

NOTES

Penile Polygraphy: The Admissibility of Penile-Plethysmograph Results at Sentencing in Tennessee

State judges in Tennessee currently consider the results of penile plethysmograph (“PPG”) evaluations when sentencing convicted sex offenders. These highly intrusive physical tests purport to identify whether an offender’s arousal is considered “deviant” by measuring the change in penis size after viewing various stimuli. Because the results are usually buried in psychosexual evaluations that are part of general presentence assessments of recidivism risk, PPG evaluations suffer from a lack of standardization and little attention under the rules of evidence. Interestingly, PPG testing is similar to polygraphy in a number of ways, although studies have shown that PPG results are more reliable than polygraph tests in determining whether a subject was truthful in reporting. For that reason, and the heightened importance of alternative sentencing decisions that prevent recidivism among individuals who cannot control their deviant sexual arousal, PPG results should be considered by judges only in limited circumstances. This Note provides a new rule of evidence modeled after New Mexico’s polygraph-admissibility rule, which provides practical standards to avoid unreliable results, consent requirements to ensure voluntariness, and opportunities to retake poorly conducted evaluations if good cause is shown. The proposed rule strikes a balance between society’s interest in safeguarding citizens from potentially dangerous sex offenders and the offender’s interests in protections from unwarranted government intrusion.

INTRODUCTION.....	354
I. RISK ASSESSMENTS, PSYCHOSEXUAL EVALUATIONS, AND PPG TESTING AT SENTENCING	358
A. <i>Methods and Science of PPG Testing</i>	359
B. <i>Risk Assessments at Sentencing</i>	363

C.	<i>Psychosexual Evaluations and PPG Results as Factors in Risk Assessments</i>	365
II.	THE ADMISSION OF PPG RESULTS AS EVIDENCE AT SENTENCING HEARINGS	368
A.	<i>Evidentiary Limitations to PPG Results Introduced at Sentencing</i>	369
1.	Rules Applicable to Scientific Evidence.....	370
2.	Admission of Polygraph Evidence	372
3.	Admission of PPG Evidence	375
B.	<i>The Ethical and Practical Limitations of PPG Testing</i>	381
III.	A RECOMMENDED RULE FOR ADMISSION OF PPG RESULTS AT SENTENCING	384
	CONCLUSION	388

INTRODUCTION

During the late 1950s, the Soviet Czechoslovakian government subjected countless homosexual men to an astonishingly intrusive physical examination to discover whether the men's professed sexual preferences were what they claimed.¹ Using a device developed by Czech researcher Kurt Freund,² officials placed a cylinder around each man's penis and forced him to view photographs of nude men, women, and children. Each man's erectile response was measured by the air displaced from the cylinder, thus exposing which pictures sexually aroused him. The goal was to discern the "true homosexuals" from those who were faking their homosexuality only to avoid service in the Czechoslovak People's Army.³ This device, known as the penile plethysmograph ("PPG"), has been continuously used since the 1950s in sexual-behavior research and has evolved into a tool for determining deviant sexual arousal (e.g., pedophilia and sadism).⁴ In Tennessee as well as many other states, PPG testing is used extensively in the

1. See Nathan Ha, *Detecting and Teaching Desire: Phallometry, Freund, and Behaviorist Sexology*, 30 OSIRIS 205, 206 (2015) (chronicling the development of penile plethysmography).

2. See K. Freund et al., *A Simple Transducer for Mechanical Plethysmography of the Male Genital*, 8 J. EXPERIMENTAL ANALYSIS BEHAV. 169, 169–70 (1965) (illustrating the hardware used as components of the plethysmograph). Freund and his coauthors' research was sponsored by the Psychiatric Research Institute in Prague, Czechoslovakia. *Id.* at 169.

