

University of Maryland Francis King Carey School of Law

DigitalCommons@UM Carey Law

Faculty Scholarship

Francis King Carey School of Law Faculty

2021

Junk Science at Sentencing

Maneka Sinha

University of Maryland Francis King Carey School of Law

Follow this and additional works at: https://digitalcommons.law.umaryland.edu/fac_pubs



Part of the [Criminal Law Commons](#), [Criminal Procedure Commons](#), and the [Evidence Commons](#)

Digital Commons Citation

Sinha, Maneka, "Junk Science at Sentencing" (2021). *Faculty Scholarship*. 1641.

https://digitalcommons.law.umaryland.edu/fac_pubs/1641

This Article is brought to you for free and open access by the Francis King Carey School of Law Faculty at DigitalCommons@UM Carey Law. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of DigitalCommons@UM Carey Law. For more information, please contact smccarty@law.umaryland.edu.

Junk Science at Sentencing

Maneka Sinha*

ABSTRACT

Junk science used in criminal trials has contributed to hundreds of wrongful convictions. But the problem is much worse than that. Junk science does not only harm criminal defendants who go to trial, but also the overwhelming majority of defendants—over ninety-five percent—who plead guilty, skip trial, and proceed straight to sentencing.

Scientific, technical, and other specialized evidence (“STS evidence”) is used regularly, and with increasing frequency, at sentencing. Despite this, Federal Rule of Evidence 702 and its state equivalents—which help filter unreliable STS evidence at trials—do not apply at the critical sentencing stage. In fact, at sentencing, no meaningful admissibility standard guards against junk science deciding punishment. Over ninety-five percent of defendants, therefore, do not get the basic protection against faulty STS evidence that trial defendants get. This may result in harsher sentences based on junk science that has been admitted and considered without any screening or vetting.

This Article offers the first in-depth exploration of STS evidence at sentencing. It links two bodies of literature: the first analyzing the negative effects of junk science on the reliability of trials and the second arguing for extending procedural protections to sentencing. This Article builds upon these literatures by proposing an implementable mechanism for evaluating STS evidence at sentencing while retaining special protections for criminal defendants. The Article recommends that Federal Rule of Evidence 702 (or its state analog) apply at sentencing to determine the admissibility of STS evidence offered in support of harsher punishment, but not to such evidence offered by defendants as mitigation.

* Assistant Professor of Law, University of Maryland School of Law. I am grateful to Zainab Ahmad, Jessica Brand, Nina Chernoff, Erin Murphy, Julia Sheketoff, Anthony C. Thompson, and Kim Taylor-Thompson for thoughtful comments on earlier drafts of this Article. This piece benefited greatly from faculty workshops at the University of Maryland School of Law and the University of Pittsburgh School of Law; from feedback provided by Courtney Cross, Gabriel Diaz, Mark Graber, David Gray, Daniel S. Harawa, Farhang Heydari, Vida B. Johnson, Michael Pappas, Kate Philpott, Lucius T. Outlaw III, Elizabeth Daniel Vasquez, and Kate Weisburd; and from suggestions made by attendees of the Clinical Law Review Writers' Workshop. An earlier version of this Article was selected for presentation at the ABA-AALS Criminal Justice Section Academic Roundtables, and I am thankful to attendees of that workshop for their helpful comments. Thanks also to Susan McCarty, Managing Research Fellow at the University of Maryland School of Law, and to Summer Akhtar, Allison Friedman, Diana Hanson, and Kelsey Robinson for valuable research assistance. Finally, I owe tremendous thanks to T.K. and to Rachel Cicurel, whose litigation inspired this piece and whose feedback was invaluable.

TABLE OF CONTENTS

INTRODUCTION 53

I. SCIENTIFIC, TECHNICAL, AND OTHER SPECIALIZED EVIDENCE AT SENTENCING 58

 A. *The Significance of the Sentencing Stage* 58

 1. The Impact of Junk Science in Criminal Cases . 58

 2. The Importance of Sentencing in Criminal Cases..... 63

 B. *Scientific, Technical, or Specialized Evidence at Sentencing* 66

 1. The Typical Sentencing Hearing 67

 2. Scientific, Technical, and Specialized Evidence at Sentencing 69

 a. *Predictive Tools and Aggravating and Mitigating Evidence* 69

 b. *Traditional Forensic Identification Methods* . 78

 C. *Comparing Trial-Stage and Sentencing-Stage Standards for Admission of Scientific, Technical, or Specialized Evidence* 80

 1. The Trial Phase: The Rule 702 Standard 81

 2. The Sentencing Phase: The Due Process Floor . 87

II. EXTENDING RULE 702 TO SENTENCING 92

 A. *The Proposed Procedure* 93

 B. *Analysis of the Proposed Extension of Rule 702* 95

 1. Asymmetry 96

 2. Cost, Delay, and Other Practical Concerns 100

 3. The Effectiveness of Rule 702 and *Daubert* at Trial..... 101

III. OTHER PROPOSALS TO IMPROVE SENTENCING ACCURACY 102

 A. *Proposals to Extend Rules of Evidence to Sentencing* 102

 B. *Other Proposals to Improve Sentencing Accuracy* ... 104

 C. *The Suitability of These Proposals* 106

CONCLUSION 110

INTRODUCTION

In 2018, with his sentencing hearing just around the corner, T.K.,¹ a seventeen-year-old in Washington, D.C., appeared set to head home

¹ T.K. has given the author permission to disclose certain facts about his case. At the time

for a period of probation. T.K. had pleaded guilty² to a felony, and all relevant parties—including the prosecutor and the juvenile probation agency—agreed that incarceration was unnecessary. He was a conscientious student on the cusp of high school graduation, he was surrounded and supported by a close-knit family, and he had, by all accounts, taken every opportunity to turn his life around. As a result, it appeared that the only issue left to be resolved at T.K.'s sentencing was not whether T.K. should be sentenced to a period of probation, but rather how long that period should be.

But the day before sentencing, the parties received the results of a violence risk assessment, the Structured Assessment for Violence Risk in Youth ("SAVRY"), that had been conducted as part of a routine psychological evaluation ordered by the court.³ The SAVRY purports to predict the likelihood of future violence or reoffending in adolescents.⁴ According to the assessment, T.K. was a high risk for committing future violence.⁵ It was a life-altering report that drastically changed T.K.'s prospects at sentencing.

The assessment was flawed, however. The SAVRY rates twenty-four risk factors supposedly associated with a juvenile's risk of violent reoffending.⁶ An evaluator considers each of these risk factors and assigns a rating of low, moderate, or high risk to each.⁷ Low risk translates to a numerical value of zero, moderate to one, and high to two.⁸ Once all factors are evaluated, the evaluator assigns a total risk rating.⁹

of his adjudication, T.K. was a juvenile. In consideration of his privacy, T.K.'s full name and case number have been withheld. All records relating to T.K.'s case are on file with the author.

² Because he was a juvenile, T.K. was charged in a juvenile delinquency case within the Family Division of the District of Columbia Superior Court, rather than the Criminal Division, where his case would have been heard had he been an adult at the time of his arrest. The juvenile delinquency equivalent of a sentencing hearing is a "disposition hearing." See D.C. SUPER. CT. JUV. R. 32. For reader ease, traditional criminal case terminology instead of juvenile delinquency-specific terminology is used throughout this article when referring to T.K.'s case.

³ See Order at 2, *In re T.K.* (D.C. Super. Ct. Mar. 15, 2018).

⁴ See Gina M. Vincent, Laura S. Guy, Samantha L. Fusco & Bernice G. Gershenson, *Field Reliability of the SAVRY with Juvenile Probation Officers: Implications for Training*, 36 LAW & HUM. BEHAV. 225, 226 (2012).

⁵ Order, *supra* note 3, at 3.

⁶ RANDY BORUM, PATRICK BARTEL & ADELLE FORTH, *MANUAL FOR THE STRUCTURED ASSESSMENT OF VIOLENCE RISK IN YOUTH* 7 (version 1 consultation ed. 2002). Examples include "History of Violence," "Peer Delinquency," and "Substance Use Difficulties." *Id.* at 28, 58, 84. The SAVRY also considers "protective factors" that mitigate against risk. *Id.* at 9–10.

⁷ *Id.* at 17.

⁸ *Id.* at 22.

⁹ See *id.* at 17.

But, in T.K.'s case, the examiner misapplied the tool. First, she improperly double-counted certain behaviors by listing them in multiple categories, an inappropriate method of risk evaluation using the SAVRY.¹⁰ Second, the evaluator made an arithmetic error: she miscalculated the total number of elevated risk factors, assigning T.K. more elevated risk factors than were present in her own assessment.¹¹ To top it off, the evaluator incorrectly assessed T.K.'s risk of future dangerousness as "high," when her own data justified only a "low," or at most, "moderate" risk rating.¹² In short, she never should have found that T.K. was a high risk of future dangerousness,¹³ but nevertheless, T.K.'s future changed overnight.

T.K.'s case is not atypical: judges regularly rely on risk assessments and other flawed scientific evidence at sentencing hearings.¹⁴ This is particularly troubling because today, the overwhelming majority of criminal defendants plead guilty and proceed to sentencing without ever having a trial.¹⁵ Their cases are functionally resolved at sentencing,¹⁶ where there is no meaningful way to screen out junk science and prevent judges from relying on the type of flawed evidence supplied in T.K.'s case.

After reviewing the SAVRY results, both the prosecutor and probation agency rescinded their recommendations for T.K. to be placed on probation and, instead, asked for him to be committed to the custody of the local youth rehabilitation agency.¹⁷ Commitment would mean that, rather than being free to attend one of the colleges to which he had been admitted, T.K. could spend years at a secure detention facility—a euphemistic term for what is, in effect, prison for children.¹⁸

¹⁰ See Order, *supra* note 3, at 5, 8. An example of double-counting might include listing a school fight in considering multiple risk factors, such as both the history of violence risk factor and the risk taking/impulsivity risk factor rather than listing it for one or the other. See *id.*

¹¹ Motion to Exclude Results of the Violence Risk Assessment and All Related Testimony and/or Allocation Under FRE 702 and Daubert v. Merrell Dow Pharmaceuticals at 13, *In re T.K.* (D.C. Super. Ct. Feb. 5, 2018) [hereinafter Motion to Exclude].

¹² See *id.* at 13–14.

¹³ See Order, *supra* note 3, at 8–9.

¹⁴ See discussion *infra* Section I.B.

¹⁵ See Clark Neily, *Jury Empowerment as an Antidote to Coercive Plea Bargaining*, 31 FED. SENT'G REP. 284, 284 (2019) ("[M]ore than 95 percent of criminal convictions are obtained through guilty pleas . . .").

¹⁶ See Jenny Roberts, *Informed Misdemeanor Sentencing*, 46 HOFSTRA L. REV. 171, 177 (2017).

¹⁷ See Motion to Exclude, *supra* note 11, at 2.

¹⁸ See D.C. Mun. Regs. tit. 29, § 1200.4 (2017) (giving the Department of Youth Rehabilitation Services ("DYRS") "sole discretion to make specific placement decisions for youth com-

T.K. sought to challenge the admissibility of the risk assessment under Federal Rule of Evidence 702, which governs the admissibility of expert evidence at trial.¹⁹ He argued, through an expert in psychology and risk assessments, that the risk assessment in question is unreliable in predicting future violence and that, even if the tool were generally reliable, it is not reliable as applied to his case.²⁰

But as T.K.'s judge noted, Rule 702 does not apply at sentencing.²¹ As a result, T.K.'s prospects for leniency were seriously diminished as the prosecutor and probation department requested a harsher sentence than they had sought prior to reading the unreliable report.²²

There is widespread agreement that evidentiary rules applied at trial, including expert admissibility tests, are important in that they aid in the search for truth and promote fair outcomes.²³ But because Rule 702 and other rules of evidence do not apply at sentencing,²⁴ judges often make extremely consequential punishment decisions based on unreliable or untested evidence. All over the country, in federal jurisdictions and the many states that have adopted Rule 702 or a similar rule, criminal defendants find themselves in the same shoes that T.K. did in 2018. At sentencing, defendants are vulnerable to the use of flawed scientific evidence in a way they would not be at trial, where Rule 702 and similar state rules serve to filter out unreliable scientific, technical, and specialized evidence ("STS evidence"), or junk science.²⁵ STS evidence may constitute junk science because (1) the un-

mitted to its custody"); see also *Committed Youth*, D.C. DEP'T OF YOUTH REHAB. SERVS., <http://dysr.dc.gov/service/committed-youth> [<https://perma.cc/8YJY-UH2M>] (allowing DYRS to place committed youth in secure facilities).

19 FED. R. EVID. 702. The Federal Rules of Evidence govern proceedings in federal courts. FED. R. EVID. 101. At the time of T.K.'s case, Rule 702 had recently been adopted by the District of Columbia as its expert admissibility standard. *Motorola Inc. v. Murray*, 147 A.3d 751, 752 (D.C. 2016) (en banc).

20 Order, *supra* note 3, at 3–4.

21 See *id.* at 7.

22 See Motion to Exclude, *supra* note 11, at 2.

23 See, e.g., FED. R. EVID. 102 (explaining that the Federal Rules of Evidence "should be construed so as to administer every proceeding fairly . . . to the end of ascertaining the truth and securing a just determination").

24 See, e.g., FED. R. EVID. 1101.

25 Rule 702 applies to "scientific, technical, or other specialized knowledge." FED. R. EVID. 702; see also *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147–49 (1999) (finding that Rule 702 applies not just to scientific knowledge, but scientific, technical, and other specialized knowledge). For ease, in this Article, scientific, technical, or other specialized evidence is referred to as "STS evidence." Examples of STS evidence may include psychological or psychiatric information contained in presentence investigation reports; DNA, fingerprint, ballistics or other traditional forensic identification evidence; evidence resulting from use of risk-assessment tools; and medical information and other, similarly specialized information. STS evidence that would

derlying science itself is inherently unreliable, (2) an otherwise valid method is misapplied to produce faulty results, or (3) forensic examiners exaggerate results.²⁶ Judicial reliance on such evidence may directly contribute to a criminal defendant spending more days, weeks, years, or even decades in prison.

This Article argues that the frequent use of often flawed STS information at sentencing, coupled with how critical the sentencing stage is to the administration of criminal justice, necessitates greater scrutiny of such evidence than is currently applied in sentencing decisions. It offers the first in-depth exploration of STS evidence at sentencing. It bridges the extensive scholarship on junk science at the trial stage with scholarship advocating for extension of procedural protections to sentencing. And it further builds upon these literatures by proposing an implementable mechanism for evaluating STS evidence at sentencing while retaining special protections for criminal defendants.

This Article proposes that Rule 702 and its state analogs be extended in a modified, asymmetrical format to sentencing. Specifically, it proposes that the same types of evidence that would be subject to Rule 702's admissibility test at the trial stage be subject to the Rule 702 admissibility test at sentencing, when offered by the prosecution or probation officer as support for harsher punishment, but not when offered by a defendant for mitigation purposes.

The Article proceeds in three parts. Part I explores the need for increased examination of the use of STS evidence at sentencing. It first details the role flawed scientific evidence has played in contributing to troubling outcomes at the trial phase. It then lays out the critical importance of the sentencing stage in the modern criminal legal system and describes the ways in which STS evidence is used at sentencings today. Finally, Part I explains the operation of Rule 702 at trial and contrasts the Rule with what little exists in the way of a legal standard to filter out unreliable evidence at sentencing. Part II proposes a mechanism for screening unreliable STS evidence at sentencing that (1) is calculated to improve sentencing accuracy, (2) allows judges to consider a broad range of information, and (3) avoids compromising defendants' ability to present mitigating evidence. It suggests that when STS evidence is offered by the prosecutor or a neutral party in support of an argument for an increased sentence, it be sub-

not meet the admissibility test laid out in Rule 702 is referred to throughout this article as "junk science."

²⁶ See FED. R. EVID. 702(b)–(c).

jected to the admissibility test in Rule 702 or the state equivalent—but that STS evidence offered by a defendant at sentencing need not be subjected to such a test. It then evaluates this proposal on a number of dimensions. Part III analyzes a sampling of previous recommendations to improve sentencing accuracy and increase protections for criminal defendants at sentencing. It then considers their suitability for filtering out junk science at sentencing, concluding that while valuable in their own right, none of these recommendations adequately addresses the problem of junk science at sentencing.

I. SCIENTIFIC, TECHNICAL, AND OTHER SPECIALIZED EVIDENCE AT SENTENCING

This Part provides three layers of background on the need for closer examination of the use of STS evidence at sentencing. It first explains why more attention must be paid to STS evidence at sentencing by outlining the importance of, and the stakes at issue for criminal defendants at, that phase. Second, it outlines the types of STS evidence used in sentencings today. Finally, it describes the gatekeeping role of Rule 702 at trial and compares this with the far more lenient standard governing what information can be considered at sentencing.

A. *The Significance of the Sentencing Stage*

Much attention is paid to the use of STS evidence at the trial stage of criminal cases. This Section begins to build the case that greater attention must be paid to the use of such evidence at sentencing, as sentencing hearings have become the most important phase of our modern plea-bargain-focused criminal legal system.

1. *The Impact of Junk Science in Criminal Cases*

Though wrongful convictions only represent a tiny slice of the pie, they are important to understanding the negative impacts junk science can have in criminal cases. Concrete evidence of the role that flawed STS evidence has played in contributing to miscarriages of justice stems from considerable study of wrongful convictions. That evidence helps shed light on how to approach junk science at sentencing, which represents the vast majority of the pie.

Unreliable STS evidence is a leading cause of a variety of unjust outcomes in criminal cases, including wrongful convictions.²⁷ The first

²⁷ See generally BRANDON L. GARRETT, CONVICTING THE INNOCENT 84–117 (2011) (discussing invalid and unreliable forensic science at trials of wrongly-convicted defendants).

study examining expert testimony in the cases of people who were later exonerated found that flawed expert evidence was involved in *sixty percent* of verified false convictions.²⁸ The National Registry of Exonerations lists “False or Misleading Forensic Evidence” as a contributing factor in twenty-four percent, or a total of 644 documented exonerations.²⁹ The Innocence Project calculates the percentage of cases in which problematic use of expert evidence contributed to exonerations as even higher; it lists misapplication of forensic science as a contributor to forty-three percent of DNA exonerations.³⁰ As troubling as those statistics are, however, they are likely to be underestimates; it is probable that flawed STS evidence has contributed to additional wrongful convictions that have not yet been—and may never be—uncovered.

No one type of flawed STS evidence is responsible for unjust outcomes. An abundance of examples exists across a variety of disciplines of faulty or misapplied expert evidence contributing to troubling outcomes in criminal cases. As described in Section I.C., for example, junk psychology famously contributed to the imposition of a death sentence for Duane Buck, who was convicted of a double homicide that occurred in 1995.³¹ In 2004, Oregon lawyer Brandon Mayfield was wrongfully detained in connection with the 2004 Madrid train bombings on the basis of a false fingerprint identification.³² There are many

²⁸ See Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 1–2 (2009) (finding invalid forensic testimony was offered in 82 out of 137 cases studied).

²⁹ See % *Exonerations by Contributing Factors*, NAT’L REGISTRY OF EXONERATIONS, <http://www.law.umich.edu/special/exoneration/Pages/ExonerationsContribFactorsByCrime.aspx> [https://perma.cc/R5CK-W72H].

³⁰ *DNA Exonerations in the United States*, INNOCENCE PROJECT, <https://www.innocenceproject.org/dna-exonerations-in-the-united-states/> [https://perma.cc/LY9N-X99A].

