University of Michigan Journal of Law Reform

Volume 50

2017

Bias, Subjectivity, and Wrongful Conviction

Katherine Judson
University of Wisconsin Law School

https://doi.org/10.36646/mjlr.50.3.bias

BIAS, SUBJECTIVITY, AND WRONGFUL CONVICTIONS

Katherine Judson

I'm Kate Judson, and I teach at the University of Wisconsin in the Law School. I've been asked to summarize the science for the lawyers in the room, which I have been told is a hopeless cause, but I'm a Cubs fan and I do wrongful convictions work, so I am comfortable with hopeless causes and I feel I'm up to the challenge. I'm confident that lawyers can learn science, maybe even before the Cubs win the World Series.

I want to talk about bias, subjectivity and wrongful convictions. In law, we apply lessons from the past to contemporary problems. When we look at a wrongful conviction after exoneration, we get to look back at what might have caused it. This is valuable because we learn lessons we can then apply to prevent future miscarriages of justice. Today I'd like to discuss a little about what we know from past exonerations and some of the factors that can contribute to wrongful convictions. What's great about this is that a lot of these factors have been ably analyzed by everybody who's been up here before me, so I'm really grateful for that. Thank you.

Examining wrongful convictions that were later overturned with new DNA evidence allows us to identify six factors that significantly contribute to wrongful convictions.[1] The influence of these factors may be slightly different percentagewise for DNA versus non-DNA cases, but they are significant in both contexts. Wrongful convictions often occur as a result of incentivized testimony from witnesses, like jailhouse informant or "snitch" testimony, false confessions, eyewitness identification error, bad lawyering, government misconduct and faulty forensic science.[2] The one I'm going to focus on today is faulty forensic science.

Faulty forensic science may mean a couple of different things. First, there is unverified or unvalidated forensic science, science that doesn't have a reliable evidence base. Then, there are misstatements of forensic science: misstatements about probability, misstatements about what the evidence actually shows, and misstatements about reliability. And then there is actual misconduct or fraud by experts or lab analysts, which, though alarming, happens

1. Barry Scheck, Peter Neufeld & Jim Dwyer, Actual Innocence: When Justice Goes Wrong and How to Make It Right 361 (2001).

2. *Id.*

less frequently. It makes lots of headlines when it happens,[3] but does not appear to happen very frequently.

More often what we're dealing with when we deal with faulty forensic science are issues of bias and issues of improper testimony. About sixty percent of wrongful convictions that used forensic science involve improper testimony.[4] After a comprehensive review of hair comparison analysis cases following a series of high-profile exonerations, a joint press release issued by the FBI, the National Association of Criminal Defense Lawyers (NACDL), and the Innocence Project acknowledged that FBI testimony on microscopic hair analysis contained testimonial errors in more than ninety percent of cases in an ongoing review—analysts misstated what the evidence actually showed and testified with more certainty than the evidence warranted.[5]

Often the problems lie with a lack of underlying research.[6] Many have recognized the difficulty and potential for missteps when forensic science disciplines are developed to aid prosecutors in the courtroom, rather than based firmly in hard science.[7] The *Daubert* case itself explicitly recognizes that "there are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory."[8] In the absence of adequate data, there is more room for subjectivity, and therefore bias, to creep into the process. In science-dependent child abuse cases, the underlying research is often inadequate or fraught with error.[9] This creates a climate that can lead to wrongful conviction.

It is particularly important to consider how much subjectivity is involved in child abuse cases because many child abuse cases are science-dependent, meaning that the entire case rests upon expert opinion (in some cases, science is used to show both the actus reus and the mens rea, sometimes in the absence of any other evidence),

---

3. *Id.* at 138.

4. Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 14 (2009).

5. Press Release, Fed. Bureau of Investigation, FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review (April 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.

6. NAT'L RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 187 (2009) ("[T]he forensic science disciplines suffer from an inadequate research base.").

7. *Id.*

8. Daubert v. Merrill-Dow Pharmaceuticals, 509 U.S. 579 at 596–97 (1993).

9. *See, e.g.,* Nytt Juridiskt Arkiv [NJA] [Supreme Court] 2014-11-2 B 3438-12 (Swed.), http://rffr.se/wp-content/uploads/2014/12/Swedish_supreme_court_20141016.pdf (discussing the limitations of research tracing the reasons of a child's injury to violent shaking by adult).