3. See Ha, *supra* note 1, at 206.

4. See *id.*; see also Karen Freeman, *Kurt Freund Dies at 82; Studied Deviant Sexual Arousal*, N.Y. TIMES (Oct. 27, 1996), <http://www.nytimes.com/1996/10/27/us/kurt-freund-dies-at-82-studied-deviant-sexual-arousal.html> [<https://perma.cc/5U9S-DAEG>] (explaining the impact of Dr. Freund's research on behavioral psychiatry).

assessment and treatment of any individual convicted of a sex offense, usually as part of a clinically performed psychosexual evaluation to determine the individual's sexual development, history, interests, and recidivism risk.⁵

Psychosexual evaluations are regularly included as factors in sex-offender risk assessments. Clinicians often use these assessments to provide individualized guidance to judges, who use risk assessments at different stages of the criminal justice system, such as in pretrial detention, sentencing, and parole decisions.⁶ In Tennessee, it is statutorily mandated that sentencing judges receive the results of these psychosexual evaluations and consider them when sentencing convicted sex offenders.⁷ One optional portion of the evaluation is a modern version of the PPG test first used in Czechoslovakia, which is designed to accurately represent an offender's deviant sexual arousal and thus level of risk.⁸ Its use is attributable to the Tennessee Sex Offender Treatment Board ("Board"), which endorses psychosexual evaluations that include PPG testing.⁹ The state legislature has given the Board oversight responsibility for the assessment and treatment of convicted sex offenders.¹⁰ Due to the statutory requirements for risk assessments and the Board's support of PPG testing, the legal environment surrounding the use of PPG in Tennessee's criminal justice system provides ample opportunity for reform.

PPG testing's high level of intrusiveness necessitates an equally high level of scrutiny before the results of such tests can influence the

5. See *infra* Section I.C. See generally *Clinical Assessments*, CTR. FOR SEX OFFENDER MGMT., http://www.csom.org/pubs/cap/2/2_4.htm (last visited Oct. 22, 2018) [<https://perma.cc/RS5P-QFR6>] (describing the primary forms of clinical assessment of sex offenders used throughout the United States).

6. See *infra* Section I.B ("The Center for Sex Offender Management . . . gives guidance to states on varying types of assessments and their applicability at different stages of the criminal process.").

7. TENN. CODE ANN. § 39-13-705 (2018) (directing convicted sex offenders to submit to an evaluation for treatment and risk potential).

8. See 2 DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 10:42 (2017–2018 ed.) ("Phallometry involves . . . monitoring changes in penis size, while stimuli are presented to the subject in a controlled fashion. In most studies, variations in the characteristics of the stimuli are used to test theories about how the sexual interests of sexual aggressors differ from those of normal men . . ."); Jason R. Odeshoo, *Of Penology and Perversity: The Use of Penile Plethysmography on Convicted Child Sex Offenders*, 14 TEMP. POL. & C.R. L. REV. 1, 6–9 (2004) (describing the history and practice of penile plethysmography—the most common phallometric evaluation).

9. See Tennessee Sex Offender Treatment Board, *Policy No. 2: Standards and Guidelines for Treatment of Adult Male Sex Offenders & Professional Code of Ethics*, TENN. DEPT' CORRECTION, <https://www.tn.gov/content/dam/tn/correction/documents/TSOTBPolicy2.pdf> (last visited Dec. 17, 2018) [<https://perma.cc/42SD-6L9F>] (explaining the Board's endorsement of treatment and assessment guidelines that include PPG testing).

10. See TENN. CODE ANN. § 39-13-704.

decision whether an individual is incarcerated or sentenced via Tennessee's alternative sentencing scheme.¹¹ The ethical and practical limitations of PPG testing should be considered, in addition to both the rules of evidence and the utility of the results in determining recidivism risk. Particularly, the purpose for which PPG results are admitted has significant influence on its admissibility under the rules of evidence. Unfortunately, judicial scrutiny of the admission of PPG results at sentencing is lacking in Tennessee. The sole judicial opinion to discuss the admissibility of PPG testing as a component of a sentencing risk assessment, *State v. Edwards*, was in a case that was ultimately dismissed for insufficient evidence.¹² Thus, the analysis of the presentencing risk assessment was immaterial to the decision and considered nonbinding dicta. Yet the inquiry by the Tennessee Court of Criminal Appeals is valuable because it concluded that the test was not intended to be an evaluation tool and thus failed the expert evidence threshold requirement of reliability contained in Tennessee Rule of Evidence 702.¹³ The court further analogized to polygraph evidence, which is considered inherently unreliable, and held that a psychosexual evaluation is inadmissible when it relies on either PPG or polygraph results.¹⁴