³¹ See Matthew Haag, *Texas Man’s Death Sentence Thrown Out Over Racist Testimony*, N.Y. TIMES (Oct. 4, 2017), <https://www.nytimes.com/2017/10/04/us/duane-buck-death-sentence.html> [https://perma.cc/VN3W-CUKY]. The sentence was eventually overturned by the United States Supreme Court. See *id.*

³² See U.S. DEP’T OF JUST., OFF. OF THE INSPECTOR GEN., A REVIEW OF THE FBI’S HANDLING OF THE BRANDON MAYFIELD CASE 1 (2006) [hereinafter 2006 OIG REPORT], <https://oig.justice.gov/sites/default/files/legacy/special/s0601/final.pdf> [https://perma.cc/ZDT4-EY2V]. In fact, Mayfield’s print was falsely identified by not one, but three FBI fingerprint examiners and an independent, court-appointed fingerprint expert. See *id.* at 1–3. His wrongful detention led to years-long discussion about the validity of fingerprint examination and comparison in both the legal and scientific communities. See, e.g., 2006 OIG REPORT, *supra*; Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985 (2005); U.S. DEP’T OF JUST., OFF. OF THE INSPECTOR GEN., A REVIEW OF THE FBI’S PROGRESS IN RESPONDING TO THE RECOMMENDATIONS IN THE OFFICE OF THE INSPECTOR GENERAL

additional examples of troubling outcomes based on faulty fingerprint methods.³³

Santae Tribble spent 23 years in prison after a 1980 murder conviction based largely on a claim by an FBI hair examiner that his hair matched a hair found at a crime scene.³⁴ The FBI examiner was wrong; the hair found at the crime scene was a dog hair.³⁵ Hair comparison analysis has produced some of the most widely publicized wrongful convictions;³⁶ in 2015, after years of concern over the validity of hair analysis, the Department of Justice finally admitted that its hair examiners had been giving flawed testimony for over 20 years.³⁷

Even DNA analysis—often thought of as the gold standard of forensic identification methods—is not immune to misuse and misapplication. In 2011, Adam Scott’s DNA was found to match a sperm sample from a rape kit.³⁸ But Scott lived over 200 miles away from—and had never visited—the city in which the rape took place.³⁹ Scott’s was not an isolated incident. A 2018 study by the National Institute of Standards and Technology (“NIST”) found that DNA analysts across labs analyze DNA in vastly different—and often erroneous—ways.⁴⁰

REPORT ON THE FINGERPRINT MISIDENTIFICATION IN THE BRANDON MAYFIELD CASE (2011), <https://www.oversight.gov/sites/default/files/oig-reports/s1105.pdf> [https://perma.cc/4AYK-NXME]; WILLIAM THOMPSON, JOHN BLACK, ANIL JAIN & JOSEPH KADANE, AM. ASS’N FOR THE ADVANCEMENT OF SCI., FORENSIC SCIENCE ASSESSMENTS: A QUALITY AND GAP ANALYSIS, LATENT FINGERPRINT ANALYSIS (2017), https://www.aaas.org/sites/default/files/s3fs-public/reports/Latent%2520Fingerprint%2520Report%2520FINAL%25209_14.pdf [https://perma.cc/Z9VJ-3CV7].

³³ See, e.g., Maurice Possley, *Richard Jackson*, NAT’L REGISTRY OF EXONERATIONS (Mar. 18, 2019), <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3318> [https://perma.cc/BD2Z-Y374]; Jeff Chinn, *Fingerprint Expert’s Mistake Leads to Wrongful Conviction in Indiana*, CAL. INNOCENCE PROJECT (Oct. 18, 2012), <https://californiainnocenceproject.org/2012/10/fingerprint-experts-mistake-leads-to-wrongful-conviction-in-indiana/> [https://perma.cc/5MFJ-87E7].

³⁴ *Santae Tribble*, INNOCENCE PROJECT, <https://www.innocenceproject.org/cases/santae-tribble/> [https://perma.cc/7F8P-PYUY].

³⁵ *Id.* In 2012, Tribble was finally exonerated. *Id.*

³⁶ See, e.g., *Korey Wise*, INNOCENCE PROJECT, <https://www.innocenceproject.org/cases/korey-wise/> [https://perma.cc/7YHU-2LWD].

³⁷ See Spenser S. Hsu, *FBI Admits Flaws in Hair Analysis over Decades*, WASH. POST (Apr. 18, 2015), https://www.washingtonpost.com/local/crime/fbi-overstated-forensic-hair-matches-in-nearly-all-criminal-trials-for-decades/2015/04/18/39c8d8c6-e515-11e4-b510-962fcfabc310_story.html [https://perma.cc/2HYZ-78MF].

³⁸ Naomi Elster, *How Forensic DNA Evidence Can Lead to Wrongful Convictions*, JSTOR: DAILY (Dec. 6, 2017), <https://daily.jstor.org/forensic-dna-evidence-can-lead-wrongful-convictions/> [https://perma.cc/C39M-37H8].

³⁹ *Id.*

⁴⁰ See John M. Butler, Margaret C. Kline & Michael D. Coble, *NIST Interlaboratory Studies Involving DNA Mixtures (MIX05 and MIX13): Variation Observed and Lessons Learned*,

In the study, which asked examiners from 108 labs across the country to analyze various DNA mixtures, analysts found a person's DNA to be present where it was not more than seventy times.⁴¹ DNA mistakes are not confined to governmental studies; there are many real-life cases where misunderstood or misapplied DNA analysis contributed to concerning results in criminal cases.⁴²

There are a myriad of other examples of STS evidence leading to unreliable convictions that cost people decades of their lives.⁴³ Troublingly, each type of STS evidence described above continues to be widely used in criminal cases today.⁴⁴ The role played by flawed expert evidence in false and unreliable convictions has been the subject of increasing national attention by academics, the media, and others. Scholars from the legal and scientific communities have gone to great lengths to bring attention to the role junk science has played in unfair trial outcomes.⁴⁵ As these communities have continued to sound the

FORENSIC SCI. INT'L: GENETICS, NOV. 2018, at 81, 92–93 (discussing the designs, observations, and findings of NIST studies of DNA mixture interpretation in 2005 and 2013).

⁴¹ *Id.* at 87–91; see Greg Hampikian, Opinion, *The Dangers of DNA Testing*, N.Y. TIMES (Sept. 21, 2018), <https://www.nytimes.com/2018/09/21/opinion/the-dangers-of-dna-testing.html> [<https://perma.cc/S9BB-8T8Z>].

⁴² See, e.g., Katie Worth, *Framed for Murder by His Own DNA*, THE MARSHALL PROJECT (Apr. 19, 2018, 7:00 AM), <https://www.themarshallproject.org/2018/04/19/framed-for-murder-by-his-own-dna> [<https://perma.cc/GAK3-Z82U>]; Marina Medvin, *Framed by Your Own Cells: How DNA Evidence Imprisons the Innocent*, FORBES (Sept. 20, 2018, 11:27 AM), <https://www.forbes.com/sites/marinamedvin/2018/09/20/framed-by-your-own-cells-how-dna-evidence-imprisons-the-innocent/#7762bb424b86> [<https://perma.cc/ZT85-5HEN>] (discussing the impact of DNA on Amanda Knox's now-overturned conviction).

⁴³ See, e.g., Radley Balko & Tucker Carrington, *Bad Science Puts Innocent People in Jail—And Keeps Them There*, WASH. POST (Mar. 21, 2018, 6:00 AM), https://www.washingtonpost.com/outlook/bad-science-puts-innocent-people-in-jail—and-keeps-them-there/2018/03/20/f1fffd08-263e-11e8-b79d-f3d931db7f68_story.html [<https://perma.cc/5V87-SN9E>] (describing problems with purported shaken baby syndrome evidence and explaining that bitemark analysis has led to dozens of wrongful convictions).

⁴⁴ See Jules Epstein, *Preferring the “Wise Man” to Science: The Failure of Courts and Non-Litigation Mechanisms to Demand Validity in Forensic Matching Testimony*, 20 WIDENER L. REV. 81, 101–06 (2014) (explaining that courts continue to admit pattern and impression evidence such as latent fingerprint, handwriting, and firearms analysis); see also PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFF. OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 1–14 (2016) [hereinafter PCAST REPORT], https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [<https://perma.cc/TC4M-F8P5>] (considering DNA, fingerprint, and hair analysis among other disciplines in report on validity of forensic science evidence used in “the Nation's legal system”).

⁴⁵ See, e.g., Edward J. Imwinkelried, *Debunked, Discredited, But Still Defended Revising State Post-Conviction Relief Statutes to Cover Convictions Resting on Subsequently Invalidated Expert Testimony*, 48 SETON HALL L. REV. 1095, 1098–100 (2018); M. Chris Fabricant & Tucker Carrington, *The Shifted Paradigm: Forensic Sciences Overdue Evolution from Magic to Law*, 4

alarm about the misapplication of and overreliance on STS evidence, the media, too, has joined the chorus warning of the consequences of admitting junk science evidence at trial.⁴⁶

Even former President Barack Obama entered the growing national conversation about misuse of expert evidence. In a 2017 *Harvard Law Review* commentary published just before he left the White House, the former President described a need to improve forensic sciences and described efforts made by his administration to effect such change.⁴⁷ These efforts included the publication of a landmark report by the President's Council of Advisors on Science and Technology which assessed the validity of a number of forensic science disciplines in 2016.⁴⁸ Its conclusions were jarring: a number of traditional forensic science disciplines lack foundational validity, even some like firearms and toolmark analysis, which have been regularly admitted by courts for decades.⁴⁹

That the need to improve the use of STS evidence in criminal cases is at the forefront of legal and public consciousness is due, in

VA. J. CRIM. L. 1, 93–112 (2016); Garrett & Neufeld, *supra* note 28; Jonathan J. Koehler, *Forensics or Fauxrensic? Ascertaining Accuracy in the Forensic Sciences*, 49 ARIZ. ST. L.J. 1369, 1389–98 (2017); Brandon L. Garrett & M. Chris Fabricant, *The Myth of the Reliability Test*, 86 FORDHAM L. REV. 1559 (2018); Paul C. Giannelli, *Forensic Science: Daubert's Failure*, 68 CASE W. RES. L. REV. 869 (2018); Keith A. Findley, *Reforming the 'Science' in Forensic Science*, WIS. LAW., Nov. 2015, at 32, 33. *See generally* NAT'L RSCH. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NAS REPORT] (reporting the results of a congressionally-mandated study on, and making recommendations to improve, forensic science); PCAST REPORT, *supra* note 44 (reviewing previous research on the validity of forensic science, evaluating scientific validity of common forensic evidence techniques, and making recommendations to improve the validity of forensic evidence).

⁴⁶ *See, e.g., Last Week Tonight with John Oliver: Forensic Science* (HBO Oct. 1, 2017), <https://youtu.be/ScmJvzmDcG0> [<https://perma.cc/AEF7-FR5A>]; Radley Balko, Opinion, *We Need to Fix Forensics. But How?*, WASH. POST (June 20, 2019, 8:00 AM), <https://www.washingtonpost.com/opinions/2019/06/20/we-need-fix-forensics-how/> [<https://perma.cc/PU4J-2S94>]; Hsu, *supra* note 37.

⁴⁷ Barack Obama, Commentary, *The President's Role in Advancing Criminal Justice Reform*, 130 HARV. L. REV. 811, 860–62 (2017).

⁴⁸ PCAST REPORT, *supra* note 44, at 4–6.

⁴⁹ *See id.* at 81–82 (noting that DNA analysis of complex-mixture samples lacks foundational validity); *id.* at 87 (noting that bitemark analysis lacks foundational validity); *id.* at 95–102 (noting latent fingerprint analysis has a “substantial” false positive error rate and cautioning against admission of the results of latent fingerprint analysis without informing jurors of the limitations of the discipline); *id.* at 104–12 (noting that firearms analysis lacks foundational validity); *id.* at 117 (finding “no appropriate empirical studies to support the foundational validity of footwear analysis to associate shoeprints with particular shoes based on specific identifying marks,” though not “evaluat[ing] the foundational validity of footwear analysis to identify class characteristics (e.g., shoe size or make)”).

significant part, to the innocence movement.⁵⁰ The innocence movement has done much to expose the problematic reliance on scientific evidence in criminal trials that resulted in wrongful convictions and to enlist the public in advocating for criminal justice reform.⁵¹ But because the innocence movement focuses on the ways in which errors lead to wrongful convictions, the conversation about flawed forensics generated by that movement necessarily focuses on the trial stage of criminal cases.⁵² The majority of criminal cases, however, never make it to trial. From a numerical perspective, therefore, focusing on wrongful convictions and their causes is focusing on a subset of the fraction of convictions that come after a trial.

2. *The Importance of Sentencing in Criminal Cases*

From *Law and Order* to *Legally Blonde*, we are told in the news and in books, movies, and television shows, that the most important stage of a criminal case is a trial, in which skilled advocates present evidence and make arguments, facts are revealed through robust

⁵⁰ The innocence movement can be broadly described as a coalition of lawyers, journalists, exonerees, and other activists working to free the wrongfully convicted and expose flaws in the criminal justice system that have led to wrongful convictions. See Jenny Roberts, *The Innocence Movement and Misdemeanors*, 98 B.U. L. REV. 779, 780 (2018); Abbe Smith, *In Praise of the Guilty Project: A Criminal Defense Lawyer's Growing Anxiety About Innocence Projects*, 13 U. PA. J.L. & SOC. CHANGE 315, 316–17 (2009–2010); Marvin Zalman & Julia Carrano, *Sustainability of Innocence Reform*, 77 ALB. L. REV. 955, 958, 997 (2013–2014). The key player in the innocence movement is the Innocence Project, which, in more than twenty-five years of work, has exonerated over 350 prisoners through DNA testing. *DNA Exonerations in the United States*, *supra* note 30.

⁵¹ See Daniel S. Medwed, *Innocentrism*, 2008 U. ILL. L. REV. 1549, 1550–51; Andrew M. Siegel, *Moving Down the Wedge of Injustice: A Proposal for a Third Generation of Wrongful Convictions Scholarship and Advocacy*, 42 AM. CRIM. L. REV. 1219, 1219–20 (2005).

⁵² See, e.g., Marvin Zalman & Ralph Grunewald, *Reinventing the Trial: The Innocence Revolution and Proposals to Modify the American Criminal Trial*, 3 TEX. A&M L. REV. 189, 190, 192 (2015) (“Innocence movement scholarship has minutely examined most possible causes of wrongful convictions The study of wrongful conviction narratives suggests that the standard list of wrongful conviction causes does not capture all the ways that justice can miscarry, and that trial processes may have generated or allowed wrongful convictions.”); Garrett & Fabricant, *supra* note 45, at 1559 (“Critics, however, have pointed out that judges continue to routinely admit unreliable evidence . . . including flawed forensic techniques that have contributed to convictions of innocent people later exonerated by DNA testing.”). Generally, criminal cases resulting in convictions comprise three phases: (1) the investigation phase, where law enforcement and prosecutors investigate potential charges against a soon-to-be defendant; (2) the pre-trial and trial phases, which commence once a defendant is arrested; and (3) the sentencing phase. For simplicity, this framework leaves out the variety of appeals and postconviction litigation that may take place after conviction, but which are not relevant to the discussion in this Article.

cross-examination, and jurors search for the truth.⁵³ But this is a myth. Trials occur in very, very few criminal cases.⁵⁴ Indeed, the overwhelming majority of criminal cases—over ninety percent—are resolved not after a trial, but as a result of a guilty plea.⁵⁵ In 2012, the Supreme Court acknowledged “a simple reality” that “[n]inety-seven percent of federal convictions and ninety-four percent of state convictions are the result of guilty pleas”⁵⁶ and noted that “criminal justice today is for the most part a system of pleas, not a system of trials.”⁵⁷ Even for the small slice of criminal defendants who do take their cases to trial, most are convicted and, thus, sentenced.⁵⁸ The result is that the vast majority of criminal cases end with a defendant being sentenced. This means that, based on the numbers alone, sentencing hearings are hugely important to almost all criminal defendants; they are the culminating event in the criminal process for the masses.

But it is not merely a numbers issue. As sentencings have evolved, their importance to criminal defendants has grown significantly. Around the time of the founding, judges had little discretion in determining a sentence once a jury convicted a defendant.⁵⁹ Sentencings were part of trials—the sentence was determined by the offense

⁵³ See EMILY BAZELON, *CHARGED: THE NEW MOVEMENT TO TRANSFORM AMERICAN PROSECUTION AND END MASS INCARCERATION* 134 (2019); Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 YALE L.J. 909, 911 (1992).

⁵⁴ See John Gramlich, *Only 2% of Federal Criminal Defendants Go to Trial, and Most Who Do Are Found Guilty*, PEW RSCH. CTR. (June 11, 2019), <https://www.pewresearch.org/fact-tank/2019/06/11/only-2-of-federal-criminal-defendants-go-to-trial-and-most-who-do-are-found-guilty/> [<https://perma.cc/2FWM-BC7G>] (noting that two percent of federal criminal defendants went to trial in fiscal year 2018 and perhaps an even smaller share of state criminal defendants go to trial); Jed S. Rakoff, *Why Prosecutors Rule the Criminal Justice System—And What Can Be Done About It*, 111 NW. U. L. REV. 1429, 1432 (2017); Emily Yoffe, *Innocence Is Irrelevant*, ATLANTIC (Sept. 2017), <https://www.theatlantic.com/magazine/archive/2017/09/innocence-is-irrelevant/534171/> [<https://perma.cc/ZL9G-KDZS>]; Scott & Stuntz, *supra* note 53, at 911.

⁵⁵ See Gramlich, *supra* note 54 (reporting that ninety percent of federal criminal defendants pled guilty in fiscal year 2018); Jed S. Rakoff, *Why Innocent People Plead Guilty*, N.Y. REV. BOOKS (Nov. 20, 2014), <https://www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/> [<https://perma.cc/ZC5L-99XQ>] (“[I]t is a rare state where plea bargains do not similarly account for the resolution of at least 95 percent of the felony cases that are not dismissed”); see also *Lafler v. Cooper*, 566 U.S. 156, 170 (2012) (“Ninety-seven percent of federal convictions and ninety-four percent of state convictions are the result of guilty pleas.”); Neily, *supra* note 15, at 284 (“[M]ore than 95 percent of criminal convictions are obtained through guilty pleas”).

⁵⁶ *Missouri v. Frye*, 566 U.S. 134, 143 (2012).

⁵⁷ *Lafler*, 566 U.S. at 170.

⁵⁸ Gramlich, *supra* note 54 (reporting that in fiscal year 2018, most federal criminal defendants who went to trial were convicted).

⁵⁹ *Apprendi v. New Jersey*, 530 U.S. 466, 478–79 (2000).

and was, therefore, highly predictable.⁶⁰ Because sentencing hearings were not independent of trials, the same procedural protections applied to both.⁶¹

Modern sentencings are relatively complex, by contrast, and extremely consequential. Because most defendants proceed directly to sentencing via a plea bargain, it is at sentencing hearings where factual findings regarding a defendant's conduct and characteristics are made.⁶² In addition, sentencings have evolved to allow judges significant discretion in fashioning a sentence.⁶³

The entire criminal legal system has seen unparalleled expansion.⁶⁴ The massive explosion of America's prison population over the last several decades has also increased the significance of the sentencing phase. Although the prison population has begun to stabilize,⁶⁵ the number of individuals in prison today is over five times what it was in the 1970s.⁶⁶ The United States houses the most prisoners of any country in the world, and it has the highest incarceration rate of any country in the world.⁶⁷ A whopping seventy percent of convictions result in the loss of liberty.⁶⁸

This prison explosion is largely the result of two key trends that are decided at sentencing hearings: increases in prison admissions and increases in time served in prison.⁶⁹ In the thirty-year period from

⁶⁰ *Id.*

⁶¹ Carissa Byrne Hessick & F. Andrew Hessick, *Recognizing Constitutional Rights at Sentencing*, 99 CALIF. L. REV. 47, 51 (2011).

⁶² Shaakirrah R. Sanders, *Making the Right Call for Confrontation at Felony Sentencing*, 47 U. MICH. J.L. REFORM 791, 818 (2014).

⁶³ See *id.* at 806–08; Hessick & Hessick, *supra* note 61, at 52–53.

⁶⁴ Peter Wagner & Wendy Sawyer, *States of Incarceration: The Global Context 2018*, PRISON POL'Y INITIATIVE (June 2018), <https://www.prisonpolicy.org/global/2018.html>. [<https://perma.cc/FV6D-SXD4>].

⁶⁵ See *Criminal Justice Facts*, SENTENCING PROJECT, <https://www.sentencingproject.org/criminal-justice-facts/> [<https://perma.cc/R4DG-VGAX>].

⁶⁶ See *id.* Under 200,000 people were in prison in 1972 compared to almost 1.5 million as of 2017. See JOHN F. PFAFF, *LOCKED IN: THE TRUE CAUSES OF MASS INCARCERATION—AND HOW TO ACHIEVE REAL REFORM* 1–2 (2017); Wendy Sawyer & Peter Wagner, *Mass Incarceration: The Whole Pie 2020*, PRISON POL'Y INITIATIVE (Mar. 24, 2020), <https://www.prisonpolicy.org/reports/pie2020.html> [<https://perma.cc/5H3P-ME23>] (listing a combined total of over 1.5 million people in federal and state prisons, with 226,000 people in federal prison and 1.291 million in state); *Criminal Justice Facts*, *supra* note 65. Note that the Prison Policy Initiative counts a total of almost 2.3 million people locked up in American state and federal prisons, jails, and other facilities. Wendy Sawyer & Peter Wagner, *supra*.

⁶⁷ Sawyer & Wagner, *supra* note 66.

⁶⁸ Wagner & Sawyer, *supra* note 64.

