contains Standards of Practice, which require forensic scientists: to recognise that their overriding duty is to the court and the administration of justice, to provide expert advice and evidence only within the limits of their professional competence, conduct casework using demonstrably valid methods, and report potential miscarriages of justice.⁴⁸

3. Practitioner Responses

One of the most disheartening features of both the NAS Report and the PCAST Report was their reception in the law-enforcement, forensic-science, and prosecution communities. More than a decade on, few of the recommendations of the NAS Report have been implemented.⁴⁹ Despite the fact that both NAS and PCAST are preeminent American scientific organizations with marquis members, the majority of professional forensic-science membership organizations, law-enforcement agencies, and prosecutorial professional organizations responded to the issuance of the reports with disdain or defensiveness.⁵⁰ The American Federal Bureau of Investigation (“FBI”) responded to the PCAST report by announcing that “the FBI disagrees with many of the scientific assertions and conclusions of the report. The report makes broad, unsupported assertions regarding science and forensic science practice.”⁵¹

The National District Attorneys’ Association (“NDAA”), the umbrella group for American prosecutors, responded:

The NDAA has significant disagreement with the logic and manner in which the PCAST report portrayed the pattern analysis and comparison disciplines of bitmarks, firearms/toolmarks, tire tread, shoe prints and DNA complex mixtures.

* * *

The PCAST position is that the forensic science disciplines specializing in the examination of bitmarks, firearms/toolmarks, complex DNA mixtures, tire-treads, and shoe prints each lack scientific foundational support and should not be permitted for use in the criminal courtroom.

⁴⁵ PCAST Report, *supra* note 9, at 19.

⁴⁶ Forensic Science Regulator Act 2021 (SI 2022/856) (U.K.), at Chapter 14, §§ 1-2.

⁴⁷ See Forensic Science Regulator, Code of Practice, March 2023.

⁴⁸ *Id.* at 12.1.2.

⁴⁹ Liliana Segura & Jordan Smith, *VIVA 4N6*, THE INTERCEPT, (March 26, 2016, 11:25 AM), <https://theintercept.com/2016/03/25/in-las-vegas-embattled-forensics-experts-respond-to-scandals-and-flawed-convictions/>.

⁵⁰ *Id.* (describing the response of forensic odontologists to the NAS Report as “downright aggressive”).

⁵¹ FED. BUREAU OF INVESTIGATION, COMMENTS ON: PRESIDENT’S COUNSEL OF ADVISORS ON SCIENCE AND TECHNOLOGY, REPORT TO THE PRESIDENT, FORENSIC SCIENCE IN FEDERAL CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF PATTERN COMPARISON METHODS, (2016).

However, the opinions expressed by PCAST in their report clearly and obviously disregard large bodies of scientific evidence to the contrary and rely, at times, on unreliable and discredited research.

* * *

Prosecutors ultimately serve the criminal justice system and the public at large. Forensic evidence and forensic experts are called to testify by prosecutors and members of the defense bar every day, both to exonerate the innocent and prosecute the guilty. Judges have the ultimate authority to determine whether or not a forensic discipline is valid or lacks the requisite foundation, not the PCAST report.⁵²

While the Scientific Working Group for Friction Ridge Analysis, Study, and Technology (“SWGFAST”)⁵³ did not issue a formal response to either the NAS Report or the PCAST report, it did amend its guidelines for fingerprint identifications after the issuance of the NAS Report. The 2002 SWGFAST guidelines defined a match determination as follows: “Individualization occurs when a latent print examiner, trained to competency, determines that two friction ridge impressions originated from the same source, to the exclusion of all others.”⁵⁴ The 2012 guidelines, which are still in effect, now advise:

Individualization is the decision by an examiner that there are sufficient features in agreement to conclude that two areas of friction ridge impressions originated from the same source. Individualization of an impression to one source is the decision that the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility.⁵⁵

While the newer language may seem, intuitively, to be an improvement – particularly the removal of the claim that a fingerprint match is made “to the exclusion of all others,” it is still fundamentally a semantic dodge. There is no more scientific basis for the claim that the likelihood of a random (but innocent) fingerprint match determination is “so remote” as to be “a practical impossibility” than there was for the claim that the possibility of a random match was zero. On the contrary, by definition, the “likelihood” of a random match determination is an

⁵² Testimony of Kay Chopard, Nat’l Dist. Attorneys Ass’n, before the President’s Council of Advisors on Science and Technology 1-3 (Sept. 30, 2016), <https://www.crime-scene-investigator.net/PDF/national-district-attorneys-association-response-to-the-report-forensic-science-in-criminal-courts-ensuring-scientific-validity-of-feature-comparison-methods.pdf>.

⁵³ The Scientific Working Groups (SWGs) are consortia of forensic-science subject-matter experts who generate best practices standards for various areas in the forensic sciences. See NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, *Legacy Scientific Working Groups*, available at: [https://www.nist.gov/forensic-science/interdisciplinary-topics/working-groups/legacy-scientific-working-groups#:~:text=Scientific%20Working%20Groups%20\(SWGs\)%20consist,particular%20area%20of%20forensic%20science](https://www.nist.gov/forensic-science/interdisciplinary-topics/working-groups/legacy-scientific-working-groups#:~:text=Scientific%20Working%20Groups%20(SWGs)%20consist,particular%20area%20of%20forensic%20science). (last visited June 6, 2020). SWGFAST, which was founded in 1998, was the SWG for fingerprint analysis, but it was disbanded in 2015.

⁵⁴ U.S. DEP’T OF J., A REVIEW OF THE FBI’S HANDLING OF THE BRANDON MAYFIELD CASE, 312 (2006).

⁵⁵ SWGFAST, *Standards for Examining Friction Ridge Impressions and Resulting Conclusions*, § 5.3.2.2 (2012). See generally *Id.* at § 4.3.2.2; SWGFAST, Position Clarification, June 29, 2010 (on file with Author).

inherently probabilistic concept, which, if valid, should be representable with a number (in this case, a very small fraction).

Similarly, in Australia, critics have pointed out the complete lack of effect that the NAS and PCAST reports had on the admissibility and communication of scientific evidence in criminal trials.⁵⁶ Instead, the Australia High Court loosened the admissibility rules for scientific evidence in 2016.⁵⁷ When asked about the PCAST report, specifically, the Victoria Attorney-General responded that existing laws were sufficient to “govern the reliability of forensic science in the state's courts.”⁵⁸ The director of the Australian National Institute of Forensic Science responded:

The United States PCAST report... is three years old and considerable work has been, and continues to be, undertaken to address findings contained in the report. The US PCAST findings have been raised in criminal trials in Australia and no issues were identified. The evidence provided by government forensic services providers is robust and accounts for limitations in the assessment and interpretation they perform, and in the reports they prepare.⁵⁹

4. *Scholars and Law Commissions*

Internationally, a scholarly consensus has emerged around the widespread intellectual complacency and reactionary inertia of criminal courts and lawyers, over-reliance on adversarial testing in lieu of scientific testing, and related lack of critical oversight of forensic-science evidence in criminal trials, particularly in light of the role that unreliable forensic-science evidence has played in wrongful convictions.⁶⁰ As the Law Commission for England and Wales concluded shortly after the NAS Report was released in the United States: “Cross-examination, the adduction of contrary expert evidence and judicial guidance at the end of the trial are currently assumed to provide sufficient safeguards in relation to expert evidence . . . [I]t is doubtful whether these are valid assumptions.”⁶¹ This critical dialogue, in conjunction with the role that exculpatory DNA testing played in the exoneration of hundreds of individuals

⁵⁶ Mannix, *supra* note 38 (noting that the NAS and PCAST reports resulted in no reforms in Australia and no “appeals based on questions of admissibility of forensic evidence”).