While not explored in much depth by the court in *Edwards*, PPG results serve the same evidentiary goal as polygraph results—establishing the accuracy of the offender's self-report on his own arousal or, more generally, the creditability of the offender.¹⁵ Furthermore, the science and methodology behind both procedures are incredibly similar. The questionable science and methodology of polygraph testing has justified a complete prohibition of the use of such evidence in Tennessee and, at the very least, strict requirements for admission in a few states.¹⁶ Because of this resemblance and the fact that polygraph

11. Alternative sentences that are authorized by the Tennessee legislature include probation, suspended sentencing, community programs, restitution, required rehabilitative treatment, and any combination of traditional sentences. *See id.* § 40-35-104(c).

12. No. E2010-01731-CCA-R3-CD, 2012 WL 1799025, at *1, *20 (Tenn. Crim. App. May 18, 2012), *overruled on other grounds* by *State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015).

13. *See id.* at *20–23 (citing Tennessee trial court cases that interpret TENN. R. EVID. 702).

14. *Id.* at *22.

15. Compare Michael J. Ligons, Comment, *Polygraph Evidence: Where Are We Now?*, 65 MO. L. REV. 209, 209–10 (2000) (explaining the determination of truthfulness by observing physiological reactions during questioning), with ASS'N FOR THE TREATMENT OF SEXUAL ABUSERS, PRACTICE GUIDELINES FOR THE ASSESSMENT, TREATMENT, AND MANAGEMENT OF MALE ADULT SEXUAL ABUSERS 26 (2014) [hereinafter ATSA PRACTICE GUIDELINES] (observing that PPG testing is recognized as useful for exploring the reliability of the subject's self-report). Permission to cite the ATSA Practice Guidelines was granted by the executive director of the ATSA on February 23, 2018.

16. *See State v. Campbell*, 904 S.W.2d 608, 614–15 (Tenn. Crim. App. 1995); Adam B. Shniderman, *You Can't Handle the Truth: Lies, Damn Lies, and the Exclusion of Polygraph*

evidence is subject to a per se exclusionary rule due to the unreliability of the results, some might argue that PPG results should likewise be subject to a per se exclusion.

This Note addresses the admissibility of PPG results within sex-offender risk assessments in presentencing reports that affect sentencing decisions in courts throughout Tennessee. The widespread use of comparable psychosexual evaluations at sentencing in other, similar judicial systems presents an opportunity to apply the analysis in this Note to other states as well.¹⁷ Part I explains the usefulness of PPG results in determining an offender's recidivism risk and protecting the public from dangerous offenders, as well as the threat to offenders' rights and liberties. To accomplish this, this Note explains the position of PPG testing among the risk-assessment instruments used at the sentencing phase, the governmental bodies that oversee the procedures, the methods and science behind PPG testing, and some of the constitutional issues that have been raised. Part II introduces the various evidentiary rules that govern presentencing reports in Tennessee and explains the application of these rules to both PPG testing and polygraph evidence; it ultimately asserts that the two are undeniably similar.