rather than science-reliant, where forensic science helps elucidate a particular part of the case (identity of the perpetrator, for example). Just because a case involves faulty forensic science, whether the case is science-dependent or science-reliant, does not mean that the defendant is innocent. If the importance of the science goes up, however, so does the risk of wrongful conviction. If a conviction rests entirely, or nearly so, on unvalidated, misleading, or improper forensic science, it is of particular concern. When a field of forensic science is without safeguards for validity and reliability, expert witness testimony should either be kept from the jury (as in successful *Daubert* challenges) or, if the jury will hear it, the witnesses must make the shortcomings in the data absolutely clear. In wrongful convictions, however, we often see unreliable methods or data and then a witness who testifies with more certainty than the science warrants.[10]

Proponents of the Shaken Baby Syndrome/Abusive Head Trauma hypothesis address this flaw in the research by claiming that in the absence of high quality research, it is appropriate to rely on clinical judgment.[11] In other words, the training and experience of the physician is what comes into the equation when the research is poor or lacking. That's really problematic because it introduces the subjectivity that infects the other forensic sciences.

Some of this is likely due to bias. Bias is a significant and often unrecognized problem in science-dependent cases. One of the things that we always need to address when we talk about bias, as Professor Maddox and Professor Sommers pointed out, is that no one is immune. There is nothing about law or medical school that makes a person immune to bias. Having bias doesn't make you a bad person, it makes you a person. We all have implicit bias. We can't keep ourselves from having it; it is, by definition, unintentional. What's important is to recognize it and try to minimize its effects.

Bias, far from being absent among physicians and experts, has been seen and studied in the medical care context. The National Academies Press have a publication called *Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare*.[12] Data from this

10. Garrett & Neufeld, *supra* note 4, at 14.

11. *See* Sandeep Narang, *A* Daubert *Analysis of Abusive Head Trauma/Shaken Baby Syndrome*, 11 HOUS. J. HEALTH L. & POL'Y 505, 532 (2011) (discussing that well-established medical diagnoses such as migraine headaches have no randomized controlled trials).

12. *See generally* INST. OF MED., UNEQUAL TREATMENT: CONFRONTING RACIAL AND ETHNIC DISPARITIES IN HEALTH CARE (2003) (reporting findings assessing the extent and source of racial and ethnic differences in the quality of health care received by patients not attributable to known factors such as access to care, ability to pay, or insurance coverage).

study suggests that different medical care is provided or prescribed to different groups. Race and gender can affect all aspects of healthcare, including what medications and procedures doctors prescribe. There is evidence suggesting that even wait times differ based on the racial or ethnic group of the patient.[13]

Further complicating this issue is the fact that clinical judgment in the child abuse context is different from clinical judgment in the treatment context. Relying on "clinical judgment" creates a number of problems, including a *Daubert* issue. "Clinical judgment" often means merely the say-so of the treating physician. Another term for the mere say-so of an expert witness is, of course, *ipse dixit,* which the *Daubert* decision prohibits. Legal scholar D. Michael Risinger notes that "in circumstances when experience alone does not resolve the main doubts about reliability, it would be irrational and therefore an abuse of discretion to rely upon it."[14]

Confirmation bias, or the tendency to seek out confirmatory information once a hypothesis is developed, frequently comes into play in medical decision-making. Confirmation bias may mean seeking confirmatory evidence, rather than disconfirming evidence. It also can affect the tendency to recall evidence. Confirmation bias may cause a person to recall confirming evidence in a biased manner. It can also cause people to interpret ambiguous evidence in a manner that conforms to their existing beliefs. When clear guidelines and unambiguous testing exist, confirmation bias plays a smaller role in medical diagnosis, though no one is free from it. When those standards are absent, however, confirmation bias may have a greater effect on outcomes.

Confirmation bias in police investigation has been studied. Police officers often rate discomfirming or exonerating evidence as less reliable or credible, and prefer guilt confirming evidence that supports their initial hypotheses.[15] This is another example of subjectivity or uncertainty in the process allowing bias to come in.

Bias can be a factor in any criminal case, but it's a particular problem in science-dependent child abuse cases. The lack of objective tests, standard diagnostic criteria, and strong evidence base

---

13. Kristin N. Ray et al., *Disparities in Time Spent Seeking Medical Care in the United States*, 175 JAMA INTERN. MED. 1983, 1984 (2015).

14. D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science After* Kumho Tire Co. v. Carmichael, 57 WASH. & LEE L. REV. 767, 776 (2000).