⁵⁷ *Id.*

⁵⁸ *Id.*

⁵⁹ *Id.*

⁶⁰ See generally DAVID L. FAIGMAN, *LABORATORY OF JUSTICE: THE SUPREME COURT’S 200-YEAR STRUGGLE TO INTEGRATE SCIENCE AND THE LAW* (2004); DAVID L. FAIGMAN, *LEGAL ALCHEMY: THE USE AND MISUSE OF SCIENCE IN THE LAW* (1999); See also David S. Caudill, *Lawyers Judging Experts: Oversimplifying Science and Undervaluing Advocacy to Construct an Ethical Duty?*, 38 PEPP. L. REV. 675 (2011); David L. Faigman, *Anecdotal Forensics, Phrenology, and Other Abject Lessons from the History of Science*, 59 U.C. HASTINGS L. J. 979 (2008); Brandon L. Garrett & Gregory Mitchell, *Forensics and Fallibility: Comparing the Views of Lawyers and Jurors*, 119 W. VA. L. REV. 621 (2016); Jennifer L. Groscup, *et al.*, *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL., PUB. POL. & L. 339 (2002); Jonathan J. Koehler & John B. Meixner, *An Empirical Research Agenda for the Forensic Sciences*, 106 J. CRIM. L. & CRIMINOLOGY 1 (2016); See generally Leonetti, *Innocence Checklist*, *supra* note 12.

⁶¹ Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales*, Law Com No 325, ¶ 1.20 (2009).

previously convicted of serious crimes on the basis of faulty scientific evidence, has led to the debunking of several pseudo-scientific disciplines, as well as the ongoing questioning of and challenges to both traditional and novel forensic-science techniques.⁶²

In Australia, a consortium of scholars has launched the Evidence-Based Forensic Initiative in Wollongong, New South Wales (“EBFI”), an interdisciplinary group of forensic scientists, cognitive scientists, and legal academics engaging in the evaluation and improvement of the forensic sciences.⁶³ According to EBFI’s mission statement:

Many forensic science procedures in regular use were developed by non-scientists, are devoid of underlying research and have never been formally evaluated. Even procedures with mainstream scientific foundations are often practiced in ways that are insufficiently attentive to the constraints of validation studies, standards, the effects of human factors and other limitations. In consequence, investigators and courts in Australia and around the world routinely rely upon forensic science evidence that is either of unknown probative value or less reliable than it could be. Moreover, these serious limitations are not necessarily disclosed, identified or understood by those presenting, evaluating and reviewing forensic science evidence in criminal trials and appeals.⁶⁴

C. Debunked Disciplines

The innocence movement and the thousands of exonerations of innocent prisoners that it has freed internationally has led to scholarly, scientific, and judicial recognition of the invalidity of several common forensic-science disciplines, particularly pattern-matching disciplines like fire-pattern analysis,⁶⁵ microscopic hair comparison,⁶⁶ and bite mark pattern matching,⁶⁷ as well

⁶² Leonetti, *Innocence Checklist*, *supra* note 12 at 101-02, and sources cited therein (documenting the scholarly consensus that “junk science” is a primary cause of wrongful convictions and that certain forensic sciences – including fire-pattern analysis, microscopic hair comparison, bitemark pattern matching, bloodstain pattern analysis, and handwriting comparison – are unreliable and should no longer be used in criminal trials).

⁶³ E.B.F.I., <http://evidencebasedforensics.com/> (last visited November 6, 2019).

⁶⁴ *Id.*

⁶⁵ See Parisa Denghani-Tafti & Paul Bieber, *Folklore and Forensics: The Challenges of Arson Investigation and Innocence Claims*, 119 W. VA. L. REV. 549, 557 (2016); see generally Christina Carrega & Stephen Rex Brown, *Three Men Convicted in Fatal 1980 Park Slope Arson Fire Cleared After Decades Behind Bars*, N.Y. DAILY NEWS (Dec. 17, 2015, 12:31 AM, updated Apr. 9, 2018, 5:11 AM) : <https://www.nydailynews.com/2015/12/17/three-men-convicted-in-fatal-1980-park-slope-arson-fire-cleared-after-decades-behind-bars/> (last visited February 23, 2024); David Grann, *Trial by Fire: Did Texas Execute an Innocent Man?*, NEW YORKER (Sep. 7, 2009): <https://www.newyorker.com/magazine/2009/09/07/trial-by-fire> (last visited February 23, 2024).

⁶⁶ Spencer S. Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Dept.*, WASH. POST. (Apr. 16, 2012, 9:52 PM), https://www.washingtonpost.com/local/crime/convicted-defendants-left-uninformed-of-forensic-flaws-found-by-justice-dept/2012/04/16/gIQAWTcgMT_story.html (detailing the tortuous delays in the debunking of the Federal Bureau of Investigation’s unreliable microscopic hair comparison program); see also Segura & Smith, *supra* note 49.

⁶⁷ NAS Report, *supra* note 39, at 174-76; PCAST Report, *supra* note 9, at 3 (“PCAST considers the prospects of developing bitemark analysis into a scientifically valid method to be low.”); DAVID L. FAIGMAN, *et al.*, MODERN SCIENTIFIC EVIDENCE (2023), at §§ 35:16; Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bitemark Evidence*, 30 CARDOZO L. REV. 1369, 1380 (2009); Jason M. Chin & D’Arcy White, *Forensic Bitemark Identification Evidence in Canada*, 52 U.B.C. L. REV. 57 (2019); Paul C. Giannelli, *Bite Mark Analysis*, 43 CRIM. L.

as causation testimony relating to child injuries and “shaken baby syndrome.”⁶⁸ One of the biggest forensic-science debacles internationally involved testimony relating to the FBI’s composition of bullet lead analysis technique (“CBLA”), which purported to allow a chemical signature match of the composition of lead bullets such that a firearms examiner could testify that an evidentiary bullet (e.g., one retrieved from a crime scene) had to have come from the same manufacturing batch or even box of ammunition as bullets linked to a suspect (e.g., a box of ammunition seized from a suspect’s home).⁶⁹

Many traditional forensic disciplines involve the same feature comparison and pattern matching as the debunked pattern-matching disciplines. Fingerprint comparison, handwriting comparison, and firearms (fired-bullet and tool-mark) comparison are all pattern-matching disciplines with methodological similarities to fire-pattern analysis, microscopic hair comparison, and bite-mark comparison. What defines the pattern-matching disciplines is that the features or characteristics of two items are compared to one another – an evidentiary sample (the fired bullet from the crime scene, the facial features of the perpetrator caught on CCTV) and an exemplar (a known sample of the same item that is taken from or connected with the suspect) – to determine whether the two samples match, which results in the inference that the suspect is therefore the source of the unknown item of evidence connected with the crime, or at least more likely to be so.⁷⁰