Part III proposes a new rule to govern the admissibility of PPG results at sentencing in Tennessee. While a per se exclusionary rule akin to the rule regulating polygraph evidence might be appealing, the value of a focused assessment of sexual arousal to protect the public from sexual predators justifies using this evidence. PPG testing, however, must include protections against the intrusiveness on subjects. Therefore, drawing from a unique rule of evidence used in New Mexico to admit polygraph evidence, this Note presents the groundwork for a rule designed to allow the admission of PPG results if certain conditions are met. These conditions would govern the qualifications of the clinicians administering PPG tests and establish minimum testing requirements to ensure reliability. Most importantly, the rule would create procedural protections that include an opportunity to challenge the use of the results at the sentencing hearing and potentially force a second PPG test if the first is not administered in a reliable fashion. Finally, this Note concludes with practical recommendations to the

Evidence, 22 ALB. L.J. SCI. & TECH. 433, 441–43 (2012) (“Twenty-nine states bar the admission of polygraph evidence under any circumstance (per se). Currently, fifteen states admit polygraph results at trial if both the prosecution and defense stipulate to its use prior to the administration of the test. Only New Mexico allows for the routine admission of polygraph evidence.” (footnotes omitted)).

17. See, e.g., COLO. REV. STAT. § 16-11-102(b) (2018); GA. CODE ANN. § 17-10-6.2(d) (2018); IDAHO CODE § 18-8316 (2018); MONT. CODE ANN. § 46-18-111 (2017); NEV. REV. STAT. § 176.139 (2018); VT. STAT. ANN. tit. 28, § 204a (2018); WASH. REV. CODE § 9.94A.670(3) (2018).

Board that supplement the proposed evidentiary rule; the recommendations emphasize the need for enhanced privacy protections due to the inherent intrusiveness of PPG.

I. RISK ASSESSMENTS, PSYCHOSEXUAL EVALUATIONS, AND PPG TESTING AT SENTENCING

A risk assessment is a relatively recent tool that allows for empirically based judicial decisionmaking.¹⁸ A central feature of these assessments is their use of software guided by complex algorithms that predict the likelihood that an individual will reoffend.¹⁹ Risk assessments are used for a broad range of decisions where recidivism is a major consideration, such as pretrial detention, sentencing, and probation.²⁰ They are already being developed to provide guidance to police on whether to focus a criminal investigation on a particular individual.²¹ It seems entirely likely that risk assessments will become the norm throughout the United States and extend to other phases of the criminal justice system, especially in the context of what many consider among the most heinous crimes: sex offenses.

This Part discusses the general use of assessment instruments that measure a sex offender's recidivism risk at sentencing hearings in Tennessee courts. The implementation of evidence-based sentencing in Tennessee has largely been statutory, shaped by the Criminal Sentencing Reform Acts of 1989 and 2005.²² Thus, this Part also lays out the current statutory scheme and describes how risk assessments and psychosexual evaluations are conducted in practice. PPG results are considered as an element of a risk assessment in conjunction with

18. J.C. Oleson, *Risk in Sentencing: Constitutionally Suspect Variables and Evidence-Based Sentencing*, 64 SMU L. REV. 1329, 1341 (2011) (explaining that, after 2005, federal judges were authorized to draw upon empirical data and impose evidence-based sentences).

19. Danielle Kehl et al., *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, DIGITAL ACCESS TO SCHOLARSHIP HARV. 2 (July 2017), https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf [<https://perma.cc/5H52-AG29>] (describing software used to assess risk as “powered by sophisticated and sometimes proprietary algorithms”).

20. See Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 805 (2014) (“[T]he majority of states now . . . direct parole boards to consider demographic and socioeconomic factors.”).

21. See Jessica Saunders et al., *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. EXPERIMENTAL CRIMINOLOGY 347, 349 (2016) (describing Chicago's Strategic Subjects List, which identifies individuals at higher risk of being involved in gun violence).

22. The provisions governing sentencing procedures that have been affected by these reforms are located in title 40, chapter 35 of the Tennessee Code. See *infra* Section I.A (“Every sentencing judge uses a risk assessment that includes a psychosexual evaluation when making their decision . . .”).

other factors, such as self-reporting by the offender and details of any criminal history. Such information will affect the evaluating clinician's estimation of the likelihood of recidivism. Prior to examining the use of psychosexual evaluations within risk assessments in Tennessee, it is important to understand how PPG testing is conducted and the scientific basis for its utility.