⁶⁹ See, e.g., NAT'L RSCH. COUNCIL, *THE GROWTH OF INCARCERATION IN THE UNITED STATES* 50–55 (Jeremy Travis et al. eds., 2014) [hereinafter NAS REPORT ON MASS INCARCERA-

1980 to 2010, the *rate* at which criminal defendants were sent to prison increased by up to 350 percent.⁷⁰ In that same time period, sentence *length* increased by up to 238 percent.⁷¹

The takeaway is simple: sentencing has become such a critical phase in criminal prosecutions not just because so many cases end with sentencing, but also because sentencing outcomes are much harsher today than they were fifty years ago. Criminal defendants are more likely to be sent to prison at sentencing, and they may spend much longer there than in years past.⁷²

There is more at stake for criminal defendants at sentencing than just prison time. Convictions and the sentences that result may generate a whole slate of secondary consequences too: loss of jobs, income, and access to benefits and voting rights; displacement of families; reduced access to health care; stunted education, not to mention stigmatization and much more.⁷³

B. Scientific, Technical, or Specialized Evidence at Sentencing

The focus on the use of flawed scientific evidence on trial outcomes would be one matter if this type of information were presented only at trial. But STS evidence is regularly presented at sentencing hearings. This Section begins by briefly describing the typical sentencing hearing. It then categorizes varieties of STS evidence considered at sentencings today and examines how they are treated at that stage.

TION]; STEVEN RAPHAEL & MICHAEL A. STOLL, WHY ARE SO MANY AMERICANS IN PRISON? 26, 27 (2013).

⁷⁰ The National Research Council reports the rate at which defendants were committed to state prison for murder rose by more than 120% in that period. NAS REPORT ON MASS INCARCERATION, *supra* note 69, at 50. The rates at which defendants were committed to prison for sexual assault, aggravated assault, and drug crimes all rose by over 200%: 275% increase for sexual assault, 250% increase for aggravated assault, 350% increase for drug offenses. *Id.* at 50–51.

⁷¹ *Id.* at 53. Average time served for murder rose from five years in 1981 to almost 17 years in 2000 (a 238% increase). *Id.* Average time served for other crimes rose as well, but not by as much. *Id.*

⁷² See, e.g., RAPHAEL & STOLL, *supra* note 69, at 50–51.

⁷³ See generally MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS 144–69 (rev. ed. 2012) (describing the wide variety of collateral consequences associated with convictions and loss of liberty). There is a vast body of literature on the subject of collateral consequences of a conviction. For an in-depth examination of these consequences, see generally AM. BAR ASS'N CRIM. JUST. SECTION, COLLATERAL CONSEQUENCES OF CRIMINAL CONVICTIONS: JUDICIAL BENCH BOOK (2018), <https://www.ncjrs.gov/pdffiles1/nij/grants/251583.pdf> [<https://perma.cc/PBR4-GZ2U>]; Michael Pinard, *Collateral Consequences of Criminal Convictions: Confronting Issues of Race and Dignity*, 85 N.Y.U. L. REV. 457 (2010).

1. *The Typical Sentencing Hearing*

Sentencing hearings follow a conviction, whether after a trial or entry of a guilty plea. Typically, before the hearing, a governmental agency conducts a presentence investigation that includes an interview of the defendant.⁷⁴ Following the investigation, a report containing background information such as employment, educational and family history, contacts with the criminal legal system, and more is provided to the court and parties.⁷⁵ The report may also contain a sentencing guidelines calculation, if applicable, and may make a sentencing recommendation to the court.⁷⁶ These reports increasingly also contain STS information in the form of mental health information and the results of risk assessments, a growing trend reflected in T.K.'s case.⁷⁷

Sentencing hearings tend to be informal as compared to trials or other robust evidentiary hearings.⁷⁸ The parties may supply letters or reports about the defendant's character, history, and other topics relevant to the punishment determination.⁷⁹ After the parties supply evidence and make arguments, a judge imposes the sentence.⁸⁰ Judges tend to have significant leeway in imposing sentences.⁸¹ Even in jurisdictions that utilize some form of sentencing guidelines, judges typically retain significant discretion in sentencing decision making.⁸² This discretion is by design.⁸³

⁷⁴ See Kimberly A. Thomas, *Sentencing: Where Case Theory and the Client Meet*, 15 CLINICAL L. REV. 187, 197 (2008).

⁷⁵ See *id.* at 196–97; 18 U.S.C. § 3552 (“A United States probation officer shall make a presentence investigation of a defendant that is required pursuant to the provisions of Rule 32(c) of the Federal Rules of Criminal Procedure”); FED. R. CRIM. P. 32 (noting that a presentence report must include an interview of the defendant as well as information on the defendant's prior criminal record, financial condition, and other characteristics); see also, e.g., *Presentence Investigation*, U.S. PROB. OFF., S. DIST. IND., <https://www.insp.uscourts.gov/presentence-investigation> [<https://perma.cc/7JPK-WXXJ>] (noting that the probation officer typically gathers information on family history, educational background, employment history, and more).

⁷⁶ See Thomas, *supra* note 74, at 196–97.

⁷⁷ See Erin Collins, *Punishing Risk*, 107 GEO. L.J. 57, 57–58 (2018).

⁷⁸ See Thomas, *supra* note 74, at 205 (“Sentencing hearings are not as constrained by the careful presentation of evidence that characterizes a trial. In fact, they often seem like free-for-alls.”).

⁷⁹ See Deborah Young, *Fact-Finding at Federal Sentencing: Why the Guidelines Should Meet the Rules*, 79 CORNELL L. REV. 299, 311, 314 (1994).

⁸⁰ See, e.g., *id.* at 304.

⁸¹ See, e.g., Thomas, *supra* note 74, at 206.

⁸² See *id.*; Lindsey Webb, *Slave Narratives and the Sentencing Court*, 42 N.Y.U. REV. L. & SOC. CHANGE 125, 135–36 (2018).

⁸³ See *infra* Section I.C. Capital sentencings proceed differently. Capital trials are split into two stages: the guilt-or-innocence phase (the equivalent of the trial in ordinary, non-capital

At the typical sentencing, witnesses may never take the witness stand; even if they do, their testimony is not confined by trial-stage evidentiary procedures and most often, they are not subject to cross-examination.⁸⁴ This is not just a feature of practice; sentencing hearings are largely informal because procedural and evidentiary rules at this stage are lax or lacking entirely.⁸⁵ Hearsay and character evidence are admissible; rules of evidence do not apply.⁸⁶

Most trial stage constitutional protections do not apply at sentencing either. At trial, for example, the Due Process Clauses of the Fifth and Fourteenth Amendments require that a criminal defendant's guilt be proven beyond a reasonable doubt.⁸⁷ But at sentencing, the government is not required to prove its case for punishment beyond a reasonable doubt.⁸⁸ Rather, its burden is the lowest it can be in criminal cases—a preponderance of the evidence.⁸⁹ The Constitution also guarantees a right to trial by jury; there is no constitutional right to be

cases) and the penalty phase, in which a jury has already found the defendant guilty but is deciding whether to impose a death sentence. See Deborah W. Denno, *The Myth of the Double-Edged Sword: An Empirical Study of Neuroscience Evidence in Criminal Cases*, 56 B.C. L. REV. 493, 502 (2015). In most jurisdictions, the jury must consider both aggravating and mitigating evidence during the penalty phase. *Id.* Aggravating factors typically must outweigh mitigating factors before a death sentence can be imposed. *Id.* In general, an aggravating factor is a fact or circumstance relating to the offense or the defendant that would justify imposition of the death penalty. Carissa Byrne Hessick, *Why Are Only Bad Acts Good Sentencing Factors?*, 88 B.U. L. REV. 1109, 1125 (2008). A mitigating factor is a fact or circumstance, usually relating to the defendant's character or culpability, that justifies withholding a death sentence. *See id.* In capital states, specific aggravating circumstances must be proven before death can be imposed. *See* Robert Alan Kelly, *Applicability of the Rules of Evidence to the Capital Sentencing Proceeding: Theoretical & Practical Support for Open Admissibility of Mitigating Information*, 60 UMKC L. REV. 411, 459 (1992). Moreover, unlike in noncapital sentencings, the burden of proof in the penalty phase of a death cases is the trial-stage standard of proof beyond a reasonable doubt. *See id.*

⁸⁴ See John G. Douglass, *Confronting Death: Sixth Amendment Rights at Capital Sentencing*, 105 COLUM. L. REV. 1967, 1976 (2005) (“[T]he Court has held that due process at capital sentencing does not require the trial-like procedure of open-court testimony by witnesses subject to cross-examination.”); Shaakirrah R. Sanders, *The Value of Confrontation as a Sentencing Right*, 25 WIDENER L.J. 103, 127–28 (2016) (noting that counsel at sentencing hearings are “unable to cross-examine testimonial statements that are material to punishment, which ties defense counsel’s hands and leaves the defendant with no meaningful opportunity to test the evidence that is material to punishment”).

⁸⁵ See Thomas, *supra* note 74, at 205.

⁸⁶ Nancy Gertner, *Neuroscience and Sentencing*, 85 FORDHAM L. REV. 533, 541–43 (2016).

⁸⁷ See, e.g., *In re Winship*, 397 U.S. 358, 361–64 (1970).

⁸⁸ See Alan C. Michaels, *Trial Rights at Sentencing*, 81 N.C. L. REV. 1771, 1833–34 (2003).

⁸⁹ *McMillan v. Pennsylvania*, 477 U.S. 79, 91–93 (1986) (concluding that the state need only prove facts by a preponderance of the evidence at the sentencing stage to satisfy due process). *But see* *Apprendi v. New Jersey*, 530 U.S. 466, 490 (2000) (holding that any fact that can increase a sentence beyond the statutory maximum, other than a fact of a prior conviction, must be decided by a jury and proven beyond a reasonable doubt).

sentenced by a jury.⁹⁰ And, under the Sixth Amendment, criminal defendants enjoy the right to confront the witnesses against them at that trial, but this right does not apply at sentencing.⁹¹

2. *Scientific, Technical, and Specialized Evidence at Sentencing*

The lack of trial-stage protections makes sentencing hearings ripe for admission of unreliable or unvalidated scientific evidence.⁹² The major categories of STS evidence offered at sentencings today are described next.

a. *Predictive Tools and Aggravating and Mitigating Evidence*

EVIDENCE-BASED TOOLS. T.K.'s case reflects a growing trend towards use of risk assessment tools at sentencing.⁹³ Risk assessments aim to provide objective, accurate, data-driven predictions of an individual's likelihood to engage in future criminal behavior.⁹⁴ A movement to reduce the prison population and recidivism⁹⁵ gave rise to a

⁹⁰ *McMillan*, 477 U.S. at 93 (“[T]here is no Sixth Amendment right to jury sentencing . . .”); see Michaels, *supra* note 88, at 1813.

⁹¹ See Michaels, *supra* note 88 at 1835–42 (discussing how “the Supreme Court has never decided whether the Confrontation Clause applies at sentencing” but noting existing limitations on its application in the sentencing context); Sanders, *supra* note 62, at 793 (“[There is] reason to reconsider the applicability of confrontation principles at felony sentencing . . .”).

⁹² A useful illustration is found in *People v. Collins*, 438 P.2d 33 (Cal. 1968) (in bank). In that case, the prosecution offered baseless statistical evidence about the likelihood of the defendants' involvement in the alleged crimes. *Id.* at 38. The court found that the evidence, which lacked an evidentiary foundation as well as a mathematical foundation, “distracted the jury from . . . weighing the evidence . . . , encouraged the jurors to rely upon an engaging but logically irrelevant expert demonstration, foreclosed the possibility of an effective defense . . . , and placed the jurors and defense . . . at a disadvantage in sifting relevant fact from inapplicable theory.” *Id.* Though the evidence in question was presented to a jury during a trial, *id.* at 33, it demonstrates the power of persuasion that even utterly baseless STS evidence can have when it is admitted without scrutiny.

⁹³ See Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 809 (2014); John Monahan & Jennifer L. Skeem, *Risk Assessment in Criminal Sentencing*, 12 ANN. REV. CLINICAL PSYCH. 489, 495–96 (2016); see also Matt Henry, *Risk Assessment: Explained*, APPEAL (Mar. 25, 2019), <https://theappeal.org/risk-assessment-explained/> [<https://perma.cc/9T8P-AAZS>] (providing an in-depth explanation of risk assessments). Analysis of risk assessments for those purposes is beyond the scope of this Article. Risk assessments are also frequently used in making pretrial detention, bail, probation, and parole decisions. See Collins, *supra* note 77, at 57–58.

⁹⁴ See Collins, *supra* note 77, at 59; see also Jessica M. Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59, 62 (2017) (“Advocates contend that, because risk tools more objectively and consistently predict the likelihood of recidivism than the inevitable human guesswork of judges, using the tools at sentencing will improve accuracy.” (footnote omitted)).

⁹⁵ See Monahan & Skeem, *supra* note 93, at 489; Roger K. Warren, *Evidence-Based Sentencing: The Application of Principles of Evidence-Based Practice to State Sentencing Practice and Policy*, 43 U.S.F. L. REV. 585, 586 (2009).

desire for increased use of “Evidence-Based Sentencing,” or sentencing practices based on individualized rehabilitation plans that rely on the conclusions of risk assessment tools.⁹⁶ At least 20 states utilize risk assessment determinations in sentencing decision making.⁹⁷ In some jurisdictions, risk assessments are ordered in every case that proceeds to sentencing.⁹⁸

Risk assessment tools are generally actuarial or statistical in nature; historical or group data about the behavior of arrested or convicted individuals is used to determine the likelihood of a particular individual recidivating.⁹⁹ Risk assessment tools primarily come in one of two styles: structured professional judgment tools (“SPJs”), like the SAVRY, and instruments that utilize a computer software program to predict risk.¹⁰⁰ An SPJ is essentially a guided interview aimed at assessing risk.¹⁰¹ The evaluator conducts an assessment using a checklist of factors that theoretically correlate to risk.¹⁰² Typically, a score or rating is assigned for each factor,¹⁰³ and a total risk judgment is made based on the ratings assigned for each factor.¹⁰⁴ Software-based tools, sometimes referred to as algorithmic risk assessments, eliminate the need for a human evaluator. They operate as a “black box”; a com-

⁹⁶ See Warren, *supra* note 95, at 585–87; Starr, *supra* note 93, at 809.

⁹⁷ Collins, *supra* note 77, at 60; Starr, *supra* note 93, at 809.

⁹⁸ See, e.g., KY. REV. STAT. ANN. § 532.007 (West 2011); Monahan & Skeem, *supra* note 93, at 495; Starr, *supra* note 93, at 809 n.11.

⁹⁹ See Eaglin, *supra* note 94, at 68; Dawinder S. Sidhu, *Moneyball Sentencing*, 56 B.C. L. REV. 671, 671 (2015); Cade Metz & Adam Satariano, *An Algorithm That Grants Freedom, Or Takes It Away*, N.Y. TIMES (Feb. 7, 2010), <https://www.nytimes.com/2020/02/06/technology/predictive-algorithms-crime.html> [<https://perma.cc/H9BA-2WEY>].

¹⁰⁰ Stephen D. Hart, Kevin S. Douglas & Laura S. Guy, *The Structured Professional Judgment Approach to Violence Risk Assessment: Origins, Nature, and Advances*, in 2 THE WILEY HANDBOOK ON THE THEORIES, ASSESSMENT AND TREATMENT OF SEXUAL OFFENDING 643, 643–44, 647–53 (Douglas P. Boer et al. eds., 2017); Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, PROPUBLICA (May 23, 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [<https://perma.cc/45M4-QYK3>].

¹⁰¹ See Douglas Mossman, Allison H. Schwartz & Elise R. Elam, *Risky Business Versus Overt Acts: What Relevance Do “Actuarial,” Probabilistic Risk Assessments Have for Judicial Decisions on Involuntary Psychiatric Hospitalization?*, 11 HOUS. J. HEALTH L. & POL’Y 365, 392–93 (2011).

¹⁰² See Collins, *supra* note 77, at 63–64; BORUM ET AL., *supra* note 6, at 7–11.

¹⁰³ See Collins, *supra* note 77, at 63.

¹⁰⁴ See *id.* at 63–64; John Monahan & Jennifer Skeem, *Introductory Discussion of the Science—Research Approaches in Violence Risk Assessment*, in 2 DAVID L. FAIGMAN, EDWARD K. CHENG, JENNIFER L. MNOOKIN, ERIN E. MURPHY, JOSEPH SANDERS & CHRISTOPHER SLOBOGIN, MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 9:11 (2019).

puter program, rather than a human evaluator, analyzes risk factors and spits out the final risk prediction.¹⁰⁵

Several concerns have been raised in both the scientific and legal communities about the wisdom of relying on risk assessment tools to aid sentencing decisions.¹⁰⁶ These include that use of risk assessments at sentencing is contrary to their intended purpose, that the tools are embedded with problematic biases, and that risk assessments have limited, if any, predictive value.

Perhaps the biggest criticism of risk predictions is that risk assessment tools were never intended for use at sentencing.¹⁰⁷ Risk assessment tools are intended to assist authorities to design appropriate rehabilitation programming after a sentence is imposed,¹⁰⁸ and the developers of risk assessment tools have warned against using them at sentencings to determine punishment. The result of this improper use of risk assessments is that reliance on such tools at sentencing can result in wholly unwarranted punishments.¹⁰⁹

Factors considered in both types of risk assessment include criminal history, family criminal history, educational background, age, gender, employment status, substance abuse history, and a number of others.¹¹⁰ But in order for a risk assessment to be accurate, these factors, which are presumed to correlate with risk, must in fact be risk predictors. Scholars have expressed concern that certain factors considered in risk assessments—like gender and other personal characteristics—should not be treated as risk predictors or factored into punishment decisions at all.¹¹¹

Moreover, risk assessments rely on population data to make risk predictions.¹¹² This means that in order for a risk prediction to be accurate, risk assessments must be validated¹¹³ on populations that re-

¹⁰⁵ See Angwin et al., *supra* note 100; Metz & Satariano, *supra* note 99.

¹⁰⁶ See Cecelia Klingele, *The Promises and Perils of Evidence-Based Corrections*, 91 NOTRE DAME L. REV. 537, 576–77 (2015); Collins, *supra* note 77; Monahan & Skeem, *supra* note 93, at 501–08.

¹⁰⁷ See Collins, *supra* note 77, at 61.

¹⁰⁸ *Id.* at 61.

¹⁰⁹ *Id.* at 62–63.

¹¹⁰ See *id.* at 64; Klingele, *supra* note 106, at 577.

¹¹¹ See Collins, *supra* note 77, at 60–62.

¹¹² *Id.* at 60.

¹¹³ Validation refers to the process of demonstrating, through empirical testing, that a particular test or tool functions as intended and determining under what range of conditions it functions properly. See Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659, 688 (2018).

flect the populations to which the tool is applied.¹¹⁴ When risk assessments are used to assess individuals who are insufficiently similar to the populations they were validated on, their predictive value diminishes.¹¹⁵ For instance, the Static-99, a sex offender risk assessment tool discussed below, was developed and validated largely using incarcerated populations in Canada.¹¹⁶ But the rate of sex offending in Canada is more than double the rate in the United States,¹¹⁷ diminishing the predictive value of the Static-99 for evaluating Americans.¹¹⁸

Given the foundation in aggregate data, commenters have expressed concern that reliance on risk assessment tools will lead to sentencing determinations that are not individualized and that risk assessment instruments may not be accurate predictors of recidivism with respect to any given individual.¹¹⁹ Courts have started to warn against overdependence on risk assessments for this very reason.¹²⁰ Former Attorney General Eric Holder worried that reliance on risk sentencing would result in sentences based on “unchangeable factors that a person cannot control, or on the possibility of a future crime that has not taken place.”¹²¹ Among the biggest concerns over the reliance on risk assessment tools, particularly algorithmic tools, is that racial bias is embedded in the algorithms used to make risk predictions.¹²² For example, instruments give high weight to characteristics that correlate with race, like criminal history, zip code, and socioeconomic status.¹²³ Indeed, actuarial risk assessments have been found to

¹¹⁴ See Melissa Hamilton, *Public Safety, Individual Liberty, and Suspect Science: Future Dangerousness Assessments and Sex Offender Laws*, 83 TEMP. L. REV. 697, 730 (2011).

¹¹⁵ *Id.*

¹¹⁶ John A. Fennel, *Punishment by Another Name: The Inherent Overreaching in Sexually Dangerous Person Commitments*, 35 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 37, 53–54, 59 (2009).

¹¹⁷ *Id.* at 59.

¹¹⁸ *Id.* at 59–61; Hamilton, *supra* note 114, at 729–31.

¹¹⁹ See Collins, *supra* note 77, at 95–96; Klingele, *supra* note 106, at 577.

¹²⁰ See, e.g., *State v. Loomis*, 2016 WI 68, ¶¶ 62–64, 371 Wis. 2d 235, 881 N.W.2d 749; *Commonwealth v. Redacted*, No. 2015-590, 2016 WL 1460521, at *2 (Mass. Super. Ct. Apr. 12, 2016).

¹²¹ Eric H. Holder, Jr., U.S. Att’y Gen., Remarks at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference (Aug. 1, 2014), <https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-association-criminal-defense-lawyers-57th> [<https://perma.cc/9QNW-JGRL>].