15. Karl Ask & Par Anders Granhag, *Motivational Bias in Criminal Investigators' Judgments of Witness Reliability*, 37 J. APPLIED SOC. PSYCHOL. 561, 579–80 (2007); Karl Ask, Anna Rebelius & Par Anders Granhag, *The "Elasticity" of Criminal Evidence: A Moderator of Investigator Bias*, 22 APPLIED COGNITIVE PSYCHOL. 1245, 1253–55 (2008).

conspire to ask doctors, police officers and lawyers to make subjective judgments that are crucial to the outcome of the case. We heard earlier today that bias might sneak into the process when someone has too great a workload or when a person feels stressed or threatened. I'm sure that everybody in here who's been in court, either as a lawyer or an expert witness, has had that feeling. People tend to fall back on their assumptions—engage in thought that is comfortable for them—under these circumstances.

Investigators also show marked confirmation bias when they're asked to form a hypothesis of guilt early in the evaluation of evidence.[16] Witnesses often claim that they weren't biased or didn't rush to judgment but, of course, people can't will away bias or necessarily stop the rush to judgment.

One of the things we see a lot in science-dependent child abuse cases is a bias in the medical history and in the taking of the medical history. You might hear something like "inconsistent history," where the history doesn't match the expectations of injury. This is significant especially if the only history given is history of a fall. You might have a caregiver who says the child fell, but the physician says, "These injuries can't have happened from this kind of a fall." Why? "Because every time I see it, I diagnose abuse." And that's where circularity enters into that process.

For example, the Wisconsin Innocence Project is handling a case in which a physician claimed a decedent's injuries could not have resulted from a fall, even when a witness reported a fall from a chair.[17] Because the medical experts had a preconceived notion that the decedent's injuries were inconsistent with a fall, doctors testified at trial that the fall either did not occur or could not have been fatal. This is an example of history that falls victim to bias. Rather than examine assumptions about what kinds of injuries could be expected from a fall (and that while rare, fatal falls from furniture occur) and fully consider the history given (a witnessed fall), expert witnesses instead testified that a collection of medical findings was always the result of abuse.

It's fair to point out that in a child abuse case, the history may not always be correct. A caregiver could misremember an incident or be unclear. An investigator could misremember or misrecord what the caregiver said. A physician could receive evidence second- or third-hand from caregivers and investigators, and portions of it could be incorrect by the time it gets to them. The history given by

16. Barbara O'Brien, *Prime Suspect: An Examination of Factors that Aggravate and Counteract Confirmation Bias in Criminal Investigations*, 15 PSYCHOL., PUB. POL'Y, & L. 315, 328 (2009).

17. People v. Bailey, 47 Misc.3d 355, 360 (Monroe Cty. Ct. 2014).

the caregiver could be a lie, that is to say, the person could have made up a story or abused the child and lied about it. But the history could also be true and exonerating. Sometimes when someone says they don't know what happened, they really don't know what happened. And the reason that they're at the doctor's office or at the hospital is to get help in figuring out what occurred. So just from the injuries alone, and just from the statements alone, it's often impossible to tell whether what you have is a true or a false statement.

Something that always needs to be addressed in a criminal investigation is tunnel vision. Tunnel vision is the "'compendium of common heuristics and logical fallacies' to which we are all susceptible, that lead actors in the criminal justice system to 'focus on a suspect, select and filter the evidence that will build a case for conviction while ignoring or suppressing evidence that points away from guilt.'"[18] Tunnel vision bias is in play to some degree in many, if not most, wrongful convictions. Tunnel vision may occur when investigators home in on one suspect rather than others, or, in child abuse prosecutions, where investigators and experts focus on one cause of a child's injuries, ignoring other possibilities. We can all think of an example of this. We can't always avoid it, but we can try to be aware of it.

Bias is exacerbated by factors in the system, including role effects. People's perceptions of their role can influence their decisions.[19] This is one of the risks of embedding forensic science within a law enforcement agency. If an analyst sees his or her role as being a part of a police department and to catch criminals, that role effect might have an impact on their decision making.[20] This

---

18. Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 WIS. L. REV. 291, 292 (2006) (citing Dianne L. Martin, *Lessons About Justice from the "Laboratory" of Wrongful convictions: Tunnel Vision, the Construction of Guilt and Informer Evidence*, 70 UMKC L. REV. 847, 848 (2002)).

19. *See generally* James W. Pichert & Richard C. Anderson, *Taking Different Perspectives on a Story*, 69 J. EDUC. PSYCHOL. 309 (1977) (finding that an idea's significance in terms of a given perspective determined whether an idea would be learned and recalled a week later); Richard C. Anderson, James W. Pichert & Larry L. Shirey, *Effects of the Reader's Schema at Different Points in Time*, 75 J. EDUC. PSYCHOL. 271 (1983) (finding that perspectives assigned before, shortly after, and long after reading all had substantial effects on recalling a story); D. Michael Risinger et al., *The* Daubert/Kumho *Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CALIF. L. REV. 1 (2002) (revealing the extent to which reliability is undermined depends on observer effects, the characteristics of the expertise at issue, especially the degree to which it depends on subjective human judgment).