A. Methods and Science of PPG Testing

PPG is a phallometric test—a test that measures changes in penile size—that determines a subject's erectile response to various stimuli.²³ The methods used in PPG testing vary throughout Tennessee; the most common tests use either a volumetric or circumferential device.²⁴ Volumetric devices completely enclose the offender's penis while leaving air space between the device and the skin.²⁵ After the offender is given various visual or audio stimuli as a control, to measure the air volume surrounding the penis while flaccid, more deviant materials (e.g., photographic depictions of children or audio recordings of violent sexual encounters) are presented to evaluate the change in air volume and thus level of arousal.²⁶ Circumferential devices use a rubber sensor that wraps around the shaft of the penis and measures the size change between flaccid and erect conditions through electronic recordings.²⁷ According to researchers, the volumetric measure appears to be more sensitive than circumferential measurements, thus generating more precise results.²⁸

In both types of tests, the change between the size of the penis while flaccid and the size while erect is measured to determine arousal level.²⁹ Deviant arousal exists when the subject shows significant erectile response after viewing stimuli that depict children or violent

23. See Hannah L. Merdian & David T. Jones, *Phallometric Assessment of Sexual Arousal*, in INTERNATIONAL PERSPECTIVES ON THE ASSESSMENT AND TREATMENT OF SEXUAL OFFENDERS: THEORY, PRACTICE, AND RESEARCH 141, 141–42 (Douglas P. Boer et al. eds., 2011) (reciting the clinical process of plethysmography).

24. See *id.* at 142 (“Both [volumetric and circumferential devices] are commonly used in correctional settings.”).

25. See Odeshoo, *supra* note 8, at 6 (briefly explaining volumetric PPG testing).

26. See *id.* at 6–9 (explaining how arousal is measured).

27. See *id.* (briefly explaining circumferential PPG testing).

28. William L. Marshall, *Clinical and Research Limitations in the Use of Phallometric Testing with Sexual Offenders*, 1 SEXUAL OFFENDER TREATMENT (2006), <http://www.sexual-offender-treatment.org/marshall.html> [<https://perma.cc/TZD5-M6CP>] (“The volumetric measure . . . appears to be the more sensitive of the two devices . . .”).

29. See Merdian & Jones, *supra* note 23, at 142 (explaining how volumetric devices measure changes in both length and diameter of the penis, while circumferential tests measure changes in diameter).

sexual encounters.³⁰ The specific type of stimulus that produces an erectile response allows the assessment to clarify which deviant sexual preference is present. These deviant preferences include those of pedophilia (sexual arousal from children) and sexual sadism (sexual arousal from humiliation, pain, or suffering of an individual).³¹

Multiple studies link erectile response to arousal levels to identify men who are more likely to reoffend than others.³² The connection between erectile response and arousal is a straightforward one. Two issues, however, threaten the relationship between arousal and recidivism. The first is the individualized nature of arousal and the difficulty of attributing the arousal to the deviant aspect of the specific stimulus, as opposed to some other nondeviant feature of the stimulus.³³ For example, a subject could be aroused by the test administrator's action rather than the stimulus itself. The lack of standardization across testing centers contributes to this concern.³⁴ Even so, there is clearly a connection when a subject responds to children but not to adults.³⁵ The second issue is the possibility of offenders faking arousal or disinterest.³⁶ As Professor William Marshall pointed out, "Several studies have shown that normal subjects can significantly inhibit their

30. See *id.* ("One attaches a device to the penis of a subject, and measures what happens when the person is exposed to a variety of possibly arousing stimuli, either visual or auditory.").

31. See Dominique Bourget & John M.W. Bradford, *Evidential Basis for the Assessment and Treatment of Sex Offenders*, 8 BRIEF TREATMENT & CRISIS INTERVENTION 130, 131 (2008) (referencing rape proneness, pedophilia, and sexual sadism as examples of deviant sexual preferences).