The availability of exculpatory biological evidence for DNA testing was the primary cause of the debunking of microscopic hair analysis, and to a lesser extent bite-mark pattern matching. The availability of a more conclusive test of identity, however, was merely fortuitous. Human hair contains mitochondrial DNA (and sometimes nuclear DNA at the follicle).⁷¹ Microscopic hair comparison was debunked when a suspect hair that an analyst claimed had “matched” one recovered from the crime scene was proven definitively by DNA testing not to have come from the same person as the crime-scene hair.⁷² The availability of biological evidence for DNA testing in old bite-mark cases was less consistent, but there have been cases in which the saliva contained in a bite-mark that had previously been declared to “match” the defendant’s bite exemplar later turned out to contain DNA that excluded the defendant as its source.⁷³

Similarly, the availability of chemical tests definitively to detect whether (and identify) accelerants used to start arson fires contributed significantly to the debunking of fire-pattern

BULL. 930 (2007); Michael J. Saks, *et al.*, *Forensic Bite-mark Identification: Weak Foundations, Exaggerated Claims*, 3 J. L. & BIOSCIENCES 538 (2016); Segura & Smith, *supra* note 49.

⁶⁸ See Keith Findley, *et al.*, *Examining Shaken Baby Syndrome Convictions in Light of New Medical Scientific Research*, 37 OKLA. CITY L. REV. 219 (2012); Keith A. Findley, *et al.*, *Shaken Baby Syndrome, Abusive Head Trauma, and Actual Innocence: Getting It Right*, 12 HOUS. J. HEALTH L. & POL’Y 209 (2012).

⁶⁹ Leonetti, *Innocence Checklist*, *supra* note 12, at 121-22 and sources cited therein.

⁷⁰ PCAST Report, *supra* note 9, at 23.

⁷¹ Leonetti, *Junk Science*, *supra* note 6, at 64.

⁷² *Id.*

⁷³ In other cases involving convictions secured with bite-mark-match evidence, DNA from another source of evidence excluded the defendant as the perpetrator and so, by logical extension, from having bitten the crime victim; see also Hsu, *supra* note 66.

analysis.⁷⁴ Arson investigators can now subject evidence taken from the scenes of suspicious fires and subject it to chemical analysis through gas chromatography / mass spectrometry. The absence of the chemical signature of an accelerant at the location where a fire is suspected to have started, while not conclusive, is powerful circumstantial evidence that no accelerant was used (and, therefore, decreases the possibility of arson).

The pattern-matching forensic disciplines that continue to be used both internationally and in New Zealand – fingerprint comparison, handwriting comparison, and firearm (fired-bullet and tool-mark) comparison – have the same problems with reliability that microscopic hair comparison, bitmark-pattern matching, and fire-pattern analysis had, but without a newer, more accurate chemical test to debunk them. While it is possible to extract low-template DNA (“LTDNA”), also referred to as “touch” DNA, from fingerprints sometimes, there is no logical connection between the fingerprint and the LTDNA collected. Because it is not possible to “date” fingerprints or trace biological-evidence samples, in most fingerprint cases, even if LTDNA that could not have been left by the person deemed to have left the fingerprint was recovered, it would not conclusively exclude the individual as the source of the fingerprint, because the two evidence samples could have been left by different individuals and different times or because the LTDNA was transferred from some other original source.

The selective debunking of traditional pattern-matching disciplines that can be disputed through unequivocal DNA evidence has led to a strange division among the pattern-matching disciplines: those that have been foiled by DNA analysis and those that both have not and cannot be so foiled. Logically, the DNA exoneration of individuals convicted by microscopic hair-comparison evidence should cast doubt on the security of all the pattern-matching disciplines, but it is sometimes instead viewed as vindication by practitioners of other forensic sciences, evidence that their disciplines are merely superior by dint of having survived (albeit unchallenged because no challenge is possible).⁷⁵

III. THE PROBLEM WITH THE FORENSIC SCIENCES AND JUDICIAL SCREENING

A. Communicating Results: Match Testimony

When a forensic examiner believes that an exemplar associated with a suspect matches an evidentiary sample, at some point this match will have to be communicated to a jury (or judge or lawyers negotiating a guilty plea). This testimony necessarily takes one of two forms: (1) the ambitious alternative: a suggestion (either explicit or implicit) that the match between the known exemplar from the suspect and the crime-scene sample is unique, in the sense that no other gun would leave matching marks on a fired bullet or no other person’s finger would make a print that matched the latent from the window sill at the scene of the burglary);⁷⁶ or (2) the ambiguous

⁷⁴Findley, *supra* note 2 at 199-201.

⁷⁵Leonetti, *Innocence Checklist*, *supra* note 12, at 102-03.

⁷⁶Wallace v. Commissioner of Police (2016) N.Z.H.C. 1338, at ¶ 117 (describing testimony from an ESR firearms examiner that all of the bullets and shell casings used during the shooting that killed Wallace could only have come from the pistol of a single constable).

alternative: the watered-down claim that there is some similarity between or correlation among the crime-scene sample and the suspect exemplar, such that the suspect's item of evidence is "included" as a possible source of the crime-scene evidence (the torn tape used to bind the victim's wrists "could have" come from the roll found in the suspect's van, the pubic hair found at the scene of the rape is "consistent with" samples taken from the suspect), without specifying the probability of the coincident occurrence of the similarities (how many other rolls of duct tape in the world might have visually indistinguishable tear patterns from the suspect's). For example, the testimonial changes recommended by SWGFAST after the PCAST Report reflect a change from the former to the latter.⁷⁷

Both types of testimony ("match" and "inclusion") are fundamentally flawed. The former type of claim (a unique match) has never been validated in any of the forensic-science disciplines in which it is commonly deployed. For example, there is no scientific evidence to support the intuition that no two people have a finger with the same pattern of friction ridges on its surface, let alone the subsidiary claims that the uniqueness of human finger-ridge patterns transfers to surfaces or is sufficiently detectable to human fingerprint examiners such that fingerprints from two or more fingers would never be indistinguishable to the world's most proficient examiner.⁷⁸

The latter approach (a match or similarity that may not be unique), while a more conservative and intuitively appealing claim, is no more helpful to a jury (or scientifically reliable) because it lacks a related piece of information – if the crime-scene tape could have been torn from the suspect's roll of duct tape but the examiner cannot say whether or how many other rolls of duct tape from which it also could have come – then the jury has no sense of the significance of the inclusion of the suspect's roll of tape in the world of possible sources. Worse yet, in the absence of any evidence of the probability of coincidental occurrence, the jurors (or appellate court judges) may assign a very high one for themselves (assuming that it must have been a one-in-a-billion match).⁷⁹ An exclusion of the suspect as the source of the crime-scene evidence is definitive. The inclusion of the suspect in a pool of possible suspects who could have been, but were not necessarily to the exclusion of all others, the source of the crime-scene evidence is only helpful if one also knows the probability of a random match (*i.e.*, the likelihood that the edge of the suspect's roll of tape would have the same shape as the edge of the torn lengths from the crime scene without being the roll from which the crime-scene segments were torn). This is what gives DNA analysis its powerful probative value: random-match probabilities

⁷⁷ See *infra* notes 97-99 and accompanying text.