20. Risinger et al., *supra* note 19, at 18 (quoting James E. Starrs, *The Ethical Obligations of the Forensic Scientist in the Criminal Justice System*, 54 J. ASS'N OFFICIAL ANALYTICAL CHEMISTS 906, 910 (1971) for the proposition, "It is quite common to find . . . laboratory facilities and personnel who are, for all intents and purposes, an arm of the prosecution. They analyze material submitted, on all but rare occasions, solely by the prosecution. They testify almost

arguably can also be a problem of creating child abuse investigation teams. We've heard a lot of conflicting information today about child abuse investigation teams. They can work, they can not work. A frank discussion of role effects isn't really to say whether these things are good or bad; it's really more to recognize that when you create a group of people who have a particular role and who see their role in a certain way, that does open them up to the possibility of bias.

Then there are conformity effects: the tendency to conform to the perceptions, beliefs, and behavior of others.[21] This is really significant and something that we see an awful lot of the time in child abuse cases because, again, emotions run very, very high. People have strong opinions and they come down hard on one side or the other. I want to give one very brief example of conformity pressure. So after testifying as a defense expert for parents accused of abuse, an expert received an email from a colleague. The colleague wrote, "You are deluded at best and criminal at the worst since you are aiding and abetting in the crime, after the fact to be sure, but nevertheless actively supporting this type of deviant behavior."[22]

This isn't to call somebody out about writing a mean email. But it's to understand that incredible pressure can be brought to bear on individuals who do not conform. That allows bias into the equation. It can affect the person's judgment in the future, and it can affect the person's willingness to stand up for a patient or for themselves, when they have a different opinion from the rest of the group. It also shows what can happen when a professional does not recognize the importance of feedback from downstream. Here, the physician who testified for the government had an opportunity to revisit his diagnosis in light of what he learned from experts for the other side. Instead, he took the opportunity to insult those experts.

Which brings us to groupthink. Working together, having an identity as part of a group, can undermine the ability to catch errors or reveal alternatives. A really classic example that happens in the medical context is the resident disagrees with her attending, but is afraid to say anything because the attending is the teacher, the attending is the boss.

---

exclusively on behalf of the prosecution. They inevitably become part of the effort to bring an offender to justice. And as a result, their impartiality is replaced by a viewpoint colored brightly with prosecutorial bias").

21. *Id.* (revealing the extent to which reliability is undermined depends on observer effects, the characteristics of the expertise at issue, especially the degree to which it depends on subjective human judgment).

22. E-mail from Stephen Done, MD, to Julie Mack, MD (Dec. 5, 2014) (on file with author).

This is just an example of this occurring in a child abuse case from a transcript. "Q: And had anyone made a determination at this time what the cause or potential causes of the injuries were?" "A: It would be unfair if I didn't say that everybody who had seen her felt very strongly that this was likely to be abuse."[23]

Some people can do this, and some people have difficulty with it. But if you're the lone voice in a group of five or six or seven or eight or nine people who doesn't agree, it might be difficult to articulate your concerns, it might be difficult to stick up for yourself if you're attacked, and it might be difficult to verbalize what you think is going on.

Diagnosis momentum is a particular type of cognitive bias in medical diagnosis. It can be experienced by a doctor on his or her own, but it also goes hand in hand with the challenges of working in groups. Dr. Jerome Groopman addresses the problem in his book *How Doctors Think*.[24] Once a particular diagnosis becomes fixed in a physician's mind, despite incomplete evidence or discrepancies in evidence, the first doctor passes on this diagnosis to his peers or subordinates. He compares it to a boulder rolling down a mountain gaining enough force to crush anything in its way.[25] We know that error in medical diagnosis is an enormous problem: the Journal of the American Medical Association has published articles showing that cases of delayed, missed, and incorrect diagnoses are common, with an incidence range of about ten to twenty percent.[26] This represents an average; certainly some diagnoses are more difficult to make than others. If a condition can be diagnosed with a test that is easy to perform and frequently accurate, the error rate might be lower, and when diagnoses are ambiguous, without standardized diagnostic criteria or gold standard tests, the rate might be higher.