¹²² See, e.g., Sandra G. Mayson, *Bias in, Bias Out*, 128 YALE L.J. 2218, 2222 (2019) (“[A]lgorithmic crime prediction . . . appears poised to entrench the inexcusable racial disparity so characteristic of our justice system, and to dignify the cultural trope of black criminality with the gloss of science.”).

¹²³ See Collins, *supra* note 77, at 102 (“Nevertheless, as scholars have noted, some of the seemingly less suspect actuarial risk factors correlate with race so strongly that they are essen-

“disproportionately classify minorities and the poor as higher risk, often due to factors outside their control.”¹²⁴ One study found that, after isolating for race, a widely used risk assessment tool, the Correctional Offender Management Profiling for Alternative Sanctions (“COMPAS”), was seventy-seven percent more likely to find African Americans to be at higher risk for committing future violence.¹²⁵

Many risk assessment instruments have very low predictive value.¹²⁶ For example, the Static-99 is among the most commonly used sex offender risk assessments.¹²⁷ Researchers have described it as “not much ‘better than a coin flip.’”¹²⁸ As a result, it has been suggested that a determination by the Static-99 that a person is a “high risk” of reoffending is unreliable.¹²⁹

Proper application of the tool is another key concern. Indeed, the heart of T.K.’s challenge to the admissibility of the SAVRY results at his sentencing was that the evaluator in his case misapplied the tool, resulting in an unreliable prediction of his potential for future violence.¹³⁰

Risk assessment tools like the COMPAS, SAVRY, and the Static-99 are being used more frequently and will likely continue to be relied

tially a proxy for this impermissible factor.”); Starr, *supra* note 93, at 819 (wondering “whether risk prediction instruments that classify defendants by demographic, socioeconomic, and family characteristics can be . . . justified”); Metz & Satariano, *supra* note 99 (discussing “worry that the biases—involving race, class and geography—of the people who create the algorithms are being baked into these systems”); Mayson, *supra* note 122 at 2222 (citing “profound concern about the racial impact of risk assessment”); cf. Will Douglas Heaven, *Predictive Policing Algorithms Are Racist. They Need to Be Dismantled*, MIT TECH. REV. (July 17, 2020), <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/> [<https://perma.cc/MTZ7-6DCP>] (“Though by law the algorithms do not use race as a predictor, other variables, such as socioeconomic background, education, and zip code, act as proxies.”).

¹²⁴ Klingele, *supra* note 106, at 577; see generally Mayson, *supra* note 122, at 2218 (“Algorithmic risk assessment has revealed the inequality inherent in all prediction . . .”).

¹²⁵ See Angwin et al., *supra* note 100. The study found that the same tool was forty-five percent more likely to find African Americans “predicted to commit a future crime of any kind.” *Id.*

¹²⁶ See Collins, *supra* note 77, at 95–96.

¹²⁷ See R. Karl Hanson, Leslie Helmus & David Thornton, *Predicting Recidivism Amongst Sexual Offenders: A Multi-Site Study of Static-2002*, 34 LAW & HUM. BEHAV. 198, 198 (2010); Heather Ellis Cucolo & Michael L. Perlin, “Far from the Turbulent Space”: Considering the Adequacy of Counsel in the Representation of Individuals Accused of Being Sexually Violent Predators, 18 U. PA. J.L. & SOC. CHANGE 125, 140–41 (2015).

¹²⁸ See Hamilton, *supra* note 114, at 728 (quoting Fred S. Berlin, Nathan W. Galbreath, Brendan Geary & Gerard McGlone, *The Use of Actuarials at Civil Commitment Hearings to Predict the Likelihood of Future Sexual Violence*, 15 SEXUAL ABUSE 377, 381 (2003)).

¹²⁹ Hamilton, *supra* note 114, at 727–28.

¹³⁰ See *supra* notes 9–13 and accompanying text.

on by sentencing courts.¹³¹ Indeed, despite all these concerns and red flags, the Model Penal Code instructs states to develop risk assessment instruments to aid in sentencing decision making.¹³²

TRADITIONAL MENTAL HEALTH AND MENTAL STATUS EVIDENCE. Psychological or psychiatric information about a defendant's mental health or mental illness, cognitive functioning, or other mental status data points has become ubiquitous in criminal sentencings, capital and noncapital.¹³³ In capital cases, psychological or psychiatric evidence is regularly offered at the penalty phase, both as aggravating evidence by prosecutors and as mitigating evidence by defendants. The same is true in noncapital cases, and courts themselves often order psychological evaluations to gather information about a defendant prior to sentencing, like the judge did in T.K.'s case.¹³⁴ Some jurisdictions go so far as to *require* psychological reports at sentencing under certain circumstances.¹³⁵ Likewise, the parties often supply their own psychological or psychiatric reports for consideration by the court in the form of affidavits, letters, or testimony.¹³⁶

This type of information is used in myriad ways at sentencing. It can shed light on a defendant's mental illnesses, his learning or educational disabilities, the impact of drug use on the brain, or many other aspects of a defendant's mental condition.¹³⁷ It can also be used to inform treatment decisions and rehabilitation programming, assist in assessing a defendant's likelihood of recidivism, and help understand

¹³¹ See Collins, *supra* note 77, at 57, 60.

¹³² See MODEL PENAL CODE § 6B.09 (AM. L. INST., Proposed Final Draft 2017); Monahan & Skeem, *supra* note 93, at 494–95 (describing the Model Penal Code mandate); Richard E. Redding, *Evidence-Based Sentencing: The Science of Sentencing Policy and Practice*, 1 CHAP. J. CRIM. JUST. 1, 7–8 (2009) (same).

¹³³ See Shari N. Spitz, *Psychiatric and Psychological Examinations for Sentencing: An Analysis of Caselaw From the Second Circuit in Comparison with Other Federal Circuits and the Governing Federal Statutes*, 6 QUINNIPIAC HEALTH L.J. 133, 135 (2003) (analyzing “the various federal statutes which enable a criminal defendant to obtain a psychiatric or psychological examination”).

¹³⁴ See, e.g., ARIZ. R. CRIM. P. 26.5 (allowing court to order mental health evaluation of defendant before sentencing); KY. REV. STAT. ANN. § 504.140 (West 2020) (allowing court to appoint a psychologist or psychiatrist to report on defendant's mental condition at the time of sentencing where the defendant is found guilty but mentally ill).

¹³⁵ See, e.g., IDAHO CODE ANN. § 19-2522 (2020) (requiring, where mental condition will bear significantly on the sentencing decision, a psychiatric or psychological examination of a defendant prior to sentencing).

¹³⁶ See, e.g., Thomas, *supra* note 74, at 200 (“The lawyer could develop effective sentencing advocacy around [the] fact [of a client's mental illness], including the testimony of relatives or doctors, records of treatment or evaluations, or letters from counselors.”).

¹³⁷ See *id.*

how mental health conditions influence criminal conduct.¹³⁸ Evidence of cognitive abilities is particularly crucial in capital sentencings; evidence of intellectual disability is a key aspect of capital mitigation because a death sentence has been found to be a violation of the Eighth Amendment's Cruel and Unusual Punishment Clause when imposed against the intellectually disabled.¹³⁹

Despite their prevalence, not all types of psychological and psychiatric information are equally reliable or accurate. For example, prosecutors frequently call psychiatrists or psychologists to testify on the question of a defendant's future dangerousness, which is treated as an aggravating factor in most capital jurisdictions.¹⁴⁰ Whether predictions of future dangerousness are reliable, accurate, or have any basis in science, however, is, at best, hotly debated.¹⁴¹ Fifth Circuit judge Emilio Garza commented that "[t]he scientific community virtually unanimously agrees that psychiatric testimony on future dangerousness is, to put it bluntly, unreliable and unscientific."¹⁴² Even so, the Supreme Court has allowed the admission of psychiatric testimony on future dangerousness in capital sentencings.¹⁴³

Some well-known cases offer concrete evidence that sentencing decisions have in fact been based on unreliable mental health evidence and illustrate how that type of information—if unchecked—can lead to unjust sentencing decisions. In 1997, Duane Buck was convicted of capital murder in Texas for killing his ex-girlfriend and her friend.¹⁴⁴ At the penalty phase, Buck's own lawyer called a psychologist who testified that Buck was statistically more likely to be violent in the future because he is Black.¹⁴⁵ Despite the blatant falsity of this claim,¹⁴⁶ the jury was allowed to consider that testimony and imposed

¹³⁸ See *id.*

¹³⁹ See *Hall v. Florida*, 572 U.S. 701, 708 (2014).

¹⁴⁰ See, e.g., TEX. CRIM. PROC. CODE ANN. § 37.071.2(a)(2)(b)(1) (West 2019) (requiring a finding be made by a capital jury "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society"); see also Denno, *supra* note 83, at 526; Gertner, *supra* note 86, at 543 ("Neuroscience testimony could well be offered by prosecutors to show aggravating factors and by defense lawyers to show mitigating factors. While [a recent series of] studies found no instance of the prosecution using neurogenetic evidence as an aggravating factor in capital cases, the past may not predict the future, particularly in noncapital cases where the rules are more relaxed.").

¹⁴¹ See Abram S. Barth, Comment, *A Double-Edged Sword: The Role of Neuroimaging in Federal Capital Sentencing*, 33 AM. J.L. & MED. 501, 507–08 (2007).

¹⁴² *Flores v. Johnson*, 210 F.3d 456, 463 (5th Cir. 2000) (Garza, J., specially concurring).

¹⁴³ See, e.g., *Barefoot v. Estelle*, 463 U.S. 880, 898–99 (1983).

¹⁴⁴ See *Buck v. Davis*, 137 S. Ct. 759, 767 (2017); Haag, *supra*, note 31.

¹⁴⁵ See *Buck*, 137 S. Ct. at 767–69.

¹⁴⁶ See, e.g., Grant H. Morris, *Defining Dangerousness: Risking a Dangerous Definition*, 10

the death penalty.¹⁴⁷ Buck was ultimately saved from execution when the Supreme Court found his counsel ineffective for admitting such damaging testimony against his own client.¹⁴⁸

The vast majority of criminal defendants are not so fortunate. The introduction of unreliable scientific evidence at Buck's sentencing was, in quite literal terms, the difference between life and death. The reversal of his death sentence did not lead to improvements to the types of scientific evidence that can be considered at sentencing. Buck was successful due to the introduction of "especially pernicious" racially discriminatory evidence presented by his own lawyer at his capital sentencing—a set of facts so egregious that they are unlikely to ever be repeated.¹⁴⁹ For those whose cases do not turn on such obviously discriminatory and unlawful testimony presented by their own lawyers, there is no meaningful system in place to screen out unreliable STS evidence and prevent it from being relied upon to enhance punishment.

NEUROSCIENCE EVIDENCE. Neuroscience, the study of the brain and nervous system, is also being used with increasing frequency in sentencings.¹⁵⁰ Typically, neuroscience evidence takes the form of neuroimaging data, computer-generated images of the brain, or nonimaging evidence resulting from tests administered by a medical

J. CONTEMP. LEGAL ISSUES 61, 90 n.175 (1999) (describing how research has shown that, after controlling for socioeconomic status, race is not a risk factor for violence); Jeffrey S. Swanson, Charles E. Holzer III, Vijay K. Ganju & Robert Tsutomu Jono, *Violence and Psychiatric Disorder in the Community: Evidence from the Epidemiologic Catchment Area Surveys*, 41 HOSP. & CMTY. PSYCHIATRY 761, 764–68 (1990) ("Race was not related to violence when socioeconomic status was controlled."); JOHN MONAHAN, HENRY J. STEADMAN, ERIC SILVER, PAUL S. APPELBAUM, PAMELA CLARK ROBBINS, EDWARD P. MULVEY, LOREN H. ROTH, THOMAS GRISSO & STEVEN BANKS, *RETHINKING RISK ASSESSMENT: THE MACARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE* 58 (2001) ("African Americans and whites residing in comparably disadvantaged neighborhoods showed no differences in their rates of violence."). The Supreme Court described the so-called expert testimony as reflecting "a particularly noxious strain of racial prejudice." *Buck*, 137 S. Ct. at 776.

¹⁴⁷ *Buck*, 137 S. Ct. at 775–76 (explaining that defense counsel's decision to call the psychologist despite being aware of his view that being Black increased the likelihood that Buck would be a future danger "fell outside the bounds of competent representation.").

¹⁴⁸ *Id.* at 779–80. After reversal, Mr. Buck reached a deal in which his sentence was reduced to life in prison. Alex Arriaga, *Texas Death Row Inmate Duane Buck Has Sentence Reduced to Life After Supreme Court Orders Retrial*, TEX. TRIB. (Oct. 3, 2017, 5:00 PM), <https://www.texastribune.org/2017/10/03/high-profile-death-row-case-comes-end-guilty-plea/> [<https://perma.cc/SM4F-C9RZ>].

¹⁴⁹ *Buck*, 137 S. Ct. at 778 (quoting *Rose v. Mitchell*, 443 U.S. 545, 555 (1979)).

¹⁵⁰ See Denno, *supra* note 83, at 495; Bernice B. Donald & Erica Bakies, *A Glimpse Inside the Brain's Black Box: Understanding the Role of Neuroscience in Criminal Sentencing*, 85 FORDHAM L. REV. 481, 483 (2016).

professional¹⁵¹ in order to attempt to measure an individual's brain function as compared to normal brain function.¹⁵² Both imaging and nonimaging data seek to shed light on brain operation and functioning. But neuroimaging allows visualization of injuries and other abnormalities in a brain; it allows us to *see* brain damage.

Neuroimaging can support a defendant's arguments to mitigate his sentence by highlighting physical causes for, or contributors to, the offending conduct in question¹⁵³ and is, therefore, increasingly being offered as mitigating evidence in sentencings, especially in capital cases.¹⁵⁴ While not yet regularly used by prosecutors against defendants,¹⁵⁵ there is concern that prosecutors, especially in the capital context, will introduce neuroscience evidence to support the argument that defendants are likely to commit future crimes and, thus, more severe penalties or death sentences are appropriate.¹⁵⁶ For example, brain scans revealing abnormalities in the amygdala, which regulates moral socialization,¹⁵⁷ or injuries to the frontal lobe, which "regulate[s] socially appropriate behavior" and impulsivity,¹⁵⁸ could be used to fuel arguments that a defendant is likely to be dangerous in the future.¹⁵⁹ Though this fear has not yet been borne out, commentators are concerned that this will change in the future, especially in noncapital sentencings.¹⁶⁰

Using neuroscience-based evidence to attempt to predict future behavior or to evaluate past conduct to determine punishment poses significant concerns. Brain scans can be manipulated by certain movements or even by the subject thinking about specific things during a scan.¹⁶¹ There is also an expectation that every brain structure controls a particular aspect of behavior and that, as a result, an injury to a

151 See Donald & Bakies, *supra* note 150, at 484.

152 See *id.* at 484–85.

153 See Elizabeth Bennett, *Neuroscience and Criminal Law: Have We Been Getting It Wrong for Centuries and Where Do We Go from Here?*, 85 *FORDHAM L. REV.* 437, 449 (2016).

154 See Lyn M. Gaudet & Gary E. Marchant, *Under the Radar: Neuroimaging Evidence in the Criminal Courtroom*, 64 *DRAKE L. REV.* 577, 627 (2016); Denno, *supra* note 83, at 499.

155 See Denno, *supra* note 83, at 496–97.

156 Donald & Bakies, *supra* note 150, at 495; see Gertner, *supra* note 86, at 543–44.

157 Bennett, *supra* note 153, at 449.

158 Barth, *supra* note 141, at 503.

159 Denno, *supra* note 83, at 496–97.

160 See, e.g., Gertner, *supra* note 86, at 543.

161 See Henry T. Greely & Judy Illes, *Neuroscience-Based Lie Detection: The Urgent Need for Regulation*, 33 *AM. J.L. & MED.* 377, 404–05 (2007) ("Simple movements of the tongue or jaw will make fMRI scans unreadable. . . . [S]imply thinking about other things during a task may activate other brain regions in ways that interfere").

particular structure would impact a specific type of behavior.¹⁶² But the degree to which a behavior is regulated by a specific structure varies significantly, both in general and from person to person.¹⁶³ Different individuals can exhibit similar behaviors despite a brain scan revealing significantly different patterns.¹⁶⁴ Additionally, interpreting imaging results is complex and data can be misinterpreted.¹⁶⁵ All of this is to say that, like other forms of STS, faulty, misinterpreted, or otherwise unreliable neuroscience evidence can lead to unjust sentencing decisions if not properly screened for reliability.

b. Traditional Forensic Identification Methods

Traditional forensic identification methods such as DNA analysis, fingerprint comparisons, and firearms and toolmarks comparisons represent a wholly different category of scientific or specialized evidence used at sentencings. Unlike the forms of STS evidence described previously, forensic identification techniques do not offer information directly related to a defendant's background, education, risk, or personal characteristics. Rather, they seek to link a defendant to evidence of other crimes to make the case that a defendant deserves a higher sentence because of uncharged or even acquitted conduct. For example, during a sentencing hearing, a prosecutor may highlight DNA results—perhaps from a “hit” in a DNA database¹⁶⁶—that suggest that the defendant has committed additional uncharged crimes beyond those for which she was convicted. The prosecutor could use this to argue that the defendant is especially dangerous and, therefore, merits a more severe sentence, notwithstanding the fact

¹⁶² See Adina L. Roskies & Kimberly Farbota, *Evaluating Neuroimaging Results*, in 2 DAVID L. FAIGMAN, EDWARD K. CHENG, JENNIFER L. MNOKIN, ERIN E. MURPHY, JOSEPH SANDERS & CHRISTOPHER SLOBOGIN, *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 20:49 (2019).

¹⁶³ *Id.*

¹⁶⁴ See Greely & Iles, *supra* note 161, at 382 (“Inter-subject variability is also a consideration . . . [T]wo independent subjects [may] show different patterns of activation while their behavioral performances are comparable.”).

¹⁶⁵ See A.C. Pustilnik, *Imaging Brains, Changing Minds: How Pain Neuroimaging Can Inform the Law*, 66 ALA. L. REV. 1099, 1156 (2015).

¹⁶⁶ See Kimberly A. Polanco, Note, *Constitutional Law—The Fourth Amendment Challenge to DNA Sampling of Arrestees Pursuant to the Justice for All Act of 2004: A Proposed Modification to the Traditional Fourth Amendment Test of Reasonableness*, 27 U. ARK. LITTLE ROCK L. REV. 483, 489–90 (2005) (describing how DNA databases are maintained); *Frequently Asked Questions on CODIS and NDIS*, FED. BUREAU OF INVESTIGATION, <https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet> [<https://perma.cc/PM6X-5W9N>]; Kimberly A. Wah, *A New Investigative Lead: Familial Searching as an Effective Crime-Fighting Tool*, 29 WHITTIER L. REV. 909, 916 (2008).

that she was not charged with or convicted of the other crimes.¹⁶⁷ But as described previously, even the results of DNA analysis cannot be assumed to be reliable.¹⁶⁸

Prosecutors might similarly offer evidence of uncharged misconduct based on fingerprint or ballistics matches to argue that the defendant is guilty of additional crimes beyond those for which he has been convicted. For example, in a burglary case, a prosecutor may offer evidence that the defendant's fingerprints were found at additional crime scenes to support a claim that the defendant committed additional burglaries.¹⁶⁹ The same is true with respect to ballistics evidence—prosecutors may argue for higher sentences by linking ballistics patterns from a gun associated with the accused to ballistics recovered in uncharged cases.¹⁷⁰

Fingerprint and ballistics comparisons are two examples of what are commonly described as “pattern matching” evidence, in which markings, or patterns, on an item with an unknown source, such as a latent fingerprint lifted from a crime scene, are compared to patterns on a known source, such as a fingerprint deposited by an identified suspect.¹⁷¹ The reliability of pattern matching techniques is frequently questioned, largely because pattern analysis relies on the subjective

¹⁶⁷ See, e.g., *Alix v. Quarterman*, 309 F. App'x 875, 877 (5th Cir. 2009) (describing prosecution's introduction, in sentencing phase of capital murder trial, of faulty DNA evidence linking defendant to unrelated homicide); *People v. Pitts*, 2013 IL App (3d) 110676-U, ¶ 6 (describing persuasiveness of State's presentation of defendant's uncharged sexual assaults based on DNA linkages); *State v. Gatewood*, 2001 WI App 280, ¶¶ 6–9, 248 Wis. 2d 982, 638 N.W.2d 394 (finding due process violation when trial court relied on DNA results that were not provided to defendant but which prosecutor argued at sentencing linked him to two uncharged assaults).

¹⁶⁸ See *supra* notes 40–42 and accompanying text.