⁷⁸ Cf. Segura & Smith, *supra* note 49, at 7-8 ("The 'science' of bite-mark analysis relies on two conceits — first, that human dentition is unique, and second, that human skin is a sufficient and reliable substrate on which to record that uniqueness. The problem is that neither proposition has ever been proven — and the only empirical research attempting to do so has shown neither assumption to be true. Nonetheless, the subjective conclusions of bite-mark analysts have been allowed into evidence in criminal cases since the 1950s, when a Texas grocery store burglary was solved with the help of a dentist who matched a suspect's teeth to a bite mark left in a piece of cheese found at the crime scene.").

⁷⁹ *Hurrell v. R.*, (2018) N.Z.H.C. 1631, at ¶ 23 (describing the testimony by an ESR analyst pieces of duct tape found at the scene of a violent assault and the car park near where the victim's body was found were "consistent and could have come from the same roll" on which Hurrell's DNA was found and interpreting it as evidence that Hurrell was likely present when the assault took place).

in the range of one in one million or even one hundred million.⁸⁰ No such likelihood of an innocent match is known for any of the other forensic sciences.⁸¹

B. Under-Validation & Overselling of Traditional Disciplines

In light of the exonerations and debunking of the practices described in Section II (A) above, scholars and reformers have also focused increasing criticism on forensic-science techniques that continue to be prevalent in many countries, including: the interpretation of DNA mixtures;⁸² traditional pattern-matching or feature-comparison disciplines (like fingerprint comparison);⁸³ firearm (fired-bullet, shell-casing, tool-mark, and gun-shot residue) comparison;⁸⁴ foot, shoe, and tyre-impression comparison, and fiber and tape-edge “matching”;⁸⁵ bloodstain pattern analysis, forensic document analysis and handwriting comparison;⁸⁶ forensic

⁸⁰Jeremy George Edward McLaughlin v. R. (2015) N.Z.C.A. 339, at ¶ 18 (describing the testimony of a prosecution DNA analyst that it was one million times more likely that a DNA sample taken from underneath a murder victim’s fingernails came from McLaughlin than from another male individual selected at random from the population of New Zealand); R. v. K. (2019) N.Z.H.C. 1794, at ¶ 35 (“It is at least 300,000 million times more likely that the DNA originated from [K’s codefendant] than a random member of the public.”); New Zealand Police v. E.T. (2015) N.Z.Y.C. 412, at ¶ 4 (describing ESR’s conclusion that DNA samples taken from one of the knives used in an aggravated robbery were “seven hundred thousand million times more likely to have originated from E.”).

⁸¹ This is not to suggest that there are not valid methodological critiques of the population genetics statistics that underlie (and arguably understate) the probability of random DNA matches, but only to point out that DNA scientists at least attempt to quantify the statistical likelihood of a coincidental match between the DNA profile of biological evidence at a crime scene and the DNA profile of the suspect.

⁸²Itiel E. Dror & Greg Hampikian, *Subjectivity and Bias in Forensic DNA Mixture Interpretation*, 51 SCI. & JUST. 204 (2011) (documenting the inconsistency in the interpretations of data by North American expert DNA examiners and the role of contextually biasing information in producing disagreements in DNA mixture interpretation).

⁸³ See SIMON COLE, SUSPECT IDENTITIES: A HISTORY OF FINGERPRINTING AND CRIMINAL IDENTIFICATION (2001); Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again*, 41 AMER. CRIM. L. REV. 1189 (2004); Simon A. Cole, *Scandal, Fraud, and the Reform of Forensic Science: the Case of Fingerprint Analysis*, 119 W. VA. L. REV. 523 (2016); Gabel, *supra* note 2, at 285 (“The common—yet unrealistically romantic—starting point is that there are no two fingerprints exactly alike in the world. That assumption produces the further assumption that fingerprint analysis must be correspondingly reliable. This logic is erroneous.”); Tamara F. Lawson, *Can Fingerprints Lie? Re-Weighing Fingerprint Evidence in Criminal Jury Trials*, 31 AMER. J. CRIM. L. 1, 65 (2003) (“[F]ingerprint analysis is under attack because of the lack of study done on the accuracy of the examiners coupled with the unwillingness or inability of the forensic science community to detect, acknowledge, and correct mistakes.”).

⁸⁴ Wallace, *supra* note 76, at ¶20, ¶117, ¶121.

⁸⁵Itiel E. Dror & Simon A. Cole, *The Vision in “Blind” Justice: Expert Perception, Judgment, and Visual Cognition in Forensic Pattern Recognition*, 17 PSYCHONOMIC BULL. & REV. 161, 161, 165 (2010); PCAST Report, *supra* note 9, at 23.

⁸⁶ D. Michael Risinger, *et al.*, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lesson of Handwriting “Expertise,”* 137 U. PENN. L. REV. 731, 734, 739 (1989); Segura & Smith, *supra* note 49, at 5; Reinoud Stoel, *et al.*, *Bias Among Forensic Document Examiners: Still a Need for Procedural Changes*, 46 AUSTRALIAN J. FORENSIC SCI. 91 (2014).

pathology);⁸⁷ newer, more innovative forensic “sciences” (like forensic podiatry or forensic gait analysis);⁸⁸ forensic facial mapping;⁸⁹ forensic voice comparison;⁹⁰ and forensic soil analysis.

Critics have also focused on concerning but common testimonial features of forensic-science evidence, including the misuse of statistical analyses and the “overselling” of under-validated claims,⁹¹ particularly as they relate to the interpretation of forensic tests (such as claims of individualization⁹² and infallibility⁹³).

While forensic-science scandals and nonsense science often receive the most public focus,⁹⁴ the underlying problem with most forensic sciences, even those that have not played an outsize role in known wrongful convictions and which have some scientific basis in reality, is their lack of validation.⁹⁵ Traditionally, forensic analysts have been able to testify to an implausible, near infallibility of their disciplines because their true error rates were unknown, not because they were unknowable but merely because they had not been ascertained.⁹⁶

Forensic examiners and scholars who study forensic science generally distinguish between class characteristics and individualizing characteristics. For example, when a forensic firearms examiner compares two bullets – one taken from the crime scene and one fired from a gun associated with a suspect in the forensic laboratory – the caliber of the bullets and the number and direction of the lands and grooves are class characteristics. They are class characteristics because a large number of firearms will fire bullets with those particular markings

⁸⁷ Stephen Cordner, *et al.*, *The Foundations of the Comparison Forensic Sciences: Report of the President’s Council of Advisors on Science and Technology*, 24 J. L. & MEDICINE 297 301-02 (2016) (“Expert opinions in forensic pathology and clinical forensic medicine often lack the type of evidence base and scientific validity that the PCAST report suggests is required.”).

⁸⁸ See Gary Edmond & Emma Cunliffe, *Cinderella Story? The Social Production of a Forensic Science*, 106 J. CRIM. L. & CRIMINOLOGY 219 (2016).

⁸⁹ Gary Edmond, *et al.*, *Law’s Looking Glass: Expert Identification Evidence Derived from Photographic and Video Images*, 20 CURRENT ISSUES IN CRIM. JUST. 337, 338 (2009).

⁹⁰ See Gary Edmond, *et al.*, *Unsound Law: Issues with (“Expert”) Voice Comparison Evidence*, 35 MELBOURNE L. REV. 52 (2011).