Error is higher with respect to clinical diagnoses, and lower with respect to diagnostic tests.[27] In the child abuse context, what that means is that you can expect the error rate to be higher when a diagnosis is something like, this was or was not abuse, and lower with respect looking at, for example, a CAT scan and identifying a

---

23. Transcript of Record at 117, State v. Morris, (WA) (transcript on file with author).

24. JEROME GROOPMAN, HOW DOCTORS THINK 128 (2007).

25. *Id.* ("Diagnosis momentum, like a boulder rolling down a mountain, gains enough force to crush anything in its way.").

26. *E.g.* Mark L. Graber, Robert M. Wachter & Christine K. Cassel, *Bringing Diagnosis into the Quality and Safety Equations*, 308 JAMA 1211, 1211 (2012).

27. *See id.*; *but see* Eta S. Berner & Mark L. Graber, *Overconfidence as a Cause of Diagnostic Error in Medicine*, 121 AM. J. MED. S2, S6 (2008) (noting that "it is clear that an extensive and ever-growing literature confirms that diagnostic errors exist at nontrivial and sometimes alarming rates [and that] [t]hese studies span every specialty and virtually every dimension of both inpatient and outpatient care").

fracture, or looking at an x-ray and identifying a fracture, or looking at an MRI and identifying a subdural hematoma. So the diagnostic testing portion tends to be somewhat more accurate than actually coming to the clinical diagnosis. That should make sense. Because any time we have to synthesize data, there's a lot more room for error than when we're looking at a lab result without interpreting it.

It's also clear that over-confidence and other cognitive biases have a role to play in misdiagnosis. Cognitive errors are defined in this context as reflecting problems gathering data, failing to elicit accurate information from the patient.[28] So there is failure in history-taking, failure to recognize the significance of the data—like misinterpreting test results. But the most common is really the failure to synthesize or put it all together. What that means in a child abuse context is that physicians are generally fairly accurate at identifying particular medical conditions, identifying a broken bone, identifying a bruise, identifying a bleed. But asking doctors to consider causation requires them to engage in a more speculative enterprise.

When Berner and Graber discuss misdiagnosis and over-confidence, they focus in particular on faulty heuristics and premature closure.[29] Premature closure is very important in this context. It's narrowing the choice of diagnostic hypotheses too early in the process.[30] It's a very, very common mistake and it happens really frequently. So like the example I gave before from the transcript where the physician said, "Q: Had anyone made a determination at this time what the cause or potential causes of the injuries were?" "A: It would be unfair if I didn't say that everybody who had seen her felt very strongly that this was likely to be abuse."[31] That's a really great example of premature closure.

Another type of bias frequently seen in wrongful convictions is context bias.[32] Context bias is another thing that's really tricky, because it can be useful, but it also can be misleading. Testimony in a case might include claims from investigators or physicians that a parent or caregiver delayed medical treatment to an injured child because they took the child to a family member rather than immediately to the emergency room, or chose to drive a child to the hospital rather than waiting for an ambulance. While a delay in

28. Berner & Graber, *supra* note 27, at S7.
29. *Id.* at S8.
30. *Id.*
31. *See* Transcript of Record at 117, State v. Morris (WA), *supra*, note 23.
32. Berner & Graber, *supra* note 27, at S8.

seeking care may be significant, it could also be a genuine misunderstanding of the seriousness of a child's condition, or a genuine belief that driving to the hospital would be faster than waiting for an ambulance. Investigators might describe an accused person as flat or unemotional after a child is seriously injured or dies, which might lead to a conclusion that the person was uncaring or not particularly upset about the injured child. While that is one explanation for a flat affect, shock or grief might also produce numbness.

I once had an emergency room physician say to me, "Look, you know, every time I give a family bad news, it goes differently. It's never the same." Different people, even within the same family, can have radically different reactions to the same bad news. Some might cry when confronted with a tragic situation; others might become angry, or even appear not to react at all. These reactions do not necessarily give the investigator reliable evidence about how the person feels about the loss of the decedent.

The other really important thing that we're dealing with here is the problem of inadequate feedback.[33] Inadequate feedback is a major obstacle to correcting diagnostic error. Feedback that is delayed or absent may not be recognized for what it is and then the perception that misdiagnosis is not a big problem remains unchallenged. When there is an absence of information that the diagnosis is wrong, the expert is allowed to assume it must be correct.[34]

We see this really frequently in the legal context when somebody makes a diagnosis of abuse. When supporters of SBS/AHT have been challenged by expert witnesses for the defense, they have at times chosen to marginalize or attack the experts and pretend there is no debate.[35] Some attacks are extreme; after a high-profile child abuse case was dismissed recently, a child abuse pediatrician told the Boston Globe that physicians who assist defense attorneys are "hired out as defense whores."[36]

And that's really an inability recognize the feedback that the legal system is giving you, right? Juries will come to a conclusion

---

33. *Id.* at S10.