¹⁶⁹ See B.H. Glenn, Annotation, *Court's Right, in Imposing Sentence, to Hear Evidence of, Or to Consider, Other Offenses Committed by Defendant*, 96 A.L.R.2d 768, § 10(f) (1964) (“Rejecting the objection of lack of proper authentication, the court in *Murphy v State* (1944) 184 Md 70, 40 A2d 239, . . . approved the sentencing judge's use of an FBI record to show the defendant's previous criminal record, stating that while that record was not certified in any manner, it showed on its face that it was based on fingerprints, of which the court took judicial notice as an infallible means of identification, and pointing out that it was not claimed that the record was false or inaccurate.”); Eang Ngov, *Judicial Nullification of Juries: Use of Acquitted Conduct at Sentencing*, 76 TENN. L. REV. 235, 235 (2009) (“In deciding a defendant's sentence, a court may consider conduct that has not been proven beyond a reasonable doubt and even conduct of which the jury has acquitted the defendant. Consideration of acquitted conduct has resulted in dramatic increases in the length of defendants' sentences . . .”).

¹⁷⁰ See, e.g., *United States v. Wilson*, 624 F.3d 640, 651 n.10 (4th Cir. 2010).

¹⁷¹ See Itiel 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 (2010).

interpretations of an examiner.¹⁷² In 2016, the President's Council of Advisers on Science and Technology, in a landmark report assessing the validity of several traditional forensic identification methods, found bitemark analysis and firearms and toolmark analysis to lack foundational validity.¹⁷³ The same report noted several issues with respect to the reliability of fingerprint analysis including a substantially high false positive rate.¹⁷⁴

Courts have started to take notice of the flaws of such disciplines and, in some cases, have begun limiting their admissibility at trial.¹⁷⁵ But admissibility rules do not apply at sentencing, which allows prosecutors to seek higher sentences based on purported linkages to other crimes using pattern-matching techniques for which no reliability assessments have been conducted.

C. Comparing Trial-Stage and Sentencing-Stage Standards for Admission of Scientific, Technical, or Specialized Evidence

Rules of evidence and other procedural rules promote fair trials and just outcomes by ensuring the reliability and trustworthiness of evidence admitted at the trial stage of a criminal prosecution.¹⁷⁶ That function is especially important in the context of STS information because STS evidence, by its very nature, contains conclusions that laypeople—including judges—cannot reach on their own.¹⁷⁷ Jurors and judges view information presented with the label of “scientific,” “technical,” or “expert” as having a stamp of authority and reliability.¹⁷⁸ Courts often warn that this “aura of special reliability and trustworthiness” creates the potential for confusing the factfinder and prejudicing the defendant.¹⁷⁹ As a result, faulty STS evidence has

¹⁷² See Brandon L. Garrett, *The Reliable Application of Fingerprint Evidence*, 66 UCLA L. REV. 64, 66 (2018).

¹⁷³ See PCAST REPORT, *supra* note 44, at 87 (assessing bitemark analysis); *id.* at 104–12 (assessing toolmark and firearm analysis).

¹⁷⁴ *Id.* at 95–102.

¹⁷⁵ See, e.g., *Williams v. United States*, 210 A.3d 734, 736 (D.C. 2019); *State v. McPhaul*, 808 S.E.2d 294, 303 (N.C. Ct. App. 2017), *disc. review allowed* by 812 S.E.2d 847 (N.C. 2018), *disc. review improvidently allowed* by 818 S.E.2d 102 (N.C. 2018).

¹⁷⁶ Kelly, *supra* note 83, at 416.

¹⁷⁷ Garrett & Fabricant, *supra* note 45, at 1581 (experts can form opinions and reach conclusions that lay witnesses cannot).

¹⁷⁸ See, e.g., *Flores v. Johnson*, 210 F.3d 456, 465–66 (5th Cir. 2000) (Garza J., specially concurring); *Motorola Inc. v. Murray*, 147 A.3d 751, 753, 755 (D.C. 2016) (en banc).

¹⁷⁹ See, e.g., *United States v. Fosher*, 590 F.2d 381, 383 (1st Cir. 1979) (finding that testimony related to eyewitness identification “would create a substantial danger of undue prejudice and confusion because of its aura of special reliability and trustworthiness”); *United States v.*

great potential to lead to unfair outcomes: unlike other evidence which jurors and judges are well equipped to weigh and consider, evidence that purports to be based in scientific or technical fields brings with it a veneer of authority that creates a significant possibility of causing undue prejudice. Because of all this, STS evidence is treated differently than other evidence at the trial stage.

Rule 702 and state equivalents operate to screen out unreliable STS evidence to avoid unwarranted reliance on junk science merely due to its “aura” of scientific authority. But as described previously, STS evidence is frequently used at sentencing as much, if not more, than at trial.¹⁸⁰ Still, no similar mechanism to screen STS evidence exists at sentencing.

This Section explains the critical role Rule 702 plays in screening out unreliable STS evidence at the trial stage of a criminal prosecution. It first briefly describes how the Rule came to include a substantive reliability test that functions to keep unreliable evidence out of trials. In doing so, it reveals parallels between what motivated the development of the modern version of Rule 702 and what motivates the need for extension of Rule 702 to sentencing: the increased use of STS evidence at each stage. It then compares Rule 702 to the minimal standard applied at the sentencing phase to determine what evidence may be considered in punishment decisions.

1. *The Trial Phase: The Rule 702 Standard*

The danger of unwarranted reliance on STS evidence simply because it is labeled “scientific” exists for judges as well as juries, all of whom, as nonexperts, are often ill equipped to grapple with scientific evidence and tend to over-rely on information that is marked as having the stamp of scientific authority when acting as factfinders. In its landmark 2009 report, the National Research Council cautioned that “judges and lawyers . . . generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner.”¹⁸¹ Legal and scientific commentators have similarly warned

Kozminski, 821 F.2d 1186, 1199 (6th Cir. 1987) (warning that “‘the aura of special reliability and trustworthiness’ which attaches to expert testimony admitted without proper foundation . . . may confuse or mislead the trier of fact and thus defeat a defendant’s right to a fair trial” (quoting *United States v. Brown*, 557 F.2d 541, 556 (6th Cir. 1977))), *aff’d and remanded*, 487 U.S. 931 (1988); *United States v. Amaral*, 488 F.2d 1148, 1152 (9th Cir. 1973) (“Scientific or expert testimony particularly courts the second danger [of undue prejudice or of confusing the issues or of misleading the jury] because of its aura of special reliability and trustworthiness.”).

¹⁸⁰ See *supra* notes 92–98 and accompanying text.

¹⁸¹ NAS REPORT, *supra* note 45, at 110.

that judges do not have the training or skill to adequately assess scientific information¹⁸² and are also likely to be taken in by the “aura” of information labeled as scientific.¹⁸³

At the trial stage, federal jurisdictions and the many states that have adopted the Federal Rules of Evidence attempt to avoid the problems created by misapplication of expert evidence by applying the admissibility test codified in Rule 702.¹⁸⁴ The rule plays an essential gatekeeping function at this stage, limiting admissible expert evidence to only that which meets the Rule’s relevance and reliability requirements. The modern iteration of Rule 702 provides that:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

(a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;

(b) the testimony is based on sufficient facts or data;

(c) the testimony is the product of reliable principles and methods; and

¹⁸² See Kayla Marie Mannucci, Note, *Framed by Forensics: Fulfilling Daubert’s Gatekeeping Function by Segregating Science from the Adversarial Model*, 39 CARDOZO L. REV. 1947, 1955 (2018); see also Epstein, *supra* note 44, at 97–98 (discussing a “knowledge deficit” among lawyers and judges regarding scientific methods); Barbara A. Spellman in Radley Balko, Opinion, *How Much Should Juries Rely on Expert Testimony?*, WASH. POST (July 16, 2019, 8:00 AM), <https://www.washingtonpost.com/opinions/2019/07/16/how-much-should-juries-rely-expert-testimony/> [<https://perma.cc/5SR8-USBP>] (noting that judges along with jurors highly value forensic evidence).

¹⁸³ See Boaz Sangero, *Safety from Flawed Forensic Sciences Evidence*, 34 GA. ST. U. L. REV. 1129, 1135 (2018) (“[T]he ‘aura’ of science can be expected to blind judges and jurors and lead them to overestimate the real probative strength of scientific evidence.”).

¹⁸⁴ Jurisdictions which have adopted Rule 702 in its current iteration or a substantially similar rule are: Alabama, Arizona, Delaware, the District of Columbia, Florida, Georgia, Indiana, Kansas, Kentucky, Louisiana, Massachusetts, Michigan, Mississippi, Missouri, New Hampshire, North Carolina, Ohio, Oklahoma, South Dakota, Utah, Vermont, West Virginia, Wisconsin and Wyoming. See ALA. R. EVID. 702; ARIZ. R. EVID. 702; DEL. R. EVID. 702; Motorola Inc. v. Murray, 147 A.3d 751 (D.C. 2016); FLA. STAT. ANN. § 90.702 (2019); GA. CODE ANN. § 24-7-702 (2020); IND. R. EVID. 702; KAN. STAT. ANN. § 60-456 (2019); KY. R. EVID. 702; LA. CODE EVID. ANN. art. 702 (2019); MASS. R. EVID. § 702; MICH. R. EVID. 702; MISS. R. EVID. 702; MO. REV. STAT. § 490.065 (2020); N.H. R. EVID. 702; N.C. GEN. STAT. § 8C-1, R. 702 (2019); OHIO R. EVID. 702; OKLA. STAT. tit. 12, § 2702 (2013); S.D. CODIFIED LAWS § 19-19-702 (2016); UTAH R. EVID. 702; VT. R. EVID. 702; W. VA. R. EVID. 702; WIS. STAT. § 907.02 (2020); WYO. R. EVID. 702; see also Garrett & Fabricant, *supra* note 45, at 1572 n.74, app. II (listing the states that have adopted and use Rule 702 in criminal cases).

(d) the expert has reliably applied the principles and methods to the facts of the case.¹⁸⁵

But the Rule did not always have such clear requirements. For many years, two competing standards were used to assess the admissibility of expert evidence in trials. Until Rule 702 was first adopted as part of the Federal Rules of Evidence in 1975,¹⁸⁶ the prevailing standard courts applied was the “general acceptance” test proposed over five decades earlier in *Frye v. United States*.¹⁸⁷ Under that test, courts admitted scientific testimony if the technique had been generally accepted by the relevant scientific community.¹⁸⁸

The *Frye* standard is lax. Its approach is a “hands off” one that allows judges to avoid meaningfully engaging with whether a given scientific discipline is valid or not. Under the *Frye* test, that question is left to the relevant scientific community; trial courts must decide merely whether the evidence has gained general acceptance in that community.¹⁸⁹ *Frye* is not concerned with actual methodological reliability.¹⁹⁰

The adoption of the Federal Rules of Evidence complicated the question of how to determine the admissibility of expert evidence at trials. The original version of Rule 702 was far more lenient than today’s version, containing only the substance of subsection (a) of the modern version of the rule: “If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise”¹⁹¹

It contained only two requirements, neither of which was designed to address the scientific validity of expert evidence. These were, first, that the evidence must be relevant and, second, that the

¹⁸⁵ FED. R. EVID. 702. This version contains stylistic amendments added in 2011. *See* FED. R. EVID. 702 advisory committee’s note to 2011 amendment.

¹⁸⁶ *See* FED. R. EVID. 702 advisory committee’s note to 2011 amendment.

¹⁸⁷ 293 F. 1013, 1014 (D.C. Cir. 1923).

¹⁸⁸ *See id.*

¹⁸⁹ *See* Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half-Century Later*, 80 COLUM. L. REV. 1197, 1205 (1980).

¹⁹⁰ *See, e.g.,* State v. Coon, 974 P.2d 386, 393–94 (Alaska 1999) (“*Frye* . . . admits scientifically unreliable evidence which although generally accepted, cannot meet rigorous scientific scrutiny.”), *abrogated on other grounds by* State v. Sharpe, 435 P.3d 887, 899–900 (Alaska 2019).

¹⁹¹ 5 JACK B. WEINSTEIN & MARGARET A. BERGER, WEINSTEIN’S FEDERAL EVIDENCE, § 702 App.01, App.02 (Mark S. Brodin & Joseph M. McLaughlin eds., Matthew Bender 2d ed. 2020).

expert providing such evidence have basic qualifications.¹⁹² As a result, under the original Rule 702, admission of expert evidence at trials was virtually unlimited. The problem with such a lenient standard quickly became evident as courts frequently admitted allegedly scientific or specialized evidence even if it was not based on objectively sound methodology.¹⁹³

Until the mid-to-late 1970s, expert evidence was thus liberally admitted in both civil and criminal trials.¹⁹⁴ But two developments made the need for better gatekeeping clear. The first was a surge in toxic tort litigation, in which plaintiffs claimed that exposure to a variety of substances caused various ailments, including cancer.¹⁹⁵ Expert evidence in these cases was frequently based on less-than-scientific claims.¹⁹⁶ The second development was the increased use of expert evidence in criminal cases.¹⁹⁷ Evidence based on a number of newly developing disciplines like voice identification and bite mark comparison, among others, saw increasing use in criminal trials.¹⁹⁸ As a result of these two developments, courts were flooded with claims based on unscientific evidence and began to seek more stringent admissibility standards than those provided by the liberal Rule 702,¹⁹⁹ with some even returning to the *Frye* test's "general acceptance" standard.²⁰⁰

As some judges reverted to using the *Frye* standard to respond to the problems caused by the rapid acceleration of toxic tort litigation on the civil side and the increased use of novel forensic identification techniques on the criminal side, the question of whether *Frye* or Rule 702 should govern expert admissibility quickly came to a head. It was finally resolved in the landmark 1993 case, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*²⁰¹

In *Daubert*, the Supreme Court found that Rule 702 had superseded *Frye* and held that trial judges must assess the reliability of purportedly scientific evidence before admitting it.²⁰² The Court noted

¹⁹² See *id.*

¹⁹³ See David E. Bernstein, *The Misbegotten Judicial Resistance to the Daubert Revolution*, 89 NOTRE DAME L. REV. 27, 31–32 (2013).

¹⁹⁴ See *id.* at 34–35; David E. Bernstein & Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 57 WM. & MARY L. REV. 1, 4 (2015).

¹⁹⁵ Bernstein, *supra* note 193, at 34–35.

¹⁹⁶ See Bernstein & Lasker, *supra* note 194, at 4.

¹⁹⁷ See Bernstein, *supra* note 193, at 34–35.

¹⁹⁸ See Giannelli, *supra* note 189, at 1198.

¹⁹⁹ See Bernstein & Lasker, *supra* note 194, at 4.

²⁰⁰ See *id.*; Bernstein, *supra* note 193, at 34–35, 40.

²⁰¹ 509 U.S. 579 (1993).

²⁰² *Id.* at 589–90.

that, although helpfulness to the factfinder and general acceptance were pertinent to an admissibility inquiry, the focus of the assessment must be on the validity of the principles and methodology of any given technique.²⁰³ In *Daubert*, the Court provided five substantive factors that judges can use to evaluate scientific reliability.²⁰⁴ In a later case, the Court emphasized that the *Daubert* standard applies not just to scientific evidence, but to all expert evidence, including scientific, technical, and specialized evidence.²⁰⁵ In 2000, Rule 702 was amended to codify the *Daubert* standard.²⁰⁶ The current version of Rule 702 reflects that codification.²⁰⁷ Ultimately, *Daubert* and the subsequent amendment of Rule 702 created a meaningful mechanism for courts to reject the flawed scientific evidence that had become an increasingly common feature of trials.

The Supreme Court later described the *Daubert* standard as containing “exacting” reliability requirements for expert evidence.²⁰⁸ Subsections (c) and (d) of Rule 702, which codify *Daubert*’s reliability test, aim to ensure that STS evidence presented at trial is scientifically valid.²⁰⁹

Rule 702(c) requires the principles and methodology underlying STS evidence—whether DNA analysis, a psychological test, or anything else—to be scientifically sound.²¹⁰ Scientists have described this requirement as “foundational validity.”²¹¹ For a technique to be considered foundationally valid, its conclusions must be repeatable, reproducible, and accurate.²¹² In *Daubert*, the Supreme Court outlined a list of nonexhaustive factors that may be used to evaluate whether a methodology is reliable: (1) whether it can be or has been tested; (2) whether the theory or practice has been published in scientific, peer-reviewed journals; (3) whether the technique has a known or po-

²⁰³ See *id.* at 594–95.

²⁰⁴ See *id.* at 592–95.

²⁰⁵ See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147 (1999).

²⁰⁶ FED. R. EVID. 702 advisory committee’s note to 2000 Amendment.

²⁰⁷ *Id.*

²⁰⁸ See *Weisgram v. Marley Co.*, 528 U.S. 440, 455 (2000).

²⁰⁹ See *Daubert*, 509 U.S. at 591 n.9 (“In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*.”).

²¹⁰ See FED. R. EVID. 702(c).

²¹¹ See PCAST REPORT, *supra* note 44, at 43.

²¹² *Id.* at 5, 46–47. Repeatability refers to the idea that the same examiner will obtain the same result upon multiple applications of a method under the same circumstances. See *id.* at 47. Reproducibility refers to the idea that different examiners will get the same results upon applications of a method under the same circumstances. See *id.* Accuracy refers to obtaining the correct result using a particular method. See *id.*

tential rate of error, and what that rate of error is; (4) whether standards exist to control the technique's operation; and (5) the technique's degree of acceptance within the scientific community.²¹³

Subsection (d) focuses on a different aspect of reliability: the reliability of the application of a method to a given set of facts. In other words, assuming the reliability of the principles and methods of a given discipline, subsection (d) asks if the method was reliably *applied* to a particular case.²¹⁴ Scientists have referred to this question as whether a method has "validity as applied."²¹⁵ For STS evidence to be admitted under Rule 702, the proponent must establish by a preponderance of the evidence that all admissibility requirements, including those of subsections (c) and (d), are met.²¹⁶

Consider an example in which a lab has validated its DNA testing procedures only for samples containing mixtures of up to three contributors of DNA but a prosecutor seeks to present evidence that the defendant's DNA is included in a mixture of *five* contributors—two more than the procedure the lab has validated for. Although DNA analysis is a generally reliable technique that might satisfy the requirements of Rule 702(c), the method may not have been reliably applied to this particular case because it was used on a five-person mixture, more than the lab's validation allows for.

Parties may attempt to convince a judge of the reliability of proffered expert evidence under Rule 702(c) and (d) in a variety of ways.²¹⁷ For example, they might offer literature supporting the validity of a method, hire an expert to discuss the method and its use in the instant case, or provide the court with an affidavit or declaration containing the expert's conclusions.²¹⁸ In cases involving complex issues or novel evidence, the trial court may hold an evidentiary hearing with experts testifying about a method or its application in order to determine admissibility.²¹⁹ Rule 702 limits the admission of expert evidence that is unreliable by requiring the party proffering such evidence to demonstrate it is objectively verifiable.²²⁰

²¹³ 509 U.S. 579, 592–95.

²¹⁴ FED. R. EVID. 702(d).

²¹⁵ PCAST REPORT, *supra* note 44, at 5–6.

²¹⁶ See FED. R. EVID. 702 advisory committee's note to 2000 Amendment (citing *Bourjaily v. United States*, 483 U.S. 171 (1987)).

²¹⁷ See *id.*

²¹⁸ See *id.*

²¹⁹ See *id.*

²²⁰ See Bernstein, *supra* note 193, at 31.

Much as with the lead-up to *Daubert*, supposedly scientific evidence is being used with increased frequency at sentencing today²²¹ and is nearly unregulated. In the DNA example discussed above, were the evidence offered at trial, the defendant would have a means for arguing its lack of reliability under Rule 702(d) or the state equivalent. But, at sentencing, the defendant would have no meaningful basis for making such an argument and might suffer a harsher penalty as a result. Just as a more stringent admissibility standard became necessary to stem the flow of unreliable expert evidence at trials decades ago, one is needed now to ensure that critical sentencing decisions are not based on unsound STS evidence.

2. *The Sentencing Phase: The Due Process Floor*

So what keeps out junk science at sentencing? Not much. By and large, trial-stage evidentiary rules do not apply at sentencing.²²² In federal cases, the federal Sentencing Guidelines allow sentencing courts to consider any “relevant information without regard to its admissibility under the rules of evidence applicable at trial,” so long as the information has only “sufficient indicia of reliability to support its probable accuracy.”²²³ The standard is similar in federal capital cases²²⁴ and

²²¹ See *supra* notes 92–98 and accompanying text.