32. See, e.g., *id.* at 132 ("Results of studies indicate that PPG testing discriminated child molesters from other sex offenders and nonoffenders; nonfamilial child molesters from incest offenders; and homicidal child molesters from nonhomicidal child molesters and nonoffenders."); Grant T. Harris et al., *Explaining the Erectile Responses of Rapists to Rape Stories: The Contributions of Sexual Activity, Non-consent, and Violence with Injury*, 41 ARCHIVES SEXUAL BEHAV. 221, 226 (2012) (finding that sexual interest in lack of consent, as studied by PPG, was a distinguishing factor between men who have committed sexual assault and those who have not); D.R. Laws et al., *Classification of Child Molesters by Plethysmographic Assessment of Sexual Arousal and a Self-Report Measure of Sexual Preference*, 15 J. INTERPERSONAL VIOLENCE 1297, 1298 (2000) ("The most well-established method for assessing sexual preferences is penile plethysmography . . .").

33. See Marshall, *supra* note 28 ("The fact is there have been no empirical determinations of what constitutes the appropriate content of stimuli for preference testing.").

34. See *id.* ("There have been repeated calls for the standardization of phallometric tests but to date this has not been done." (citations omitted)).

35. See Max B. Bernstein, Note, *Supervised Release, Sex-Offender Treatment Programs, and Substantive Due Process*, 85 FORDHAM L. REV. 261, 278 & n.158 (2016) (interviewing a professor of psychiatry who asserted that "[i]f a patient shows erectile responses to children, yet no erectile response to adults, that has meaning").

36. See Judith V. Becker & William D. Murphy, *What We Know and Do Not Know About Assessing and Treating Sex Offenders*, 4 PSYCHOL. PUB. POL'Y & L. 116, 122–23 (1998) (identifying the faking of erectile responses as a threat to the validity of PPG testing); Merdian & Jones, *supra* note 23, at 160 ("[I]t remains questionable whether the use of a [PPG] is justified when that test is not statistically validated and where the theoretical basis of the test is unclear.").

arousal by using mental activities to distract themselves, despite a clear indication that they were attending to the stimuli.”³⁷

The connection between levels of deviant arousal and sexual-recidivism risk depends entirely on the quality of the underlying data of the arousal. PPG testing was created because of the need for an objective measurement of sexual interest.³⁸ Without an objective measure, the data used in determining recidivism risk would come solely from the offender’s self-report, which the offender is required to disclose to the evaluators.³⁹ Multiple metastudies have emphasized this point in concluding that deviant sexual interest is “among the strongest predictors of sexual recidivism.”⁴⁰ Indeed, one analysis of sixty-one studies on sex-offender recidivism recognized that “[s]exual interest in children as measured by phallometric assessment was the single strongest predictor found in the meta-analysis.”⁴¹ However, the test is not without criticism—some scholars have questioned the statistical validity of the connections to recidivism on the basis of the unreliable underlying data (discussed above with respect to faking).⁴² On a theoretical level, the ability to choose whether to act on any deviant arousal is the true intervening cause of sexual recidivism, not the deviant arousal alone.⁴³

Furthermore, there are various practical problems with PPG. While a major problem with PPG testing is the lack of standardization across testing centers, there are positive trends toward more standardized practices across jurisdictions.⁴⁴ With respect to the

37. Marshall, *supra* note 28.

38. See Freund et al., *supra* note 2, at 169 (“The availability of an objective method to determine the objects of sexual arousal in the male is of considerable importance for experimental research on sexual behavior.”).

39. See Wesley B. Maram, *Psychophysiological Assessment of Sexual Offenders: A Practitioner’s Perspective*, in *SEXUAL OFFENDING: PREDISPOSING ANTECEDENTS, ASSESSMENTS AND MANAGEMENT* 331, 332 (Amy Phenix & Harry M. Hoberman eds., 2016) (outlining the procedural steps in the administration of plethysmography).

40. *Id.*

41. R. Karl Hanson & Monique T. Bussière, *Predicting Relapse: A Meta-analysis of Sexual Offender Recidivism Studies*, 66 *J. CONSULTING & CLINICAL PSYCHOL.* 348, 351 (1998).