⁹¹ See Simon A. Cole, *Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents’ Discourse*, 28 L. & POL. 109 (2006); Dawn McQuiston-Surrett & Michael J. Saks, *The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear*, 33 L. & HUMAN BEHAV. 436 (2009); see, e.g., Segura & Smith, *supra* note 49 (describing the exoneration of Steven Chaney in Texas, who was wrongfully convicted on the basis of “expert” testimony that there was a “one in a million chance” that bitemarks on a murder victim came from any source other than Chaney).

⁹² Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 VAND. L. REV. 199, 205 (2008).

⁹³ Simon A. Cole, *More than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985, 987 (2005) [hereinafter “Cole, *More Than Zero*”]; Jonathan J. Koehler, *Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter*, 59 HASTINGS L. J. 1077, 1091 (2008) [hereinafter “Koehler, *Fingerprint Error Rates*”].

⁹⁴ Jordan Michael Smith, *Forget CSI: a Disaster Is Happening in America’s Crime Labs*, THE INSIDER, (Apr. 30, 2014, 1:00 PM), <https://www.businessinsider.com/forensic-csi-crime-labs-disaster-2014-4>.

⁹⁵ Lyn Haber & Ralph Norman Haber, *Scientific Validation of Fingerprint Evidence under Daubert*, 7 L., PROBABILITY & RISK 87, 105 (2008) (“We have reviewed available scientific evidence of the validity of the ACE-V method and found none.”).

⁹⁶ See Cole, *More Than Zero*, *supra* note 93; Koehler, *Fingerprint Error Rates*, *supra* note 93; Saks & Koehler, *supra* note 2.

on them – e.g., a .22 caliber bullet with five rightward twisting lands and grooves. If the defendant’s gun “matches” the class characteristics of the gun that fired the crime-scene bullet, it is in a class of guns that could have done so – circumstantial evidence but hardly conclusive. Because gun barrels are machine manufactured, made of metal, and fire explosive projectiles, however, their barrels undergo processes of minute change, either through microscopic factory imperfections, wear and tear, and chemical processes like corrosion. Like in the Pursley case discussed above, firearms examiners believe (and routinely testify) that, due to the sheer number of ways that gun barrels can be affected by these processes, it is impossible for any two to leave the same (or microscopically indistinguishable) pattern of striations on a fired bullet.⁹⁷ It is the latter claim – that the number of minute striations caused by these forces of alteration are so numerous that their remaining signature is “unique,” such that a match of a bullet from a suspect gun to an evidentiary bullet must mean that the suspect gun, and only the suspect gun to the exclusion of all others, must be the source of the fired evidentiary bullet – that lacks validation, while simultaneously being treated as a definitive “identification” of the gun that is the source of the evidentiary bullet.

C. Human Factors

Reformers have also pointed to the role of unnecessary contextually biasing information in subjective forensic analyses and its ability to contribute to cognitive biases and interpretive error by analysts.⁹⁸ Probably the most prolific psychology scholar of the decision making of forensic examiners is Itiel Dror, a cognitive neuroscientist at University College London, who

⁹⁷ Possley, “Pursley,” *supra* note 29.

⁹⁸ See Itiel Dror, *The Paradox of Human Expertise: Why Experts Get It Wrong*, in THE PARADOXICAL BRAIN (Narinder Kapur, ed., 2011); Itiel Dror & Robert Rosenthal, *Meta-Analytically Quantifying the Reliability and Biasability of Forensic Experts*, 53 J. FORENSIC SCI. 900 (2008); Itiel E. Dror, *Cognitive Forensics and Experimental Research About Bias in Forensic Casework*, 52 SCI. & JUST. 128 (2012); Itiel E. Dror, *Human Expert Performance in Forensic Decision Making: Seven Different Sources of Bias*, 49 AUSTL. J. FORENSIC SCI. 541 (2017); Itiel E. Dror, *Practical Solutions to Cognitive and Human Factor Challenges in Forensic Science*, 4 FORENSIC SCI. POL. & MGT. 105 (2013); Itiel E. Dror & David Charlton, *Why Experts Make Errors*, 56 J. FORENSIC IDENTIFICATION 600 (2006); Itiel E. Dror & Jennifer L. Mnookin, *The Use of Technology in Human Expert Domains: Challenges and Risks Arising from the Use of Automated Fingerprint Identification Systems in Forensic Science*, 9 L., PROBABILITY & RISK 47 (2010); Itiel E. Dror, *et al.*, *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 FORENSIC SCI. INTL. 74 (2006); Itiel E. Dror, *et al.*, *The Impact of Human-Technology Cooperation and Distributed Cognition in Forensic Science: Biasing Effects of AFIS Contextual Information on Human Experts*, 57 J. FORENSIC SCI. 343 (2012); Gary Edmond, *et al.*, *Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals*, 14 L., PROBABILITY & RISK 1 (2015); Peter A.F. Fraser-Mackenzie, *et al.*, *Cognitive and Contextual Influences in Determination of Latent Fingerprint Suitability for Identification Judgments*, 53 SCI. & JUST. 144 (2013) [hereinafter “Fraser-Mackenzie, *et al.*, Contextual Influences”]; Peter A.F. Fraser-Mackenzie, *et al.*, *Forensic Judgment and Decision Making*, in COMPARATIVE DECISION MAKING (Thomas R. Zentall & Philip H. Crowley, eds., 2013); Amy M. Jeanguenat, *et al.*, *Strengthening Forensic DNA Decision Making Through a Better Understanding of the Influence of Cognitive Bias*, 57 SCI. & JUST. 415 (2017); Sherry Nakhaeizadeh, *et al.*, *Cascading Bias of Initial Exposure to Information at the Crime Scene to the Subsequent Evaluation of Skeletal Remains*, 63 J. FORENSIC SCI. 403 (2018); D. Michael Risinger, *et al.*, *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1 (2002); Reinoud Stoel, *et al.*, *Bias Among Forensic Document Examiners: Still a Need for Procedural Changes*, 46 AUSTL. J. FORENSIC SCI. 91 (2014).

has published dozens of studies on the effect of contextually biasing information on the decision making of forensic analysts in the pattern-matching disciplines. In several studies, performed in 2006, Dror's research team presented fingerprint examiners with fingerprints that they had previously matched but gave them contextual information that strongly suggested that the prints should not match.⁹⁹ Faced with the manipulated contextual information, which was contrary to their previous findings, the examiners reached the opposite decision as when they had examined the prints the first time most of the time.¹⁰⁰

Scholars largely divide scientific error into two types: methodological error and practitioner error.¹⁰¹ Practitioner error is the rate at which certified practitioners of the various fields declare false positives and negatives due to lack of training or proficiency; it is typically ascertained through well-designed proficiency tests. Methodological error is the rate at which error is inherent in a discipline, even when practiced by the most proficient practitioners, usually due to subjectivity or lack of controlling standards. These two types of error track the two types of validity examined in the PCAST Report: the foundational validity of a forensic methodology and the applied validity of the method in practice.¹⁰² The role of contextually biasing information in subjective decision making blurs the distinction between the two types of error in the forensic sciences.¹⁰³ Nonetheless, comprehensive and accurate rates of either or both of these types of errors simply do not exist for most forensic sciences, and rudimentary techniques for ascertaining them often suffer from profound methodological flaws.