²²² See, e.g., FED. R. EVID. 1101; ALA. R. EVID. 1101; ALASKA R. EVID. 101; *People v. Peterson*, 511 P.2d 1187, 1193 (Cal. 1973) (in bank); COLO. R. EVID. 1101; *State v. Young*, 778 A.2d 1015, 1019–20 (Conn. App. Ct. 2001); DEL. R. EVID. 1101; *Chandler v. State*, 534 So. 2d 701, 702–03 (Fla. 1988) (per curiam); HAW. REV. STAT. § 626-1, R. 1101 (2019); *State v. Hall*, 419 P.3d 1042, 1089 (Idaho 2018); *People v. Olson*, 2019 IL App (2d) 170334, ¶ 26; IND. R. EVID. 101; IOWA R. EVID. 1101; *State v. Sims*, 887 P.2d 72, 80 (Kan. 1994); KY. R. EVID. 1101; LA. CODE EVID. ANN. art. 1101 (2019); ME. R. EVID. 101; MD. R. EVID. 5-101; MASS. R. EVID. § 1101; *People v. Uphaus*, 748 N.W.2d 899, 905 (Mich. Ct. App. 2008); *State v. Sanchez-Sanchez*, 879 N.W.2d 324, 329 (Minn. 2016); MISS. R. EVID. 1101; *Martin v. State*, 291 S.W.3d 846, 849 (Mo. Ct. App. 2009); MONT. R. EVID. 101; *State v. Jenkins*, 931 N.W.2d 851, 885 (Neb. 2019); NEV. REV. STAT. § 47.020 (2019); N.H. R. EVID. 1101; *State v. Hawkins*, 719 A.2d 689, 692 (N.J. Super. Ct. App. Div. 1998); *State v. Smith*, 797 P.2d 984, 987 (N.M. Ct. App. 1990); *People v. Murray*, 709 N.Y.S.2d 806, 812 (N.Y. Sup. Ct. 2000); N.C. GEN. STAT. § 8C-1, R. 1101 (2019); N.D. R. EVID. 1101; *State v. Wright*, 2018-Ohio-965, 108 N.E.3d 1109, at ¶ 21 (8th Dist.); *Wade v. State*, 624 P.2d 86, 91 (Okla. Crim. App. 1981); OR. REV. STAT. ANN. § 40.015 (2019); *Commonwealth v. Medley*, 1999 PA Super 20, ¶ 13; R.I. R. EVID. 101; *State v. Gullledge*, 487 S.E.2d 590, 594 (S.C. 1997); S.D. CODIFIED LAWS § 19-19-1101 (2016); *Toledo v. State*, 519 S.W.3d 273, 289 (Tex. App. 2017); UTAH R. EVID. 1101; VT. R. EVID. 1101; *State v. Hixson*, 973 P.2d 496, 499 (Wash. Ct. App. 1999); *State ex rel. Dunlap v. McBride*, 691 S.E.2d 183, 193 (W. Va. 2010); *State v. Arredondo*, 2003 WI App 7, ¶ 53, 674 N.W.2d 647; *Peden v. State*, 2006 WY 26, ¶ 12, 129 P.3d 869, 872 (Wyo. 2006).

²²³ U.S. SENTENCING GUIDELINES MANUAL § 6A1.3 (U.S. SENTENCING COMM’N 2018); see also 18 U.S.C. § 3661 (“No limitation shall be placed on the information concerning the background, character, and conduct of a person convicted of an offense which a court of the United States may receive and consider for the purpose of imposing an appropriate sentence.”); Gert-

most state courts.²²⁵ That standard is derived from the extremely low constitutional due process standard, which requires only that information be “minimally reliable” in order to be considered at sentencing hearings.²²⁶

The “minimal reliability” rule stems from *Williams v. New York*,²²⁷ a seventy-year-old case widely considered to be the seminal case on the question of evidentiary standards at sentencing.²²⁸ In *Williams*, the Supreme Court considered whether, as a constitutional matter, due process requires that traditional rules of evidence apply at sentencing in the context of a defendant’s challenge to his sentencing judge’s consideration of out-of-court statements contained in his pro-

ner, *supra* note 86, at 543 (discussing *Daubert* and Rule 702). This means that even prejudicial or confusing evidence can be admitted at sentencing. See FED. R. EVID. 403.

224 See 18 U.S.C. § 3593(c) (“Information is admissible regardless of its admissibility under the rules governing admission of evidence at criminal trials except that information may be excluded if its probative value is outweighed by the danger of creating unfair prejudice, confusing the issues, or misleading the jury.” (emphasis added)).

225 See, e.g., ALA. CODE § 13A-5-45(d) (2019) (“Any evidence which has probative value and is relevant to sentence shall be received at the sentence hearing regardless of its admissibility under the exclusionary rules of evidence”); FLA. STAT. § 921.141(1) (2019) (“Any such evidence that the court deems to have probative value may be received, regardless of its admissibility under the exclusionary rules of evidence”); *Malenchik v. State*, 928 N.E.2d 564, 573–74 (Ind. 2010); cf. *State v. Loomis*, 2016 WI 68, ¶ 47, 371 Wis. 2d 235, 881 N.W.2d 749 (holding that due process requires that sentencing be based upon accurate information). But note that some capital-punishment states do maintain heightened procedural protections, including application of state rules of evidence, at the penalty phase in death cases. See, e.g., *Dumas v. State*, 803 N.E.2d 1113, 1121 (Ind. 2004) (holding that the rules of evidence apply to the penalty phase in capital cases).

226 See Gertner, *supra* note 86, at 542–43 (comparing the low standard of admissibility for evidence at sentencing with the higher standard of Rule 702); *United States v. Fields*, 483 F.3d 313, 337–38 (5th Cir. 2007) (“[D]ue process requires . . . some minimal indicia of reliability” (quoting *United States v. Petty*, 982 F.2d 1365, 1370 (9th Cir. 1993))); *United States v. McCarty*, 628 F.3d 284, 289 (6th Cir. 2010) (“[D]ue process requires simply that ‘some evidentiary basis beyond mere allegation in an indictment be presented to support consideration of such [uncharged and unconvicted] conduct as relevant to sentencing.’” (quoting *United States v. Silverman*, 976 F.2d 1502, 1512 (6th Cir. 1992))); *United States v. McGowan*, 668 F.3d 601, 606–08 (9th Cir. 2012) (reiterating that challenged information requires some “minimal indicium of reliability” to be relied on in sentencing (quoting *United States v. Vanderwerfhorst*, 576 F.3d 929, 936 (9th Cir. 2009))). But the minimal reliability rule is not designed to catch the most problematic evidence. *People v. Collins* provides another useful example. 438 P.2d 33 (Cal. 1968) (in bank). In that case, the statistical theory invoked by the prosecutor was itself sound. But without a judge having first analyzed the reliability of the prosecutor’s statistical evidence, the misapplication of that theory created a “mystique” that prevented the factfinder from assessing its probative value and resulted in a “miscarriage of justice.” *Id.* at 41.

227 337 U.S. 241 (1949).

228 See Young, *supra* note 79, at 308–18; 6 WAYNE R. LAFAYE, JEROLD H. ISRAEL, NANCY J. KING & ORIN S. KERR, CRIMINAL PROCEDURE § 26.4(a) (4th ed. 2017); Michaels, *supra* note 88, at 1772.

bation report.²²⁹ The Supreme Court ultimately denied the defendant's challenge²³⁰ and, in its holding, set a very low bar for what evidence can be considered at sentencing. Specifically, the Court held that due process allows consideration of information at sentencing irrespective of whether such evidence would be admissible at trial.²³¹ The *Williams* holding has since been interpreted to mean that due process requires only that information be minimally reliable in order to be considered at sentencing.²³²

Part of the Court's rationale for creating such a lax evidentiary standard was that judges must have as much information as possible about a defendant's behavior and character in order to fashion individualized punishment and impose a proper sentence.²³³ The Court emphasized that, historically, sentencing judges retained broad discretion to consider a variety of information in making punishment decisions.²³⁴ If trial-stage evidentiary procedures were applied at sentencing, the Court reasoned, sentencing judges would be deprived of much of the information on which they had historically relied.²³⁵ Ultimately, the Court decided that due process does not mandate trial-stage evidentiary procedures at sentencing hearings, reasoning that "[t]he due process clause should not be treated as a device for freezing the evidential procedure of sentencing in the mold of trial procedure."²³⁶

But the Court's rationale for maintaining the historical status quo in *Williams* was flawed: it prioritized judges' purported need for a *breadth* of information over *accuracy* of information.²³⁷ Commentators have criticized both the premise underlying that determination and its ultimate wisdom. First, contrary to the Court's rationale in *Wil-*

²²⁹ See *Williams*, 337 U.S. at 242, 244–45.

²³⁰ *Id.* at 252.

²³¹ *Id.* at 251.

²³² See, e.g., *State v. McGill*, 140 P.3d 930, 941 (Ariz. 2006) (en banc); *United States v. Powell*, 650 F.3d 388, 392 (4th Cir. 2011); *United States v. Silverman*, 976 F.2d 1502, 1512 (6th Cir. 1992); *United States v. Petty*, 982 F.2d 1365, 1367–68 (9th Cir. 1993); *United States v. Sprauer*, 358 F. App'x 871, 872 (9th Cir. 2009); *United States v. Reme*, 738 F.2d 1156, 1167 (11th Cir. 1984); see also Edward R. Becker, *Insuring Reliable Fact Finding in Guidelines Sentencing: Must the Guarantees of the Confrontation and Due Process Clauses Be Applied?*, 22 CAP. U. L. REV. 1, 5 (1993) ("The essential nature of this regime [of indeterminate punishment] is explained in *Williams v. New York*.").

²³³ See *Williams*, 337 U.S. at 247–48.

²³⁴ See *id.* at 246.

²³⁵ See *id.* at 250. The Court also expressed concern that the adoption of trial procedures at sentencing would restrain courts from improving the administration of justice. *Id.* at 251.

²³⁶ *Id.* at 251.

²³⁷ See *id.* at 247.

liams, judges had little discretion in sentencing decisions in the era of determinate sentencing, the model in place at the time of the founding.²³⁸ Second, and perhaps more importantly, many note that access to accurate and reliable information at sentencing protects against unjust results.²³⁹ Failure to screen out inaccurate STS evidence can confuse the factfinder and cause overreliance on unfounded evidence.²⁴⁰

Williams stands for the idea that evidentiary standards are not needed at sentencing to protect defendants' right to due process because judges both need and have the ability to wade through large amounts of information to arrive at the correct punishment decision. But, the underlying facts of *Williams*, coupled with its trajectory following the Supreme Court's decision, undercut that very premise.

Samuel Tito Williams, the defendant in *Williams*, was an African American man arrested and charged with the murder of a 15-year-old white girl.²⁴¹ The lone eyewitness to the murder, the decedent's brother, was present for the attack, but he was unable to identify Mr. Williams as the perpetrator when police brought Mr. Williams to his school for identification.²⁴² In fact, the brother testified on cross-examination that the actual perpetrator was a white man who was seven inches shorter than Mr. Williams.²⁴³ This eyewitness testimony so directly undermined the government's case that, after it was completed, the government told the trial judge that it was not relying on the brother's testimony to make its case.²⁴⁴ The only evidence of Mr. Williams' guilt came in the form of coerced confessions that had been secured as Mr. Williams testified at trial, after detectives had severely beaten him.²⁴⁵ Mr. Williams nonetheless was convicted.²⁴⁶

238 Sanders, *supra* note 84, at 105–06, 117; *see also* Young, *supra* note 79, at 310–11 (describing the *Williams* Court's failure to appreciate the rationale for not applying evidentiary rules at sentencing during the era of determinate sentencing).

239 *See* Young, *supra* note 79, at 305, 310–11; Carissa Byrne Hessick & F. Andrew Hessick, *Procedural Rights at Sentencing*, 90 NOTRE DAME L. REV. 187, 222 (2014) (describing the argument that judges need “unfettered access to information” to make proper sentencing decisions as “overly broad” in part because “[a]ccess to information is not an end in and of itself at sentencing; rather, the reason for broad access to information is to allow a judge to impose a more accurate sentence. . . . [P]rocedural rights . . . enhance the accuracy of the information provided to the court.”).

240 *See* *People v. Collins*, 438 P.2d 33, 38–41 (Cal. 1968) (in bank).

241 *See* Sanders, *supra* note 84, at 111; Lewis Wood, *Judge Has Power to Impose Death*, N.Y. TIMES, June 7, 1949, at 20.

242 *See* Sanders, *supra* note 62, at 796–97.

243 *Id.* at 795, 797.

244 *See* Sanders, *supra* note 84, at 113–14.

245 *See* Robert McG. Thomas, *Man Freed in Slaying Wins Damages*, N.Y. TIMES, Feb. 18,

The jury that convicted Mr. Williams declined to impose the death penalty.²⁴⁷ The trial judge overrode that recommendation.²⁴⁸ Calling Mr. Williams “a ‘menace’ to society,” the judge sentenced him to die by the electric chair.²⁴⁹

The probation report on which the sentencing judge relied in sentencing Mr. Williams to the chair mirrors that considered almost seventy years later by T.K.’s judge. It contained the very same type of purportedly scientific information that was at issue in T.K.’s case and that has become ubiquitous in sentencings today: psychological information of questionable reliability regarding the defendant’s character.²⁵⁰ Based in no small part on that report, the judge found Mr. Williams to be a “psychopathic liar whose personality [wa]s permeated with psychosexual habits of thought and conduct” who deserved to be sentenced to death.²⁵¹

The case of a man sentenced to death based on untested information of questionable value after almost certainly being falsely convicted of murder is what forms the basis for our current standard for the admission of STS evidence at sentencing hearings.²⁵² In fact, in the decades since, that holding has consistently been cited with approval²⁵³ and has been read to imbue trial judges with effectively unfettered discretion to admit and consider untested and potentially unreliable information at sentencing hearings.²⁵⁴

After spending sixteen years in prison, Mr. Williams was granted habeas relief in 1963 and released after his conviction was over-

1973, at 57, <https://www.nytimes.com/1973/02/18/archives/man-freed-in-slaying-wins-damages-sentenced-to-death-weiss-urges.html> [<https://perma.cc/DMR5-XVMU>].

²⁴⁶ See *id.*

²⁴⁷ See *id.*

²⁴⁸ See *id.*

²⁴⁹ See *id.*

²⁵⁰ See *Youth Sentenced to Die in the Chair*, N.Y. TIMES, Mar. 3, 1948, at 48, <https://times-machine.nytimes.com/timesmachine/1948/03/03/94940055.html?pageNumber=48> [<https://perma.cc/W98M-K3ZC>].

²⁵¹ Thomas, *supra* note 245.

²⁵² See *Williams v. New York*, 337 U.S. 241, 253 (1949) (Murphy, J., dissenting) (“The record before us indicates that the judge exercised his discretion to deprive a man of his life, in reliance on material made available to him in a probation report, consisting almost entirely of evidence that would have been inadmissible at the trial.”).

²⁵³ See Young, *supra* note 79, at 316; LAFAYE ET AL., *supra* note 228; see also *Williams v. Oklahoma*, 358 U.S. 576, 584 (1959) (holding that under *Williams v. New York*, a “State’s Attorney’s statement of the details of the crime and of petitioner’s criminal record—all admitted by petitioner to be true—did not deprive petitioner of fundamental fairness or of any right of confrontation or cross-examination”).

²⁵⁴ See Sanders, *supra* note 84, at 105.

turned.²⁵⁵ Ten years after that, Mr. Williams was awarded \$120,000 in damages for malicious prosecution and false arrest.²⁵⁶ The trial judge in *Williams*—who issued a death sentence based on a sentencing report containing untested psychological information, leading the Supreme Court to establish a rule that evidentiary protections are not needed at sentencing—seems to have gotten it entirely wrong.

Over seventy years later, as T.K.'s case makes clear, judges are still getting it wrong by relying on flawed STS evidence at sentencing. The stakes at sentencing today, though, are much higher than they have ever been and are far too high to allow *Williams* to carry the day. What follows is a proposal for an alternative rule that will enable judges to adequately screen out junk science while allowing them to consider relevant, important information about defendants at sentencing.

II. EXTENDING RULE 702 TO SENTENCING

Williams set a floor, not a ceiling, for what is required before evidence can be considered at sentencing. This Section advances a straightforward proposal that states and rulemaking bodies can follow to fill the gap created by *Williams* and prevent junk STS evidence from being used against defendants at sentencings: extending Rule 702—which governs admissibility of the same type of evidence at the trial stage—or its state equivalent to sentencing hearings when such evidence is introduced against a defendant. In other words, it proposes that STS evidence offered by the prosecutor or a neutral party in support of an argument for an increased punishment—but not information offered by the defendant for mitigation—be subjected to an admissibility test.

This approach would not overrule *Williams*. It would address the shortcomings of that case while protecting the interest that it sought to advance: allowing judges to consider a broad range of information at sentencing. The sentencing process necessarily involves consideration of a broad range of factors, but a meaningful solution to the problem of junk science at sentencing must find a way to screen out unreliable STS evidence that might unfairly cost a defendant extra years in prison. It must also do so without infringing on a defendant's right to present mitigating evidence.

²⁵⁵ See Thomas, *supra* note 245.

²⁵⁶ See *id.*

A. *The Proposed Procedure*

Rules of evidence do not typically apply at sentencing,²⁵⁷ so Rule 702 (or its state equivalents) would need to be adjusted somewhat before being applied at sentencing. Most relevantly, Rule 702 applies to evidence offered through a live expert witness (e.g., a DNA analyst, psychologist, or fingerprint examiner)²⁵⁸ because hearsay is generally inadmissible at the trial stage.²⁵⁹ Because it is neither feasible nor advisable to require that any STS evidence offered at sentencing be offered through a live witness, the language of the rule proposed herein has been modified to accommodate evidence offered without a live witness. The proposed rule, consistent with the asymmetrical approach described above, is as follows:

Scientific, technical, or other specialized evidence may be admitted at sentencing by a party other than the defendant only if:

- (a) the evidence will help the sentencer to understand the evidence, to determine a fact in issue or to render a just sentence;
- (b) the evidence is based on sufficient facts or data;
- (c) the evidence is the product of reliable principles and methods;
- (d) the evidence is the product of reliable application of the principles and methods to the facts of the case; and
- (e) any witness who produced such evidence is qualified as an expert by knowledge, skill, experience, training, or education.

This proposal maintains Rule 702's substantive requirements for admission of STS evidence, but it tailors these requirements to the sentencing phase. While it allows for the admission of hearsay forms of STS evidence, it requires that, regardless of whether the witness who produced the STS evidence in question testifies at sentencing, that witness must possess the same expert qualifications (expertise by knowledge, skill, experience, training or education) required by Rule 702.²⁶⁰ It also modifies the relevance standard for the sentencing stage by allowing admission of the proposed evidence if it aids in under-

²⁵⁷ See sources cited *supra* note 222.

²⁵⁸ This is explicit in the language of the rule, which begins, "A witness who is qualified as an expert by knowledge, skill, experience, training, or education may *testify* in the form of an opinion or otherwise if" FED. R. EVID. 702 (emphasis added).

²⁵⁹ See, e.g., FED. R. EVID. 802.

²⁶⁰ See FED. R. EVID. 702.

standing evidence or determining a fact in issue or if it assists in rendering a just sentence.

T.K.'s case can be used to illustrate the ways that a judge should evaluate STS evidence at sentencing. The first two parts of the proposed rule, parts (a) and (b), are straightforward. They track the language of and are substantively identical to parts (a) and (b) of Rule 702. Part (a) of the proposed rule assesses relevance. For the SAVRY to meet part (a) of the proposed rule, a judge would ask if, taken together with all of the other information provided, the results of the SAVRY would aid in his sentencing determination or if they would be superfluous.²⁶¹ Part (b) of the proposed rule requires a quantitative analysis. It asks whether there is *enough* data to support the conclusion regarding T.K.'s future dangerousness.²⁶²

Parts (c) and (d) of the proposed rule are also substantively identical to Rule 702(c) and (d). They comprise the reliability test codified in those parts. For the SAVRY to meet part (c) of the proposed rule above—methodological reliability—several factors might be considered: Has the SAVRY been validated or tested? On what population? Does the population on which it was validated mirror the population it is being used to assess? Are the risk factors the SAVRY purports to assess based on research and data or on speculation and guesswork? Has research on the SAVRY been subjected to peer review? What is the error rate of the tool? Can the error rate of a risk assessment tool be calculated? How is examiner discretion accounted for? Are the risk factors the tool treats as predictors of future violence linked to actual violence? Take, for example, the SAVRY's treatment of prior convictions as predictors of risk for future violence. Could the SAVRY be considered reliable if the tool were to treat *any* prior conviction—even a conviction for nonviolent conduct—as a predictor of future violence? These are just a few factors that may be explored to evaluate whether the SAVRY represents a generally reliable method for predicting children's potential for future violence.

To establish prong (d) of the proposed rule—validity as applied—it must be determined that application of the SAVRY is appropriate in the specific case even if it is a sound method for assessing future violence potential generally. Two factors that might be considered in conducting this analysis are (1) whether the tool was misapplied, perhaps by assigning a higher risk rating than was warranted, and (2) whether a risk rating was mistakenly assigned to a factor uncorrelated with risk

²⁶¹ See FED. R. EVID. 702 advisory committee notes.

²⁶² See *id.*

of future dangerousness. Analysis of these factors might reveal that *application* of the risk assessment instrument was not reliable even if the tool itself were deemed to be theoretically sound if used correctly.²⁶³

An evaluation of prong (e) might focus on the training, competence, and experience of the evaluator conducting the assessment.