42. See Walter T. Simon & Peter G.W. Schouten, *The Plethysmograph Reconsidered: Comments on Barker and Howell*, 21 *BULL. AM. ACAD. PSYCHIATRY & L.* 505, 510 (1993) (“The vulnerability of the plethysmograph to voluntary control has been widely documented and is a major concern in the use of the test with offenders.”).

43. See Bernstein, *supra* note 35, at 277 (“[A] man may be aroused by sexually deviant stimuli, but engage in exclusively nondeviant activity because he is aware of social and penal sanctions that come with acting on his deviant arousal.”).

44. Lisa Murphy et al., *Standardization of Penile Plethysmography Testing in Assessment of Problematic Sexual Interests*, 12 *J. SEXUAL MED.* 1853, 1857–59 (2015) (reviewing the current state of the science and highlighting research efforts to reduce issues arising from lack of standardization); Bernstein, *supra* note 35, at 273–74 (describing the lack of standardization across PPG testing centers).

administration of the test, the level of privacy given to the offender during the test varies among testing facilities, with some offenders being completely unobserved (aside from viewing the results from a remote location) and others being observed through a one-way mirror or videotaped.⁴⁵ For these reasons, many suggest the use of PPG is best reserved for treatment of cooperating offenders that volunteer for the testing.⁴⁶ Moreover, the offender is usually required to pay for the test, which can cost up to \$1,000 per session.⁴⁷ The Tennessee Department of Correction authorizes payment of a maximum of \$175 for testing indigent offenders, which requires testing centers to make up the difference.⁴⁸ Regardless of these disparities, offenders are incentivized to undergo PPG testing because any positive results (i.e., a finding of nondeviancy) will mitigate their risk assessment and subsequent sentencing decision.⁴⁹ Every sentencing judge uses a risk assessment that includes a psychosexual evaluation when making their decisions; when conducted, PPG results are usually incorporated into those psychosexual evaluations.⁵⁰

45. See Odeshoo, *supra* note 8, at 8 (providing a general explanation of testing procedures).

46. See, e.g., *United States v. Dotson*, 324 F.3d 256, 261 (4th Cir. 2003) (holding that PPG was authorized as a condition of supervised release because it is “useful for treatment of sex offenders”). *But see* *United States v. McLaurin*, 731 F.3d 258, 260 (2d Cir. 2013) (holding that PPG as a condition was “unjustified, . . . not reasonably related to the statutory goals of sentencing, and violat[ive of the] right to substantive due process”).

47. Odeshoo, *supra* note 8, at 8.

48. See Sex Offender Treatment Board, *Sunset Public Hearing Questions*, TENN. GEN. ASSEMBLY 7–8 (2017), <http://www.capitol.tn.gov/joint/committees/gov-opps/jud/Sex%20Offender%20Treatment%20Board.pdf> [<https://perma.cc/N8NK-CC3C>] [hereinafter *Public Hearing*] (describing the statutory requirement that offenders pay and the amounts provided for indigent offenders). The Board notes that the reimbursement rates were scheduled to be raised on January 1, 2018. *Id.* at 6. While the rates rose seventy-five dollars, this amount now includes both polygraph and plethysmograph testing and therefore likely leaves less than one hundred dollars for the PPG test. See Sex Offender Treatment Board, *Reimbursement Management*, TENN. DEP’T CORRECTION 1 (July 18, 2018), https://www.tn.gov/content/dam/tn/correction/documents/TSOTB_ReimbursementManagement.pdf [<https://perma.cc/CRV2-55WH>] [hereinafter *Reimbursement Management*] (“[R]eimbursement rates for polygraphs will be reimbursed at a rate not to exceed \$175.00.”).

49. PPG testing is not explicitly mandated as part of the risk assessments used in Tennessee courts. See TENN. CODE ANN. § 39-13-705(b) (2018) (directing that a risk assessment be conducted but not identifying the specific factors to be considered). Mandating the test would implicate multiple constitutional concerns. For example, the Free Exercise Clause of the First Amendment would likely be applicable to these evaluations, which may violate an offender’s right to practice a faith that forbids masturbatory actions or viewing explicit material. See Odeshoo, *supra* note 8, at 28 (describing how PPG testing could be challenged on constitutional grounds). Also, due to the privacy interests that individuals have in their sexual fantasies, the Fourth Amendment can be implicated in certain egregious circumstances. See *id.* at 21–25 (surveying Fourth Amendment challenges to PPG testing).