In practice, the specificity (rate of false positives) and sensitivity (rate of true positives) of scientific tests tend to trade off with one another. The higher stringency at which a comparison is performed, the less likely that it is to produce both true and false positives. Conversely, lowering the stringency of a test makes it more likely to avoid both false and true positives. A test that is highly sensitive, therefore, is more likely to produce false positives. This is not a result of practitioner error, but rather a methodological flaw in the test itself.

D. Rushing to Market: Insufficiently Established New Techniques

Purported new forensic techniques pop up in criminal trials all the time, and they inevitably make their way into expert testimony, typically on the say-so of the practitioner of the new technique, without sufficient validation testing. Some of these newer techniques are pattern-matching techniques – for example, forensic paper comparison or comparison of fracture patterns in glass and glass polymers (like plastic tail-light covers on automobiles).¹⁰⁴ Other new techniques are laboratory tests revamped for forensic purposes – for example, isotopic hair

⁹⁹ Carrie Leonetti, *The Myth of the Appropriate Response to Trauma: "Abnormal Reactions" as Evidence of Guilt*, 58 GONZAGA L. REV. 379, 386 (2022).

¹⁰⁰ *Id.*

¹⁰¹ See Dror & Charlton, *supra* note 98.

¹⁰² PCAST, *supra* note 9, at 4.

¹⁰³ See Dror & Charlton, *supra* note 98; Fraser-Mackenzie, *et al.*, *Contextual Influences*, *supra* note 98.

¹⁰⁴ See Hitomi Miyata, *et al.*, *A Discrimination Method for Paper by Fourier Transform and Cross Correlation*, 47 J. FORENSIC SCI. 1 (2002).

analysis, pathogen fingerprinting, or the use of magneto-optical sensors to “detect” the remnants of obliterated serial numbers.¹⁰⁵

IV. PLAYING WITH FIRE: HOMEGROWN LAXITY AND THE FORENSIC SCIENCES

New Zealand’s criminal-justice system seems to be replicating other countries’ tendency to admit and rely upon forensic “science” evidence without sufficient skepticism or vetting, despite its clearly established role in wrongful convictions elsewhere. Many professional certification organizations for forensic scientists are transnational in nature, such as the International Association for Identification, the oldest and largest forensic-science organization in the world.¹⁰⁶

A. Forensic Science Disciplines in New Zealand

Sometimes, New Zealand is at the cutting edge of the creation of new forensic sciences. Other times, as a small country, New Zealand imports its forensic experts from abroad to testify to the results of techniques developed in other countries. The most recent iteration of the *Lundy* case is an example of the former. In *Lundy v. the Queen*,¹⁰⁷ the Crown introduced a novel form of messenger RNA (“mRNA”) analysis that purported to identify the type of tissue from which DNA cells originated.¹⁰⁸ The particular forensic mRNA analysis at issue in *Lundy* is performed only by the Netherlands Forensic Institute and the New Zealand Institute of Environmental Science and Research (“ESR”).¹⁰⁹ The technique was originally designed to identify cancer cells, then repurposed for forensic tissue analysis.¹¹⁰ Defence experts testified to scathing criticisms of both the underlying methodology of the mRNA technique, as well as the specificity threshold selected by the analysts, but the trial court admitted the evidence under section 25 anyway.¹¹¹ The court directed the jury, in pertinent part, that the Crown’s position was that

[O]n the shirt that Mr. Lundy was wearing the night his wife was killed in a manner which exposed her brain, there was found smeared on the shirt, in two separate places, pieces of central nervous system tissue; . . . when these two pieces of tissue were cut out of the shirt, and tested for human DNA, each of them yielded significant amounts of one person’s DNA, Mrs. Lundy.”¹¹²

¹⁰⁵ PRIYANKA KHATRIYA ET AL., PROGRESS TOWARDS DEVELOPING THE “PATHOGEN TOOLKIT,” 4 (2014), <https://www.ojp.gov/pdffiles1/nij/grants/246954.pdf>; BRETT J. TIPPLE, ISOTOPE ANALYSIS OF HAIR AS A TRACE EVIDENCE TOOL TO RECONSTRUCT HUMAN MOVEMENTS: COMBINING STRONTIUM ISOTOPE WITH HYDROGEN/OXYGEN ISOTOPE DATA, 4 (2015), <https://www.ojp.gov/pdffiles1/nij/grants/250339.pdf>.

¹⁰⁶ INTERNATIONAL ASSOCIATION FOR IDENTIFICATION, <https://www.theiai.org/> (last visited June 27, 2020).

¹⁰⁷ Mark Edward Lundy v. The Queen, [2019] N.Z.S.C. 152.

¹⁰⁸ *Id.* at ¶¶ 48-49.

¹⁰⁹ *Id.* at ¶ 55 n. 69.

¹¹⁰ *Id.* at ¶ 78.

¹¹¹ Lundy, N.Z.S.C. at ¶¶ 54-58, 61.

¹¹² *Id.* at ¶ 67.

It then left to the jury the job of deciding whether that scientific evidence was reliable¹¹³ – precisely the type of reliability determination that section 25 envisions the court making at the gatekeeping stage of admissibility.

*Barlow v. the Queen*¹¹⁴ is an example of the latter. In *Barlow*, the Crown called Charles Peters, a forensic analyst for the FBI, to give CBLA testimony that, according to the FBI CBLA database, the bullet that killed the victim could only have been a Geco automatic .32-caliber bullet, the precise type of bullet that the police had seized from a box of bullets associated with Barlow.¹¹⁵ Peters opined that the crime-scene bullets were “analytically indistinguishable” as the bullets associated with Barlow because they “came from the same manufacturer and are of the same type and the same source of lead at that manufacturer,” concluding: “They came from this, the same box or a box manufactured on or, manufactured or loaded on or about, that same date.”¹¹⁶ In summing up, the trial court directed the jury:

The Crown has sought to show that analysis of the materials of the bullets from the scene and from the box at the tip showed such a relationship in their content that you could conclude it was likely the bullets came from the same box, i e that the scene bullets, likely that the scene bullets had been taken from Mr Barlow’s box if you accept it was indeed his box¹¹⁷

The FBI has subsequently disavowed this type of CBLA testimony, publicly acknowledging that its technique was not scientifically valid and that CBLA testimony that its analysts gave during the CBLA era was unreliable.¹¹⁸ While the Privy Council agreed that the CBLA testimony offered in *Barlow* was inadmissible because it was “unscientific and untenable,”¹¹⁹ it ultimately dismissed the appeal under the proviso then-contained in section 385 (1) of the Crimes Act 1961.¹²⁰

Many of the forensic-science disciplines criticized in scholarly studies and government reports in other countries continue to be practiced in New Zealand with little to no challenge to their validity. ESR, which is the sole provider of forensic-science services to the New Zealand Police for use as evidence in criminal proceedings, regularly engages in crime-scene blood-pattern analysis, microscopic glass, paint, fibre, and tape-edge comparison,¹²¹ footwear and tyre impression comparison, and firearm and toolmark comparison, all techniques that have been the

¹¹³ *Id.* at ¶ 70.

¹¹⁴ John Robert Barlow v. The Queen, (2009) U.K.P.C. 30 (N.Z.).