If courts were required to perform a gatekeeping function and apply the procedure laid out above at sentencing, it is likely that psychiatric testimony on future dangerousness—like that offered in Duane Buck’s case—would fail to meet expert admissibility requirements under Rule 702.²⁶⁴ In fact, judges and commentators who have analyzed future dangerousness testimony under the *Daubert* standard have argued that, despite its continued use in criminal cases, it would be found inadmissible due to its lack of reliability.²⁶⁵

B. Analysis of the Proposed Extension of Rule 702

The proposed rule is designed to be an administrable improvement to, not an overhaul of, current sentencing practice. Though absolute uniformity in the way judges apply the standard cannot be guaranteed, this proposal asks judges and lawyers to apply a standard they already use regularly at the trial stage. The standard is clearly defined, laying out each factor to be considered in determining admissibility, and has been in existence for many years, so there is well-developed case law to guide its application where ambiguities arise.

Its scope is also narrow: it aims to restrict consideration of STS information of doubtful validity without limiting the sentencer’s ability to collect and consider non–STS information about a defendant’s character, criminal history, or other factors that inform punishment.

Still, the proposal’s asymmetrical construction, the potential for strain on judicial resources in the form of cost, and delay, and the ineffectiveness of the rule in restricting admissibility of expert evidence at the trial stage may draw criticism. This Section responds to these potential criticisms.

²⁶³ In fact, the evaluator in T.K.’s case did misapply the SAVRY in quite a few of the described ways. Order, *supra* note 3, at 8–9.

²⁶⁴ See *Flores v. Johnson*, 210 F.3d 456, 464–65 (5th Cir. 2000) (Garza, J., specially concurring).

²⁶⁵ See *id.*; Erica Beecher-Monas, *The Epistemology of Prediction: Future Dangerousness Testimony and Intellectual Due Process*, 60 WASH. & LEE L. REV. 353, 369 (2003).

1. Asymmetry

Some may hesitate to embrace this proposal because of its asymmetrical application. On its face, it may seem inequitable given that, at trial, Rule 702 and its state counterparts apply to all expert evidence, regardless of the party seeking its admission. Yet several factors necessitate and justify an asymmetrical approach in dealing with junk science at sentencing.

As discussed above, junk science has been a major contributor to wrongful convictions and other miscarriages of justice at the trial stage.²⁶⁶ This is partly because the government is more likely to produce flawed scientific information at any stage of a case and indeed, prosecutors frequently utilize untested STS evidence of questionable validity.²⁶⁷ On top of that, flawed STS evidence is increasingly being admitted at sentencings, where it can confuse and unduly influence sentencers and cause significant problems for defendants, undermining trust in sentencings.²⁶⁸

The takeaway from all of this is that there has to be some mechanism for screening STS evidence before it may be considered by a judge or capital jury. But the different stakes and competing interests of the parties mitigate against a blanket application of Rule 702 and support applying it only to STS evidence introduced to increase a sentence. J.J. Prescott and Sonja Starr note that sentencing frequently results in harsh punishment and that allowing defendants latitude in introducing mitigating evidence can lessen that harshness.²⁶⁹ As they

²⁶⁶ See *supra* Section I.A.

²⁶⁷ 1 DAVID L. FAIGMAN, EDWARD K. CHENG, JENNIFER L. MNOOKIN, ERIN E. MURPHY, JOSEPH SANDERS & CHRISTOPHER SLOBOGIN, *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 1:35 (2019). Prosecutors' resources also contributed to their use of flawed STS evidence; for the most part, it is the government, not the average criminal defendant, that has the resources to hire experts and forensic laboratories to analyze evidence. See Anthony LoMonaco, Note, *Disproportionate Impact: An Impetus to Raise the Standard of Proof at Sentencing*, 92 N.Y.U. L. REV. 1225, 1252–53 (2017); Laurel Gilbert, Comment, *Sharpening the Tools of an Adequate Defense: Providing for the Appointment of Experts for Indigent Defendants in Child Death Cases Under Ake v. Oklahoma*, 50 SAN DIEGO L. REV. 469, 487–90 (2013). In criminal cases, most expert evidence is offered by the government. *Id.* at 490 n.44. And, forensic science labs work closely with law enforcement and prosecutors, creating an atmosphere that encourages pro-prosecution bias and can result in admission of flawed STS evidence. Gilbert, *supra*, at 487–88.

²⁶⁸ See *supra* Section I.B.

²⁶⁹ J.J. Prescott & Sonja Starr, *Improving Criminal Jury Decision Making After the Blakely Revolution*, 2006 U. ILL. L. REV. 301, 321–22. Prescott and Starr make this argument in the context of determinate sentencing systems. *Id.* at 322. But harsh sentences do not only arise in determinate systems. See *supra* notes 65–71 and accompanying text (describing significant increase in sentence length since the start of the prison population explosion).

suggest, this is a forceful argument for an asymmetrical approach to adoption of evidentiary rules at sentencing²⁷⁰ because the criminal legal system, at least theoretically, places a higher premium on protecting liberty than on securing detention.²⁷¹

The harms caused by STS evidence are radically different based on why they are used. Unreliable STS evidence offered against a defendant at sentencing may contribute to a significant loss of liberty, or much worse—it may cost a defendant his life.²⁷² Evidence from a defendant, however, does not harm the government in anything close to the same way, and it is a valuable part of a system that recognizes the importance of a defendant's right to do almost everything he or she can to defend him or herself. Indeed, the right to present mitigating evidence in capital and, for juveniles, life without parole cases, is constitutionally required²⁷³ and many jurisdictions allow broad presentation of mitigating evidence by convicted defendants in noncapital cases as a statutory right, public policy matter, or principle of justice.²⁷⁴ The need to screen STS evidence at sentencing must therefore be reconciled with the need to allow defendants to fight for leniency at a stage where the significant potential for loss of liberty warrants additional protections.²⁷⁵ The instant proposal does this by structuring a solution that screens the most problematic STS evidence but recognizes defendants' interest in presenting mitigating evidence at this phase.

It is possible that if trial-stage protections that promote reliability of evidence and protect defendants in light of the power imbalance between the parties (e.g., the right to have facts proven beyond a rea-

²⁷⁰ Prescott & Starr, *supra* note 269, at 321–22.

²⁷¹ 4 WILLIAM BLACKSTONE, COMMENTARIES *357 (“[F]or the law holds that it is better that ten guilty persons escape than that one innocent suffer.”); *cf.*, e.g., Tipton v. State, 2013-CA-00415-SCT (¶ 14) (Miss. 2014) (“We find that a person who serves time in Mississippi’s Intensive Supervision Program (ISP) is entitled to compensation if wrongly convicted.”).

²⁷² See Barefoot v. Estelle, 463 U.S. 880, 916 (1983) (Blackmun, J., dissenting) (“In a capital case, the specious testimony of a psychiatrist . . . equates with death itself.”).

²⁷³ Lockett v. Ohio, 438 U.S. 586, 604 (1978) (“[T]he Eighth and Fourteenth Amendments require that the sentencer, in all but the rarest kind of capital case, not be precluded from considering, as a mitigating factor, any aspect of a defendant’s character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death.” (footnote omitted)); Miller v. Alabama, 567 U.S. 460, 489 (2012) (“[I]ndividualized sentencing decisions make clear that a judge or jury must have the opportunity to consider mitigating circumstances before imposing the harshest possible penalty . . .”).

²⁷⁴ See, e.g., People v. Bush, 238 N.E.2d 105, 108 (Ill. App. Ct. 1968); State v. Sullivan, 2018 VT 112, ¶¶ 25–26, 200 A.3d 670.

²⁷⁵ See Apprendi v. New Jersey, 530 U.S. 446, 484 (2000) (recognizing that the potential for loss of liberty and the need for protections of the defendant are linked).

sonable doubt and the right to confront) also applied at sentencing, perhaps an asymmetrical approach would not be necessary under this proposal. If prosecutors had to do more to persuade a sentencer that a harsher sentence is necessary in order for it to be imposed and defendants had greater ability to expose unreliable evidence through confrontation and cross-examination, then Rule 702 might fairly be applied to both parties at sentencing hearings. Even as sentencing hearings have evolved to become more complex and the stakes for defendants have become more serious, extension of constitutional and evidentiary trial-stage protections has lagged.²⁷⁶ As it stands, then, defendants both need to be protected and need to be afforded great flexibility to seek leniency, making asymmetrical application of the proposed rule necessary.

This is not a radical departure from how our system currently operates. As mentioned, defendants already enjoy many protections in the criminal process not extended to prosecutors, witnesses, or others. In certain capital jurisdictions, versions of this very proposal already exist: in three states, rules of evidence apply to the presentation of aggravating evidence, but not mitigating evidence, at capital sentencing proceedings.²⁷⁷ In effect, these jurisdictions preclude prosecutors from presenting aggravating evidence, if it is not admissible under the state's rules of evidence, but they do not impose the same limitations on mitigating evidence from the defense, just as the instant proposal seeks to do with STS evidence. These systems are directly analogous to the proposal advanced herein. Underlying all of these systems is a recognition of the importance of a defendant's substantial interest in seeking leniency.²⁷⁸

Criminal defendants also enjoy constitutional and evidentiary protections to ensure the reliability of evidence introduced against them at trial in non-STs contexts. The right to have facts proven beyond a reasonable doubt at trial and the right to confront witnesses are perhaps the best examples of constitutionalized protections in-

²⁷⁶ See *supra* Section I.C. But note that the Supreme Court has recognized some limited constitutional protections at sentencing. See, e.g., *Apprendi*, 530 U.S. at 490 (noting that the right to trial by jury and the right to have facts proven beyond a reasonable doubt apply to facts that may increase a sentence beyond the statutory maximum); *Ring v. Arizona*, 536 U.S. 584, 609 (2002) (holding that aggravating factors must be found by a jury to impose the death penalty under the Sixth Amendment); see also Hessick & Hessick, *supra* note 61, at 53–56 (outlining additional constitutional rights that have been recognized to apply at sentencing).

²⁷⁷ See ARIZ. REV. STAT. ANN. § 13-751 (2020); ARK. CODE ANN. § 5-4-602 (2020); NEB. REV. STAT. §§ 29-2520 to -2521 (2019).

²⁷⁸ See Kelly, *supra* note 83, at 438.

tended to promote the reliability of evidence offered against defendants.²⁷⁹ Similarly, Federal Rule of Evidence 404, which governs the admissibility of character evidence, applies to evidence against defendants but not to government evidence.²⁸⁰ While the prosecutor may not offer evidence at trial of a defendant's bad character, the *defendant* may introduce evidence of a *victim's* bad character.²⁸¹ Another example is Federal Rule of Evidence 609, which requires the prosecution to satisfy more criteria when attempting to impeach a defendant with prior convictions than it requires any party satisfy when making the same attempt with any other witness.²⁸²

Under the federal capital sentencing scheme, aggravating factors must be established beyond a reasonable doubt, while mitigating factors must only be established by a preponderance of the evidence.²⁸³ Likewise, aggravating factors must be found unanimously, while mitigating factors need only be found by a single juror.²⁸⁴

The asymmetrical approach advanced herein embraces minimizing harm to criminal defendants as a goal of sentencing reform and thus avoids creating tension with capital sentencing law and noncapital sentencing policy. A more expansive approach—for instance, one promoting extension of all rules of evidence to sentencing—would necessarily limit defendants' ability to introduce mitigating evidence and might not survive a challenge to its lawfulness under the Eighth and Fourteenth Amendments in capital sentencing schemes, potentially creating tension with sentencing policy even in the noncapital sentencing context.²⁸⁵

Finally, the asymmetry of the rule would not pose an undue burden on the government; it will not exclude *reliable* STS evidence at sentencing, only *unreliable* STS evidence. Moreover, Rule 702 and its state counterparts are rules of admissibility, not weight—the government is free to argue that a judge should not rely heavily on STS evidence of questionable validity. It has the resources to make such challenges and does so with regular success.²⁸⁶

²⁷⁹ See *Maryland v. Craig*, 497 U.S. 836, 845 (1990) (“The central concern of the Confrontation Clause is to ensure the reliability of the evidence against a criminal defendant by subjecting it to rigorous testing in the context of an adversary proceeding before the trier of fact.”).

²⁸⁰ See FED. R. EVID. 404.

²⁸¹ See *id.*

²⁸² See FED. R. EVID. 609.

²⁸³ 18 U.S.C. § 3593(c).

²⁸⁴ *Id.* § 3593(d).

²⁸⁵ See *supra* notes 273–74 and accompanying text.

²⁸⁶ Stephanie L. Damon-Moore, Note, *Trial Judges and the Forensic Science Problem*, 92

2. Cost, Delay, and Other Practical Concerns

Extension of Rule 702 or a corresponding state rule to sentencing may add costs and create delay in those cases in which a defendant seeks exclusion of STS evidence offered against her. But the potential for major adverse consequences to defendants where a sentencing decision is based on unreliable STS evidence justifies the additional cost and delay caused by screening out junk science at sentencing.

Defendants challenge the admissibility of STS evidence proffered by the government in only a small percentage of trials.²⁸⁷ It is true that many more defendants are sentenced than go to trial, which might result in an absolute increase in the number of challenges to STS evidence. But if that number is significant, it only underscores the need for the reform proposed herein. As STS evidence—and, by extension, unreliable STS evidence—is becoming more and more common in the many cases that proceed to sentencing,²⁸⁸ the response cannot be to leave the status quo in place by admitting all such evidence in order to avoid burdening the judicial system.

In any event, this proposal is tailored to limit cost and delay because it does not create a brand-new rule. Adoption of a new rule might result in greater cost and delay while parties and courts learn and adapt to the application of a new admissibility standard. Parties are already familiar with the proposed rule. Prosecutors are equipped to present information that meets admissibility standards and regularly do so at trial while courts are also experienced in applying and interpreting the rule.²⁸⁹

As cases are litigated, the contours of what is and is not admissible will be determined, limiting the need for extended litigation about the admissibility of different kinds of evidence. It will also raise the standard for scientific evidence because subjecting STS evidence to *Daubert* and Rule 702 will lead to reexamination and renewed scrutiny of the underlying science,²⁹⁰ resulting in better STS evidence being offered in the future.

N.Y.U. L. REV. 1532, 1557 (2017); FAIGMAN ET AL., *supra* note 267, § 1:35. The same is not true of the typical criminal defendant. See FAIGMAN ET AL., *supra* note 267, § 1:35; Gilbert, *supra* note 267, at 488.

²⁸⁷ See FAIGMAN ET AL., *supra* note 267, § 1:35.

²⁸⁸ See *supra* notes 92–98 and accompanying text.

²⁸⁹ See Deborah Young, *Untested Evidence: A Weak Foundation for Sentencing*, 5 FED. SENT'G REP. 63, 65 (1992).

²⁹⁰ See Garrett, *supra* note 172, at 67; Paul C. Giannelli, *Forensic Science: Under the Microscope*, 34 OHIO N.U. L. REV. 315, 320–21 (2008).

3. *The Effectiveness of Rule 702 and Daubert at Trial*

An important potential critique of the proposal made here is that, as many commentators have argued, Rule 702 and *Daubert* have not been particularly effective in filtering out problematic expert evidence at the trial phase²⁹¹ so it might make little sense to try to extend it to other contexts. In their seminal treatise on scientific evidence, David Faigman and his coauthors have complained that, with respect to forensic science evidence, “courts have been, at best, lackadaisical and, at worst, disingenuous, in carrying out their gatekeeping duties.”²⁹² Moreover, courts are far more lenient in admitting faulty scientific evidence offered by prosecutors than towards such evidence offered by criminal defendants.²⁹³

These concerns have merit. This Article does not seek to inflate the value of Rule 702 and *Daubert* and suggest that they perfectly filter out unsound STS evidence. It suggests only that, at the very least, the same admissibility standard applied at trial should be adopted at sentencing. This is because the rule is already in place and well known, and, as a result, adoption will be straightforward. Imperfect though it may be, extension of Rule 702 to sentencing will at least begin the process of filtering out faulty STS evidence at sentencing. Importantly, once applied, it can be updated as improvements are made to strengthen Rule 702 at the trial stage.

Even still, Rule 702 and *Daubert* have not been total failures. Since *Daubert* was issued and Rule 702 was amended to codify its ruling, there has been substantial renewed scrutiny of questionable scientific evidence.²⁹⁴ And when applied properly, the Rule is quite effective in exposing the unreliability of faulty STS evidence.²⁹⁵

²⁹¹ See, e.g., Epstein, *supra* note 44, at 84–86; Koehler, *supra* note 45, at 1389, 1395. See generally Garrett & Fabricant, *supra* note 45 (discussing how the *Daubert*/Rule 702 reliability test has done little to keep out unreliable evidence); Giannelli, *supra* note 45 (same).

²⁹² FAIGMAN ET AL., *supra* note 267, § 1:30.

²⁹³ See *id.*; Garrett & Fabricant, *supra* note 45, at 1561–62; Damon-Moore, *supra* note 286, at 1557.

²⁹⁴ See Bernstein, *supra* note 193, at 69; Garrett, *supra* note 172, at 67. See generally NAS REPORT, *supra* note 45 (reporting the results of a congressionally mandated study on, and making recommendations to improve, forensic science).

²⁹⁵ See, e.g., Garrett, *supra* note 172, at 71–78 (describing the North Carolina Court of Appeals’ analysis of fingerprint evidence under Rule 702(d) in *State v. McPhaul*, 808 S.E.2d 294, 303 (N.C. Ct. App. 2017), *disc. review allowed* by 812 S.E.2d 847 (N.C. 2018), and *disc. review improvidently allowed*, 818 S.E.2d 102 (N.C. 2018)).

III. OTHER PROPOSALS TO IMPROVE SENTENCING ACCURACY

Many scholars have noted the importance of the sentencing phase in today's criminal legal system,²⁹⁶ share this author's concerns over the accuracy of criminal sentences,²⁹⁷ and have called for the application of trial-stage rights and procedures to sentencing or for other sentencing reforms.²⁹⁸ This Section reviews a reflective sampling of such proposals and their suitability for addressing the effect junk science has on sentencings. It concludes that, though valuable in their own right, none of the proposals adequately addresses the precise problem considered in this Article: how to minimize the contribution of STS evidence of questionable soundness to unjust sentences without creating added harms for criminal defendants.

A. *Proposals to Extend Rules of Evidence to Sentencing*

A common suggestion for reforming sentencing hearings is to extend either all, some, or a specialized version of the rules of evidence to sentencing hearings. This Section contains an analysis of two such proposals.

Almost thirty years ago, Third Circuit Judge Edward Becker and Aviva Orenstein noted that the adoption of the Federal Sentencing Guidelines ("Sentencing Guidelines") in 1987²⁹⁹ increased the importance of the sentencing phase in federal criminal cases.³⁰⁰ They explained that adoption of the then-mandatory³⁰¹ Sentencing Guidelines

²⁹⁶ See, e.g., Sanders, *supra* note 62, at 793 ("Now more than ever, factual findings made during felony sentencing hearings are as quantitatively vital as those that were previously only made during trials."); Hessick & Hessick, *supra* note 61, at 56 ("Sentencing is a critically important phase of the criminal justice system for determining the punishment a defendant will receive, and it continues to grow in importance because most criminal defendants plead guilty rather than proceed to trial.").

²⁹⁷ See, e.g., Young, *supra* note 79, at 302–03.

²⁹⁸ See Brandon L. Garrett, *Accuracy in Sentencing*, 87 S. CAL. L. REV. 499, 539–40 (2014) (collecting sources written by scholars that propose a range of sentencing reforms); Young, *supra* note 289, at 63, 65 (same).

²⁹⁹ U.S. SENTENCING GUIDELINES MANUAL § 1A1 (U.S. SENTENCING COMM'N 2018).

³⁰⁰ See Edward R. Becker & Aviva Orenstein, *The Federal Rules of Evidence After Sixteen Years—The Effect of "Plain Meaning" Jurisprudence, the Need for an Advisory Committee on the Rules of Evidence, and Suggestions for Selective Revision of the Rules*, 60 GEO. WASH. L. REV. 857, 887–91 (1992). Though Becker and Orenstein's discussion of the impact of the Sentencing Guidelines on federal sentencing occurred when the Sentencing Guidelines were considered mandatory, their argument has continued force today. The Sentencing Guidelines continue to be highly influential. See LoMonaco, *supra* note 267, at 1246–49. This is because, *inter alia*, sentencing judges are required to consider the Sentencing Guidelines, 18 U.S.C. § 3553, and because they continue to treat the Sentencing Guidelines as a benchmark for sentencing decisions. See LoMonaco, *supra* note 267, at 1247–48.