34. *See id.*

35. Peter J. Strouse, Editorial, *Child Abuse: We Have Problems*, 46 PEDIATRIC RADIOLOGY 587, 588 (2016); Thomas L. Slovis et al., Editorial, *The Creation of Non-Disease: An Assault on the Diagnosis of Child Abuse*, 42 PEDIATRIC RADIOLOGY 903 (2012).

36. Kevin Cullen, *Nanny's Case Could Have Broad Effects on Child Abuse Prosecutions*, THE BOSTON GLOBE (Sept. 12, 2015), https://www.bostonglobe.com/news/nation/2015/09/12/nanny-case-could-have-broad-effects-child-abuse-prosecutions/doika7FUTM8W8wg0O84v7L/story.html.

that's scientifically inappropriate. It's a possibility and it's something that certainly should be considered. But to summarily ignore that when we have such an intersection of science and the law is a serious problem.

Just a little bit more on the problem of absent feedback.[37] Dr. Schiff makes a really great point about how feedback on patient response is critical for knowing not just how the patient is doing, but also how we as clinicians are doing.[38] And most doctors rely, at least in part, on patient feedback to know whether their diagnoses are correct. So you have a patient, and you think your patient has a particular disorder. You give them medication, and it works. You can be fairly confident that you diagnosed it correctly and you gave them the right medication. It doesn't work and you have to go back to the drawing board. That kind of feedback is absent in most child abuse cases.

Schiff also says that these refined signals from downstream represent an antidote to anchoring bias, which he describes as fixing on a particular diagnosis, despite clues that such persistence is unwarranted.[39] And when that's absent, you never get those clues. So the feedback gap in medicine is especially pronounced in diagnosis of child abuse. It's an enormous problem there, and the physician doesn't get the kind of feedback that they need to refine their practice, making them unable to revise their thinking or improve their skills to face the next challenge.

Any time you have subjectivity, it leaves room for bias. Without explicit criteria for decision making, individuals will disambiguate the situation using whatever information is most easily accessible including all kinds of things, including stereotypes.

We've talked about the research objectives and challenges, but it's important to note that the circularity here that we're talking about really has a lot to do with these retrospective case studies which depend on accurately sorting cases into abuse and non-abuse categories that can be difficult and can be wrong. And so that throws the entire finding into question. The methodological challenge here is the circularity challenge. When you're trying to assess the value of a diagnostic test or a diagnostic criteria, and you include as inclusion criteria the very critical findings that are being studied, you can't have confidence in the reliability of the result.

---

37. *See generally* Gordon D. Schiff, *Minimizing Diagnostic Error: The Importance of Follow-up and Feedback*, 121 AM. J. MED. S38 (2008) (discussing an open-loop system in clinical diagnoses—or "nonfeedback controlled" system where one makes decisions based solely on preprogrammed criteria and the preexisting model of the system).

38. *Id.* at S39.

39. *Id.* at S41.

Some of the recent research recognizes this.[40] Piteau says, "As there are no standardized criteria for the definition of abuse, most authors developed their own criteria, and many of these are fraught with circular reasoning."[41] Again, that tells us what the limitations of some of the best studies out there are. Diagnostic studies depend on a constellation of clinical features because there is not a gold standard diagnostic test. Any time there's a lack of a test, a lack of objective criteria, we allow bias to seep into the system. Maguire's research is in a similar vein and is significant because it acknowledges the uncertainty of even the best data available.[42]

I want to touch a little bit on solutions. We've talked a lot about the problems; I don't have all the answers. But we have some ideas about how we might prevent wrongful conviction in child abuse cases. One is overcoming cognitive biases. And certainly those who study psychology can address this better than I can. But one of the things we want to talk about is bias awareness in education. You need to know that you have bias. Bias is a thing we all need to know that we carry around. Implicit bias is something that we have. But just knowing about it doesn't help, right? You can't will it away.

And so there are a few things you can try to do. Asking individuals to consider and articulate their hypothesis's opposite can mitigate bias.[43] Asking people to articulate reasons that counter their own position can minimize the illusion of validity underlying conformation bias.[44] Discussing the evidence for and against their hypothesis can reduce bias.[45] Asking people to delay hypothesis formation can reduce bias.[46] None of these will solve the problem, but they're all important to think about.