¹¹⁵ *Id.* at ¶¶ 45-46.

¹¹⁶ *Id.* at ¶ 47.

¹¹⁷ *Id.* at ¶ 49.

¹¹⁸ See Leonetti, *Innocence Checklist*, *supra* note 12 and sources cited therein; see also *Barlow*, *supra* note 114, at ¶ 51.

¹¹⁹ *Barlow*, *supra* note 114, at ¶ 55.

¹²⁰ *Id.* at ¶¶ 75-76.

¹²¹ See, e.g., *Hurrell*, *supra* note 79, at ¶ 23.

subject of serious criticism in other countries.¹²² According to its 2019 annual report, ESR “analysed more than 50,000 pieces of forensic evidence”.¹²³ It describes itself as “one of the world’s leading experts in the field of forensic bloodstain pattern analysis,”¹²⁴ hardly an impressive distinction in light of the international criticism of the discipline. The Forensics Unit of the New Zealand Police makes “identifications” of latent fingerprints and questioned documents and testifies to matches arising out of these comparisons in criminal proceedings.¹²⁵

B. Faulty Reliance on the Adversary System

One error that courts make in adversarial criminal justice systems is overreliance on adversarial processes (cross-examination, rebuttal experts, jury addresses, and cautionary judicial directions) to screen faulty forensic-science evidence, even though scholars have long been critical of the ability of lawyers, judges, and juries to tell the difference between valid and unreliable science.¹²⁶ The case of *Shepherd v. The Queen*¹²⁷ is a good example of the New Zealand Court of Appeal falling into the same traps as courts in other countries have fallen; traps that have been clearly shown to correspond to wrongful convictions abroad. A crucial component of the Crown’s evidence on the issue of the identity of the robber was the expert testimony of a forensic analyst who opined that there was “strong support” for the identification of Shepherd as the robber.¹²⁸ The Crown’s expert derived his expertise “largely from training and experience,” rather than formal education or scientific training.¹²⁹ In response, the defence engaged in “extensive cross-examination” of the Crown’s expert and called an expert in physics and mathematics, with extensive experience in facial recognition, who opined that the Crown’s facial mapping evidence was unreliable because the features available for comparison in the CCTV footage was too poor to make a valid comparison.¹³⁰ An independent defence expert emphasized the lack of “any adequate database to provide a quantitative means to determine the statistical likelihood of a match between two facial images (assuming the two images are found to be similar, based on an accepted number of reference points and features).”¹³¹

There are several fundamental problems with the facial mapping technique in *Shepherd*. First, there is no way to know the probability of an individual chosen at random possessing any

¹²² ESR, Forensic Scene Investigations, <https://www.esr.cri.nz/our-services/testing/forensic-scene-investigations/> (last visited Nov. 6, 2019).

¹²³ ESR, Annual Report 2019, Oct. 10, 2019, at 2 (on file with Author) [hereinafter “ESR Annual Report”].

¹²⁴ *Id.* at 40.

¹²⁵ NEW ZEALAND POLICE, Forensics, <https://www.police.govt.nz/about-us/structure/teams-units/forensics> (last visited Nov. 6, 2019).

¹²⁶ See Gary Edmond & Mehera San Roque, *The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial*, 24 CURRENT ISSUES IN CRIM. JUST. 51 (2012) (discussing the failures of the adversarial system to weed out junk science in Australia); Keith A. Findley, *Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth*, 38 SETON HALL L. REV. 893 (2008) (discussing the failures of the adversarial system to weed out junk science in the United States).

¹²⁷ *Shepherd v. R.*, (2011) N.Z.C.A. 666.

¹²⁸ *Id.* at ¶¶ 3, 19, 80.

¹²⁹ *Id.* at ¶ 33.

¹³⁰ *Id.* at ¶¶ 4, 20, 32.

¹³¹ *Id.* at ¶ 22.

given point of comparison (head shape, chin shape, jaw line, moles, ear features) – *i.e.*, the likelihood of any particular distinctive facial feature’s occurrence in the human population. So far, attempts within the forensic-science community to ascertain these statistics have failed.¹³² As a result, it is also impossible to know the probability of all of the facial features identified and compared in a facial mapping comparison occurring randomly (*i.e.*, the probability of “matching” two faces from two different people).

Second, digital videos and photographs are subject to a process of compression, which can distort the identifying features that forensic examiners are analysing in their comparisons of suspect images.¹³³ For example, facial marks like freckles and moles, which are some of the best facial features for comparison, can be both removed and appear to be added through the distortion caused by compression algorithms.¹³⁴

Third, pose and expression are critical aspects of the measurements of distances between the landmark facial features that facial-mapping software makes.¹³⁵ Because of this, in preparing photographs for facial-mapping comparison, a forensic examiner first has to “normalize” the images for side-by-side comparison, by rotating, enhancing, and editing them by, for example, altering facial expressions.¹³⁶ The software used for this process engages in “pose correction,” alters “illumination,” and fills in “missing” facial features from partial facial images with an approximation of what it dictates should be there, by mirroring the visible parts of the suspect’s face or using “average” facial characteristics.¹³⁷ In the comparison that follows, the facial-mapping software does not distinguish between the facial characteristics present in the original image and those added later by the computer algorithm.¹³⁸

In rejecting Shepherd’s challenge to the admission of the facial-mapping testimony, the Court of Appeal noted that the “lack of a statistical database which could indicate the likelihood of similarities between faces within the wider population is a fundamental criticism of facial mapping evidence,” that it might “not be possible to adequately address” the inaccuracies of facial mapping through cross-examination, and that facial mapping had “been subjected to extensive academic criticism” on the grounds that it was an “epistemologically flimsy field of

¹³² Ava Kofman, *Losing Face: How a Facial Recognition Mismatch Can Ruin Your Life*, THE INTERCEPT, (Oct. 13, 2016, 8:00 AM), https://theintercept.com/2016/10/13/how-a-facial-recognition-mismatch-can-ruin-your-life/?campaign_id=3&emc=edit_MBAU_p_20200625&instance_id=19717&nl=morning-briefing®i_id=46959919§ion=backStory&segment_id=31860&te=1&user_id=862e72fe053910e76334f0e9328b7e78.

¹³³ *Id.*

¹³⁴ *Id.*

¹³⁵ Erik Sofge, *The End of Anonymity*, POPULAR SCIENCE, (Jan. 15, 2014, 1:02 AM), <https://www.popsci.com/article/technology/end-anonymity/>.

¹³⁶ Clare Garvie, *Garbage In, Garbage Out: Facial Recognition on Flawed Data*, GEORGETOWN LAW CENTER ON PRIVACY AND TECHNOLOGY, (May 16, 2019), https://www.flawedfacedata.com/?campaign_id=3&emc=edit_MBAU_p_20200625&instance_id=19717&nl=morning-briefing®i_id=46959919§ion=backStory&segment_id=31860&te=1&user_id=862e72fe053910e76334f0e9328b7e78.

¹³⁷ *Id.*; Sofge, *supra* note 135.