³⁰¹ The Federal Sentencing Guidelines were considered mandatory until the Supreme

resulted in sending a greater percentage of convicted defendants to prison and longer sentences for certain crimes,³⁰² largely because sentences depended on factual findings made at sentencing hearings, including about the commission of uncharged offenses.³⁰³ Almost thirty years ago, Judge Becker and Orenstein saw the problem clearly: reliance upon evidence inadmissible at the trial stage at sentencing increases sentences in ways that are not justified.³⁰⁴ They called sentencing hearings “arguably the most important judicial business conducted by Article III judges”³⁰⁵ and explained that “the de facto liberty interest is tried at the sentencing hearing, not the trial, and yet the evidentiary standards required in trials are absent.”³⁰⁶

Consequently, Judge Becker and Orenstein argued for the application of specialized rules of evidence written specifically for sentencing.³⁰⁷ They posited that the framers of the Federal Rules of Evidence, in deciding that the Rules would not apply at sentencing, did not anticipate the massive changes to sentences that adoption of the Sentencing Guidelines created.³⁰⁸ Judge Becker and Orenstein thus argued for the extension of selected rules of evidence to sentencing.³⁰⁹ Alternatively, they recommended that courts apply a “*strong* reliability standard” to evidence that might have a “significant impact” on a sentence.³¹⁰

Others have also suggested extending existing evidentiary rules to sentencing hearings. Deborah Young, for instance, argued that the Federal Rules of Evidence should apply at federal sentencing to minimize inaccurate sentences and to prevent defendants from being injured by such sentences.³¹¹ She notes that there are two “key harms” that courts should avoid at sentencing: inaccurate factfinding and allowing defendants to suffer the consequences of sentencing errors.³¹²

Court struck down their mandatory provisions in 2005. *United States v. Booker*, 543 U.S. 220, 226–27 (2005).

³⁰² Becker & Orenstein, *supra* note 300, at 889–90.

³⁰³ For a basic overview of how guidelines calculations are performed, see LoMonaco, *supra* note 267, at 1232–34.

³⁰⁴ Becker & Orenstein, *supra* note 300, at 887–90.

³⁰⁵ *Id.* at 889.

³⁰⁶ *Id.* at 890.

³⁰⁷ *See id.*

³⁰⁸ *See id.*

³⁰⁹ *Id.*

³¹⁰ *Id.* at 891. Becker and Orenstein suggested that this standard might be equated to a clear and convincing standard. *Id.*

³¹¹ Young, *supra* note 79, at 302–04.

³¹² *Id.* at 302.

She suggests that adoption of rules of evidence at sentencing would be both a practically feasible, efficient, and comprehensive means of improving sentencing and a way to minimize the possibility that sentencing errors harm defendants.³¹³ She emphasizes the comprehensive nature of the broad extension of rules of evidence as a beneficial contrast to taking a “piecemeal” approach to solving specific sentencing problems.³¹⁴

B. Other Proposals to Improve Sentencing Accuracy

Other critics have focused less on rules of evidence than on a variety of other reforms to the sentencing phase. Carissa Byrne Hessick and F. Andrew Hessick argue that traditional justifications for denying defendants procedural rights at sentencing are unfounded and unpersuasive³¹⁵ and that each procedural right should be examined individually to determine whether it should be enforced at sentencing.³¹⁶

They examine the rights that apply at sentencing in two types of sentencing systems: mandatory systems, in which a particular sentence is indicated based on specific facts, such as mandatory guidelines systems and capital sentencing systems; and discretionary systems, in which punishment is decided largely by a judge with broad discretion.³¹⁷ Hessick and Hessick note that courts have enforced a number of procedural rights in mandatory schemes—including the right to have facts that increase a sentence found by a jury, the right to have such facts proven beyond a reasonable doubt, the right to notice of factors that may increase a sentence, and the right not to be sentenced under retroactive laws—but that these protections have not been held to apply in discretionary systems.³¹⁸

They discuss three rationales that traditionally have been offered to justify greater procedural protections in mandatory sentencing schemes than in discretionary sentencing schemes: (1) the idea that judicial discretion negates the need for procedural protections at sentencing; (2) the notion that defendants have an expectation that they will not receive a sentencing increase without a finding of aggravating circumstances in mandatory systems; and (3) a “history-based” ratio-

³¹³ *Id.* at 302–03.

³¹⁴ *Id.* at 303.

³¹⁵ See Hessick & Hessick, *supra* note 239, at 187, 189–90.

³¹⁶ *Id.* at 234.

³¹⁷ *Id.* at 188. Hessick and Hessick are careful to clarify, however, that no sentencing system is entirely mandatory or entirely discretionary. *Id.*

³¹⁸ *Id.*

nale that aggravating facts are not merely sentencing factors, but elements of a crime requiring greater procedural protection.³¹⁹

After an in-depth analysis of each rationale, Hessick and Hessick conclude that incongruent application of procedural rights in mandatory but not discretionary schemes is not justified.³²⁰ Specifically, they note that the argument that procedures will interfere with judges' abilities to collect information about a defendant and use their discretion to fashion an appropriate sentence, much like the rationale of the *Williams* holding,³²¹ is too broad because it fails to account for the role of procedural rights in improving the accuracy of information relied upon by the sentencing judge.³²² They note, too, that both discretionary and mandatory systems require factfinding at sentencing, and thus accuracy of information is just as important in the former as it is in the latter.³²³ With regard to the second two rationales, they argue that defendants have strong expectations even in discretionary systems, and the historical aggravators-as-elements rationale is based on a faulty understanding of historical practice.³²⁴

Ultimately, Hessick and Hessick conclude that enforcing procedural rights based on the type of sentencing scheme is not defensible.³²⁵ They argue that resolving the question of what rights must be enforced at sentencing should depend on the purpose of the procedure, rather than the type of system to which it is being applied.³²⁶ Because different rights serve different purposes, they argue, each should be analyzed individually to determine whether it should apply at sentencing.³²⁷

Brandon Garrett takes a back-end approach to improving sentencing accuracy, focusing on the remedies available to correct a sentencing error after the fact. He notes that sentencing errors can occur in myriad ways but that not all types of sentencing errors can be remedied under the same standard, or at all.³²⁸ He argues that a single, consistent miscarriage of justice standard that "asks whether a reasonable judge would 'more likely than not'" have sentenced a defendant

³¹⁹ *Id.* at 189–90, 221–33.

³²⁰ *Id.* at 234.

³²¹ *See supra* Section I.C.2.

³²² Hessick & Hessick, *supra* note 239, at 222.

³²³ *See id.* at 223.

³²⁴ *Id.* at 227–33.

³²⁵ *Id.* at 233.

³²⁶ *Id.*

³²⁷ *Id.* at 233–34.

³²⁸ Garrett, *supra* note 298, at 500–03.

differently but for an error at sentencing should apply to review of all sentencing errors, regardless of the type of error or procedural context in which it is raised.³²⁹

Others have proposed various wide-ranging reforms to improve sentencing accuracy and reliability.³³⁰ Some have argued that to improve the reliability of sentencing decisions, the standard of proof at sentencing should be raised³³¹ or that confrontation rights should be recognized.³³² Shaakirrah Sanders recognizes that in the modern system of mass plea bargaining, facts pertinent to the sentencing decision are commonly established after the plea at sentencing,³³³ similar to trials, yet they do not ensure a defendant's right to confront evidence against him and thus create a risk that the defendant will be sentenced based on unreliable evidence.³³⁴ Sanders concludes that where testimonial statements are material to punishment and where cross-examination will be useful in determining the veracity of such statements, confrontation rights should extend to felony sentencing.³³⁵

C. *The Suitability of These Proposals*

Each of the proposals discussed above is likely to lead to some degree of increased accuracy and reliability in sentencing. But because their focus is different than that of the instant proposal, none quite adequately addresses the precise problem of how to minimize unjust

³²⁹ *Id.* at 504. Garrett also acknowledges that a number of potential front-end sentencing reforms may be useful in reducing sentencing errors in the first instance. *Id.* at 539–42. He suggests that these reforms may include use of properly validated risk assessments and application of the *Daubert* standard to sentencing. *Id.*

³³⁰ The listed examples encompass a subset of proposals made to improve accuracy and reliability of sentencing decisions; they are meant to provide a survey of recommendations rather than a comprehensive listing of all such recommendations made to date. Others include, for example: Prescott & Starr, *supra* note 269, at 355 (suggesting, *inter alia*, bifurcation, application of certain rules of evidence, and a number of proposals aimed at improving the factfinding abilities of juries); Sara Sun Beale, *Procedural Issues Raised by Guidelines Sentencing: The Constitutional Significance of the "Elements of the Sentence,"* 35 WM. & MARY L. REV. 147, 159–61 (1993) (suggesting that procedural protections for defendants at sentencing should not be less than those afforded parties in civil suits).

³³¹ See, e.g., LoMonaco, *supra* note 267; Stephanie C. Slatkin, Note, *The Standard of Proof at Sentencing Hearings Under the Federal Sentencing Guidelines: Why the Preponderance of the Evidence Standard Is Constitutionally Inadequate*, 1997 U. ILL. L. REV. 583; see also Becker & Orenstein, *supra* note 300, at 891.

³³² See generally Sanders, *supra* note 84, at 108 (“This Article advocates reexamining the theory that confrontation rights only apply at trial.”).

³³³ *Id.* at 127–28.

³³⁴ *Id.*

³³⁵ *Id.* at 168–69.

sentences based on junk science without hindering a defendant's ability to present mitigating evidence.

Becker and Orenstein's proposals are designed to address a broad concern about sentencing accuracy. As a result, neither is well tailored to manage the specific issue of reliance on unsound scientific evidence at sentencing. Their proposed extension of certain or specialized rules of evidence to sentencings does not include a discussion of extending Rule 702 to sentencing nor does it home in on how to deal with STS evidence at sentencing. Becker and Orenstein's alternative proposal—to apply a heightened reliability standard to evidence that might significantly impact a sentence—might lead to disparate punishment decisions for similarly situated defendants, in part because the recommended standard is not precisely defined. Judges across, and even within, jurisdictions may hold differing views on what evidence might have a “significant impact” on a sentence. As a result, the same type of evidence may be treated as requiring a heightened reliability standard in some cases and not in others. An advantage of the recommendation made herein, in contrast, is that the rule is already well-defined, and there is significant attendant case law to guide its application. Moreover, Becker and Orenstein's proposal pre-dated *Daubert*, so they cannot be faulted for not explaining how their “strong” reliability standard would correspond to the current version of Rule 702, its state analogs, or *Daubert*. Nevertheless, their proposal is silent on how a “strong reliability standard” would treat STS evidence of questionable validity.

Garrett's proposal to create a clear and consistent remedy for all sentencing errors could improve overall sentencing accuracy. But it is a back-end solution that focuses on remedying sentencing errors after the fact; it is not designed to prevent them from occurring in the first place and is certainly not specifically tailored to prevent errors arising from reliance on problematic STS evidence. Under Garrett's proposal, as with any appellate remedy, not all defendants would avail themselves of the proposed remedy and even among those who would, not all would prevail. In any event, many—if not most—of those who would seek to avail themselves of the remedy would have to wait in prison until their case was resolved. Ultimately, Garrett's proposal—as he himself recognizes by suggesting additional reforms³³⁶—does not obviate the need for a front-end approach to improving sentencing accuracy. His focus is on the wider problem of

³³⁶ See Garrett, *supra* note 298, at 539–42.

sentencing error generally, not the narrow problem of errors resulting from improper reliance on unreliable scientific evidence addressed here.

At least since *Williams*, sentencing policy has prioritized the need to consider a wide range of information about a defendant's character and background.³³⁷ Despite its analytical flaws, *Williams* continues to be cited with approval.³³⁸ This is one possible reason proposals like those described above have not been implemented. Adoption of rules of evidence at sentencing would also limit a judge's ability to rely on a wide range of other evidence, like character letters and victim impact statements containing hearsay. Extension of rules of evidence to sentencing, therefore, would result in direct tension with the sentencing goals set out in *Williams*.

The fundamental tenet of *Williams*, that judges must be able to collect and consider a broad range of information about a defendant in order to fashion an appropriate sentence,³³⁹ is not entirely without merit. The problem with *Williams* is primarily that the holding goes too far by failing to consider that, in order to avoid sentencing errors or unjustly harsh sentences, information must also be accurate.³⁴⁰ Proposals that seek to apply too many trial level evidentiary protections to sentencings are overbroad and would filter out too much information, which is what the *Williams* holding reasonably sought to prevent. For example, wholesale application of the rules of evidence might help prevent judges from relying on hearsay evidence contained in pseudo-scientific reports or on unfounded rumors, but it would also prevent them from considering evidence that is widely acknowledged to be appropriate for consideration at a sentencing hearing. If Federal Rule of Evidence 404, which restricts the introduction at trials of character evidence or evidence of a defendant's prior crimes or bad acts to demonstrate criminal propensity,³⁴¹ were held to apply at sentencing, a judge would be precluded from hearing character and character history that is considered routinely to justify an increase or decrease in sentence length.³⁴² Thus, broad extension of rules of evidence at sentencing may deprive sentencers of information that has limited rele-

³³⁷ See *supra* Section I.C.

³³⁸ See *supra* note 253 and accompanying text.

³³⁹ See *supra* note 233 and accompanying text.

³⁴⁰ See *supra* notes 237, 239 and accompanying text.

³⁴¹ See FED. R. EVID. 404.

³⁴² In fact, character and criminal history are factors that the sentencer is usually encouraged to consider. See, e.g., U.S. SENTENCING GUIDELINES MANUAL § 1B1.4 (U.S. SENTENCING COMM'N 2018).

vance at trial but that has been recognized as important to punishment decisions.

While the proponents of the recommendations above frequently frame their suggestions as ways to improve sentencing accuracy, inherent in those proposals is a recognition that a goal of sentencing reform is to minimize the harm to criminal defendants at sentencing.³⁴³ Still, many of the proposals would hurt defendants as they purport to make the hearings fairer, restricting defendants' ability to rely on character letters or other hearsay evidence at sentencing.³⁴⁴ In effect, such proposals would create the precise scenario this Article seeks to avoid: sentencing whack-a-mole, remedying one harm while allowing a new one to surface. On the other hand, by focusing on a very specific cause of unfair sentences—unreliable STS evidence—the proposed rule extension made in this Article sidesteps the overbreadth problem suffered by other recommendations.

Some of the proposals discussed above seek to balance the need for sentencing accuracy with the desire to minimize harm to defendants. These proposals focus on recognizing rights traditionally enjoyed by criminal defendants only at trial during the sentencing phase as well. For instance, increasing the burden of proof or recognizing confrontation rights at sentencing would increase the reliability of information considered at sentencing without placing a new burden on defendants.

Yet, these proposals are limited in their ability to minimize the role of faulty STS evidence in contributing to unjust sentencing outcomes. Raising the standard of proof by which the government must establish sentencing facts would certainly increase the quality of evidence relied upon at sentencing in many cases.³⁴⁵ But, just as with Becker and Orenstein's proposal for a "strong reliability standard," courts might not apply the standard in the same way.³⁴⁶ Perhaps more importantly, it does not actually screen out unreliable evidence. At

³⁴³ See Becker & Orenstein, *supra* note 300, at 888–89; Sanders, *supra* note 84, at 168. Other proponents acknowledge this goal more explicitly. Garrett, *supra*, note 298, at 543 ("Sentencing errors are inevitable, and although claims of sentencing error can raise complex issues, the practical importance of addressing their merits is simple: convicts should not serve added time based on errors."); Young, *supra* note 79, at 302 ("Two key harms to be avoided in fact-finding at guidelines sentencing are inaccuracy in fact-finding and having defendants unfairly bear the burden of errors in fact-finding.").

³⁴⁴ Of course, the same would be true with respect to prosecutors' reliance on victim impact statements or presentence investigation reports.

³⁴⁵ See Young, *supra* note 79, at 358–59.

³⁴⁶ See *id.* at 355–56 (providing examples of how one standard of proof may be applied in multiple ways).

trial, even with an elevated burden of proof, Rule 702 requires that judges perform their gatekeeping function and prevent the admission of unreliable scientific evidence at the outset, lest a factfinder over-rely on evidence labeled as “scientific.”³⁴⁷ Such evidence needs to be screened before admission. The same is true at sentencing; increasing the burden of proof would be an important step towards improving sentencing outcomes, but it would not serve as a substitute for a mechanism to screen out unreliable or invalid STS evidence.

The same applies for the right to confront witnesses. Recognizing the right to confront at sentencing would certainly aid in improving the reliability and veracity of evidence considered at sentencing³⁴⁸ but, just as confrontation does not obviate the need for application of the substantive admissibility test laid out in Rule 702 and its state counterparts at trial, it also would not at sentencing. Indeed, *Daubert’s* holding, requiring trial courts to conduct a gatekeeping function to ensure the relevance and reliability of scientific evidence despite the availability of cross-examination, rejects this idea.³⁴⁹

The proposal advanced in this Article is narrowly tailored to address one very specific, but common, cause of unfair sentences—reliance on unreliable STS evidence—and it does so while both allowing judges to consider relevant and reliable evidence that might not be admissible at trial and protecting defendants’ rights to present mitigating evidence.

CONCLUSION

Admissibility thresholds apply at trial to keep “junk science” from resulting in wrongful convictions and other miscarriages of justice. At sentencing, where the liberty interest may be greater than at trial, the same should apply. Still, there has been insufficient discussion of the extent to which STS evidence has contributed to unjust outcomes at sentencing.

³⁴⁷ See *supra* notes 177–83 and accompanying text.

³⁴⁸ See *Crawford v. Washington*, 541 U.S. 36, 61 (2004) (“The [Confrontation] Clause thus reflects a judgment, not only about the desirability of reliable evidence . . . , but about how reliability can best be determined.”).

³⁴⁹ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 589 (1993) (“[T]he trial judge *must ensure* that any and all scientific testimony or evidence admitted is not only relevant, but reliable.” (emphasis added)); see also PAUL F. ROTHSTEIN, *Rule 702: Testimony by Expert Witnesses*, in *FEDERAL RULES OF EVIDENCE* (3d ed. 2020) (explaining that *Daubert* rejects sole reliance on the adversary system for ensuring relevance and reliability of expert evidence); *United States v. Crisp*, 324 F.3d 261, 273 (4th Cir. 2003) (Michael, J., dissenting) (“Our adversarial system has much to commend it, but it is not a general substitute for the specific *Daubert* inquiry.”).

It is in no one's interest—not the defendant's, not society's, not even the prosecutor's³⁵⁰—for sentences to be based upon unreliable evidence. The criminal justice system prioritizes protecting liberty above securing detention.³⁵¹ The solution proposed by this Article targets the problematic reliance on STS evidence of questionable validity while acknowledging and promoting this principle.

Ultimately, T.K.'s story had a relatively happy ending that illustrates how the proposal made herein is administrable and can be effective in screening out junk science at sentencing. T.K.'s judge did not have to grant a hearing on the reliability of the SAVRY results or assess the admissibility of those results under *Daubert*. But he decided to anyway, holding a hearing at which he heard evidence from an expert on the SAVRY's lack of reliability and the problematic application of the tool in T.K.'s case.³⁵² That decision was pivotal. The peek behind the curtain showed T.K.'s judge that, despite the aura of trustworthiness STS evidence carries with it, not all such evidence is reliable. He found that the risk evaluator's conclusion was not supported by her data, and he decided not to rely on the assessment at sentencing.³⁵³

T.K.'s challenge was successful because the stars aligned for him in a way that is not likely to happen again; he is one of very few criminal defendants to have been represented by well-resourced defenders who were willing to litigate his claim despite knowing the law would not support it. On top of that, he was fortunate to have been assigned a judge who was willing to analyze whether the purportedly scientific evidence he was presented with was actually reliable. But T.K.'s case does not have to be a one-off. As described above, some commenters may argue against the proposal advanced here by noting that Rule 702 has proven ineffective in screening out unreliable STS evidence at the trial stage. But tell that to T.K.: T.K.'s judge essentially applied the

³⁵⁰ See, e.g., *DaCosta v. City of New York*, 296 F. Supp. 3d 569, 600 (E.D.N.Y. 2017), (“It is uncontroversial that a lawyer representing the government in the criminal context has a heightened ethical obligation that extends beyond just representing the narrow interests of her most direct client; she also must endeavor to ‘do justice.’”).

³⁵¹ See, e.g., Hon. J. Harvie Wilkinson III, *The Presumption of Civil Innocence*, 104 VA. L. REV. 589, 589 (2018) (“The presumption of innocence represents a political and moral consensus that criminal defendants should not be subject to punishment until adjudicated guilty under a strict standard of proof.”).

³⁵² See Order, *supra* note 3, at 4.

³⁵³ See *id.* at 8–9.

model advanced here; and, in doing so, changed T.K.'s future. Though T.K.'s case is only a single data point, it demonstrates that the proposal advanced here is administrable and, if adopted, *can* result in fewer miscarriages of justice.

If not for the judge's highly unusual decision to review the reliability of the report submitted against T.K., junk science would have cost T.K. his liberty and, likely, his prospects for a successful future. Few, if any, defendants are afforded the opportunity that T.K. was and are, therefore, at the mercy of unreliable STS evidence at sentencing. This Article has advanced one proposal for changing that.