Greater transparency is a big deal in this and other forensic sciences and increasing transparency may lead to better outcomes. To that end, increased electronic recording of interrogations and other contacts between investigators and people suspected of crimes may be warranted.[47] Full discovery in criminal cases—where

---

40. Shalea J. Piteau et al., *Clinical and Radiographic Characteristics Associated with Abusive and Nonabusive Head Trauma: A Systematic Review*, 130 PEDIATRICS 315, 321 (2012).

41. *Id.*

42. Sabine Ann Maguire et al., *Estimating the Probability of Abusive Head Trauma: A Pooled Analysis*, 128 PEDIATRICS e550, e558 (2011) (noting that "[d]iagnostic studies in the field are open to criticism of circularity because of their dependence on a constellation of clinical features, as opposed to a single gold-standard diagnostic test, which does not exist").

43. *See* Findley & Scott, *supra* note 18, at 382.

44. *Id.* at 388–89.

45. *Id.* at 388.

46. *Id.*

47. *See* David Oswald, Fred Sherratt & Simon Smith, *Handling the Hawthorne Effect: The Challenges Surrounding a Participant Observer*, *in* REVIEW OF SOCIAL STUDIES: METHODOLOGICAL

defense lawyers have access to open files—and increased collaboration between the government and the defense also head in the right direction. When decisions about prosecutions are made in open and observable ways, investigators and prosecutors are more able to resist biasing pressure.[48]

Proper investigation may help in these cases. And I think it makes a big difference. But a lot of significant flaws will still exist. These difficulties are exaggerated by an inadequate evidence base and imprecise testing. Clearly, in innocence work, we have a sampling bias; cases do not come to me unless there is powerful evidence of innocence. For the most part, complete, objective investigations were not done in the cases my office accepts, though incomplete investigations may have been complicated by poor lawyering, misconduct, incorrect claims in testimony, or evidence discovered after trial. Leigh Bishop, Chief of the Child Fatality Unit at the Queens County District Attorney's Office is going to talk about one of her cases that included an exhaustive investigation, after which no one was charged with a crime; I don't see that in my cases. Investigation can make a big difference, but it is not perfect. It's important to recognize that even a really good investigation is not immune to bias. Even if all appropriate tests are performed properly, even if physicians took unassailable histories, even if police investigated every alternate suspect, even if everyone involved did everything correctly, there's still a place for error to enter the system.

Outcomes cannot be improved unless actors recognize that a problem exists. Recognizing that wrongful convictions can and do happen in child abuse cases, recognizing that there's bias in the literature, that there's potential for error; these are all crucial parts of improving outcomes. It's particularly crucial for all actors in the system to recognize that false negatives (failing to recognize abuse when it is present) and false positives (concluding abuse when abuse is not present) are equally unacceptable outcomes. Some experts have suggested that overdiagnosing abuse is preferable because protection of children is paramount. But failing to properly diagnose a child is equally harmful and must be recognized as such. Failing to make a proper diagnosis has serious and sometimes irreparable consequences for children and families, which can include, but are not limited to, separation of children from loving

---

CHOICES AND CHALLENGES 55 (Latif Tas ed., 2014) (describing the Hawthorne Effect, in which individuals change their behavior because they are aware they are being observed).

48. Findley & Scott, *supra* note 18, at 391.

parents and serious medical consequences when children do not receive proper treatment as the result of a misdiagnosis.

Even a criminal accusation that results in a dismissal or an acquittal can destroy family relationships, can financially ruin a family, and can be really significant and damaging in other ways.[49] So one way to improve outcomes might be to recognize false positives and false negatives on the same footing, right on the same plane, mitigating bias.

What does not accomplish mitigating biases is to say, "I was not biased." Frequently, experts claim that they did not rush to judgment or that they have no bias. Claiming not to have them does nothing to mitigate the biases that every single person has. We want to take lessons from other forensic sciences, because there are forensic sciences that can be tested or parts of them that can be tested, and that we can learn lessons from. Then you want appropriate adversarial testing of claims, not allowing false testimony to come out in front of a jury.

There are a lot of similarities between child abuse investigation and fire science investigation. For a long time, investigations of fires relied on unsupported rules and hypotheses: if you saw certain features of a fire, you could assume that it was caused by some sort of incendiary means.[50] Those have been shown to be false, as has any idea of negative corpus—that if you rule out what you perceived to be all possible accidental causes of the fire—that you could reliably determine that the fire was arson.[51] And those of us who've worked on a lot of abuse cases can recognize that that's often something that happens. So this is from the NFPA report from 2009.[52] They, in their summary, found that many rules of thumb about fire causation typically assumed to indicate that an accelerant was used, like "alligatoring" of wood, have been shown not to be true.[53] You can draw a nice parallel between these indicators of abuse that were long thought to be exclusively diagnostic of abuse that are no longer thought to be so.