¹³⁸ Garvie, *supra* note 136.

evidence” that was too unreliable on which to base criminal convictions.¹³⁹ Nonetheless, the Court found the evidence admissible because it had been consistently admitted in British courts and has been admitted on a limited basis in Australian courts and because it was analogous to fingerprint matching.¹⁴⁰ The analogy to fingerprint matching was made without the awareness that forensic fingerprint analysis has been subject to scathing academic and scientific critique itself,¹⁴¹ and the Court of Appeal did not impose the same limitations on the admissibility of facial-mapping testimony that Australian courts have.

Despite holding that the trial court’s concluding jury directions regarding the facial-mapping testimony were inadequate because “the Judge did not make it clear that [the] facial mapping evidence was incapable of constituting positive identification[,] . . . did not say explicitly that there was no database which could give any statistical foundation for [the expert’s] opinion[, and] did not make it clear that the scale [of certainty] expressed by [the Crown’s expert] was based solely on his own expertise and experience, and was therefore subjective,” the Court of Appeal ultimately concluded that there had not been a miscarriage of justice because “[t]he jury had before it the evidence of both expert witnesses who were each rigorously and extensively cross-examined” and could understand that the Crown’s expert’s opinion “was subjective in nature,” specifically noting that the court had directed the jury that the testimony lacked any “‘science’ in relation to the likelihood of similarity of faces.”¹⁴² The Court delegated to the jury the gatekeeping function that is supposed to belong to the Court under section 25 of the Evidence Act 2006, concluding that “it was open to the jury to accept [the expert testimony] as reliable.”¹⁴³

To any observer of international critiques of the lackluster judicial screening of under-validated and oversold pseudo-scientific forensic evidence, the *Shepherd* opinion is depressingly familiar.¹⁴⁴ It contains the typical errors that American courts have made in failing to uphold their gatekeeping obligations. The Court conflated the concept of expertise generally (that the witness has some training or experience that an average juror would not) with the specific expertise necessary to render a particular opinion (*e.g.*, the interpretive significance of a “match” in the absence of a database that would allow the calculation of error rates or a random match probability). The Court allowed a (non-scientist) forensic practitioner’s unsupported opinion of the reliability of their own technique, garnered through third-party assurances at practitioner trainings rather than scientific tests of validity, to suffice as an adequate basis for the proponent’s burden of establishing the reliability of scientific evidence, even in the face of trenchant criticism from a defence witness with actual scientific training and credentials. The Court at least

¹³⁹ *Shepherd*, *supra* note 127, at ¶ 47.

¹⁴⁰ *Shepherd*, *supra* note 127 at ¶¶ 37-54 (internal footnotes & citations omitted).

¹⁴¹ See *supra* note xxx & accompanying text.

¹⁴² *Id.* at ¶¶ 95-97.

¹⁴³ *Id.*

¹⁴⁴ See, *e.g.*, Gabel, *supra* note 2, at 285 (“American courts have improperly legitimized various forensic disciplines without subjecting them to the kind of scrutiny that would be required of novel scientific or technical evidence today.”); Michael J. Saks, *Merlin and Solomon: Lessons from the Law’s Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069 (1998).

implicitly treated the absence of governing standards or protocols for a technique as a factor in favor of admissibility, since the absence of standards makes it impossible for an expert to have violated them. The Court permitted an expert's confession that a purportedly scientific technique was subjective, unscientific, and lacked controlling standards to substitute for the required showing that the technique was replicable and testable. The Court confused the concept of the admissibility of expert "scientific" testimony with the issue of the weight to be given evidence by the jury after it has met a threshold of judicial reliability.¹⁴⁵ The Court conflated the adversarial testing of cross-examination and a jury trial with the scientific testing required to establish scientific validity.¹⁴⁶ Finally, the Court failed to engage meaningfully with the substance of scholarly critiques of the testimony at issue, stemming from a likely lack of understanding of the scientific method itself.

V. CONCLUSION

In 1992, the New Zealand Court of Appeal notoriously proclaimed that there was "no good reason" to believe that wrongful convictions stemming from official misconduct ever had or would occur in New Zealand, noting the availability of ever-improving scientific evidence to ascertain guilt or innocence.¹⁴⁷ This Article takes the opposite tack, based on nearly three decades of international experience with exonerations, suggesting that there is no good reason to think that the forensic sciences in New Zealand are somehow immune to the criticisms targeted at them in other parts of the world. Unlike language or custom, scientific techniques tend to bleed seamlessly across borders. New Zealand lacks any validation studies demonstrating the reliability of its forensic analyses, like other jurisdictions where forensic-science disciplines have faced sustained criticism. These deficiencies threaten the accuracy of the criminal-justice system. New Zealand's reliance on the American *Daubert* test for scientific evidence has compounded the problem, by importing an insufficient screening mechanism for keeping unreliable scientific opinions from contaminating criminal trials.

New Zealand courts should not leave to juries, as part of the ordinary adversarial testing of the strength of evidence, the job of breaking the tie between prosecution witnesses whose expertise is limited to on-the-job training and experience and defence witnesses with advanced scientific training who testify to the validation failures of the prosecution evidence, but rather should consider such technical criticisms at the admissibility stage of the proceedings. They

¹⁴⁵ Gary Edmond, *Icarus and the Evidence Act: Section 137, Probative Value and Taking Forensic Science Evidence at Its Highest*, 41 MELB. U. L. REV. 106 (2017).

¹⁴⁶ Jane Campbell Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES J. 16, 29 (2005) ("The single most important observation about judicial gatekeeping of forensic sciences is that most judges under most circumstances admit most forensic science."); D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALBANY L. REV. 99 (2000); see, e.g., *Gilham v. The Queen*, [2012] N.S.W.C.C.A. 131, at ¶ 405 ("It is in the discharge of the different but allied obligations of the expert and the Crown Prosecutor that the jury is educated and informed about matters in issue between the Crown and the accused which are beyond the jury's experience.").

¹⁴⁷ *Burt v. Governor-General*, (1992) 3 N.Z.L.R. 672, 682 (Ct. App.).

should not allow individual forensic-science disciplines, or forensic scientists as a community, to self-police the sufficiency of the validation required of them or allow a consensus belief among practitioners that the particular form of analysis – the *ipse dixit* of the examiners themselves – that they practice is reliable to substitute for their independent gatekeeping obligation. They should not allow (non-scientist) forensic practitioners, particularly those without formal scientific training and expertise, to establish the foundation for reliability of their own techniques by assuring courts of their subjective belief in reliability, in the absence of validation evidence to buttress those beliefs.

Instead, courts in New Zealand should separate the concept of expertise, training, and experience generally with the specific question of whether a purported scientific expert witness has the necessary expertise to render a particular opinion. They should demand evidence of objective governing standards and protocols for a technique as a precondition of admissibility, rather than favoring disciplines that lack them by treating the absence of evidence of testability as evidence of infallibility. They should treat the absence of controlling standards and evidence of replicability, particularly for disciplines that fail to administer forensic tests blindly (*i.e.*, without providing the analyst access to unnecessary contextualizing biasing information), as powerful evidence that an opinion resulting from those tests is impermissibly unreliable. Courts should take seriously the substantial body of scholarly and scientific criticism of the forensic sciences, and they should treat the failure of forensic-science practitioners to engage meaningfully with the substance of scholarly and scientific critiques of their techniques as evidence of a lack of understanding of the scientific method, rather than as the type of good-faith disagreement that is usually left to the adversarial system of justice to resolve.