---

49. Rachel Bluestain, *Doctors Said They Shook Their Baby To Death. They Didn't*, DAILY BEAST (Oct. 3, 2016), http://www.thedailybeast.com/articles/2016/10/03/doctors-said-they-shook-their-baby-to-death-they-didn-t.html.

50. *See generally,* Caitlin Plummer and Imran Syed, *Shifted Science and Post-Conviction Relief,* 8 STAN. J. CIV. RTS. & CIV. LIBERTIES 259 (2012).

51. *See* NATIONAL FIRE PROTECTION ASSOCIATION (NFPA), NFPA 921—GUIDE FOR FIRE AND EXPLOSION INVESTIGATIONS § 19.6.5 (7th ed. 2014).

52. *Id.*

53. *Id.*

Then there's this tidy parallel between negative corpus and with the process of elimination in diagnosis. Negative corpus is the process of determining the ignition source for a fire by eliminating all suspected sources and then "claiming such methodology is proof of an ignition source for which there is no evidence of its existence."[54] This was something that went on for a long time, but now the NFPA, the National Fire Protection Association, calls the practice unscientific.[55] The NFPA's 2014 manual says, "[Negative corpus] is not consistent with the Scientific Method, is inappropriate, and should not be used because it generates un-testable hypotheses, and may result in incorrect determinations of the ignition source and the first fuel ignited."[56] And then they say, "[I]t is improper to opine a specific ignition source that has no evidence to support it, even though all other hypothesized sources were eliminated."[57] Process of elimination arguments rest on a number of assumptions. The negative corpus example assumes that the investigator can correctly identify and rule out all other possible causes of a fire, that the investigator will do all tests correctly, that there's a negligible error rate for each test, and then, perhaps most importantly, that all of the possible sources of the fire are, in fact, on the investigator's list. Put another way, it does not meaningfully allow for the cause to be "undetermined."

This concern also applies in the child abuse context. Experts may claim that abuse is a "rule-out" diagnosis, and that it can safely be made once the physician has "ruled out" every other hypothesized cause of the findings. Or they may answer in the affirmative questions like, "to a reasonable degree of medical certainty, were you able to rule out *all other potential causes* for the injuries?"[58] Those statements are problematic in the same way that statements about negative corpus are. And that's where we can take the lessons from other forensic sciences and apply them in this context.

The Wisconsin Innocence Project is leading a fundraising effort for an independent scientific panel to review the evidence base for SBS/AHT and other science-dependent child abuse prosecutions. An independent review, conducted by impartial experts, is crucial to getting it right. This effort, however, is ongoing and expected to take some time. So in the meantime, what steps can we take?

---

54. *Id.*

55. *Id.*

56. *Id.* at § 18.6.5.

57. *Id.* at § 18.6.5.1.

58. *See, e.g.*, Transcript of Record at 764–65, People v. Calderaro, No. 1251-06 (N.Y. Sup. Ct. April 1, 2009) (transcript on file with author).

We can start by recognizing the problem. We can recognize that wrongful convictions exist, and that they can be ambiguous. But some wrongful convictions have known, specific causes that can be addressed. In science-dependent child abuse prosecutions, bias and subjectivity can affect diagnosis decisions, which may in turn affect prosecution decisions. We need to recognize that false negatives and false positives are equally problematic. Children who are not recognized as abused can continue to suffer abuse, sometimes escalating abuse. This is tragic. It is equally tragic when an abuse diagnosis is made incorrectly and families are separated wrongfully, resulting in the suffering of all involved. Getting it right in these cases requires a comprehensive and objective medical assessment. It also requires more research and information to improve the diagnostic process. It requires changes in the process that mitigate bias. Getting it right further requires responsible and accurate testimony. When a proposition made in court is untestable or unknowable, witnesses have an obligation to say so. Witnesses must not testify with more certainty than the evidence supports, or speculate when research is lacking.

So given what we know now about the science behind SBS and other kinds of child abuse, we know that medical findings often associated with shaking are not pathognomonic of abuse;[59] the same is true of fractures, bruises and other medical findings.[60]

In order to combat known causes of wrongful conviction and ensure appropriate an accurate testimony in science-dependent prosecutions, it is crucial to limit bias and subjectivity in the investigation and prosecution of these cases to the extent possible.

59. *See, e.g.,* Christian et al, *Abusive Head Trauma in Infants and Children*, 123 PEDIATRICS 206 (2009).

60. *See* Glick et al., *Physical Abuse of Children*, 37 PEDIATRICS REV. 146 (2016).