50. See *infra* Section I.C (exploring how PPGs factor into risk assessments of offenders).

B. Risk Assessments at Sentencing

The first hurdle opponents of PPG testing face in attacking use of the test is the seemingly insurmountable discretion afforded to trial judges in sentencing decisions. The U.S. Supreme Court has held that judges have broad freedom in the types and sources of evidence they may consider at the sentencing phase of the adjudicative process.⁵¹ Nevertheless, the Court has, relying on Sixth Amendment protections, limited sentencing judges' discretion by holding that facts that extend the statutory maximum or mandatory minimum must be submitted to the jury.⁵² It appears that even though offenders enjoy diminished rights after conviction, there are protections that must be maintained, regardless of their status. For example, some courts have held that reliance on certain factors (e.g., race or gender) in sentencing are unconstitutional, even though scholars have identified those factors as predictive of recidivism.⁵³

Although the federal system does not currently utilize risk assessments to inform sentencing decisions,⁵⁴ the principle of broad sentencing discretion is the genesis of risk assessments' legitimacy. Broad discretion also provides the opportunity to use newly developed statistical and scientific methods within those assessments since sentencing-level evidentiary rules are subject to less scrutiny than their trial-level counterparts. Some states have embraced this principle directly in the context of risk assessments.⁵⁵ A recent decision by the Wisconsin Supreme Court upheld the use of a risk assessment at sentencing because the score produced by the assessment was only one

51. See *Williams v. New York*, 337 U.S. 241, 250–52 (1949) (maintaining that limiting the evidence that sentencing judges can consider hinders the administration of criminal justice).

52. See *Alleyne v. United States*, 570 U.S. 99, 103 (2013) (“Mandatory minimum sentences increase the penalty for a crime. It follows, then, that any fact that increases the mandatory minimum is an ‘element’ that must be submitted to the jury.”); *Apprendi v. New Jersey*, 530 U.S. 466, 490 (2000) (holding that any fact that increases the penalty for a crime beyond the prescribed statutory maximum must be submitted to the jury and proved beyond a reasonable doubt); see also Alan C. Michaels, *Trial Rights at Sentencing*, 81 N.C. L. REV. 1771, 1856–57 (2003) (comparing the rights at trial to those at sentencing).

53. See Oleson, *supra* note 18, at 1337 (identifying gender and race as factors that courts have held as unconstitutional in risk assessments).

54. See John Monahan, *Risk Assessment in Sentencing*, in 4 REFORMING CRIMINAL JUSTICE 77, 83 (Erik Luna ed., 2017) (“Risk assessment is *not* used to inform sentencing decisions in the federal system.”).

55. See, e.g., *id.* at 81–83 (discussing the unique development of risk assessments used in Virginia and Pennsylvania).

factor considered by the sentencing judge, who retained a high level of discretion.⁵⁶

The Tennessee Supreme Court has not reviewed the legitimacy of risk assessments to the extent that the high court in Wisconsin has, but it has declared a heightened standard of review when a trial court “fails to comply with . . . statutory sentencing directives.”⁵⁷ Moreover, following the nationwide trend, the Tennessee legislature has directed that a “presumption of correctness” accompanies a trial court’s sentencing decision in the absence of a failure to comply with sentencing statutes.⁵⁸ The established sentencing directives in Tennessee mandate that presentencing reports “include information identifying the defendant’s risks and needs as determined through the use of a validated assessment instrument.”⁵⁹ While there is no statutory direction on the specific validation process or factors within the assessments, recent action by the General Assembly and the Tennessee Department of Correction has shifted focus to a uniform validated assessment throughout the state, as opposed to various assessments that were validated in differing ways by trial judges.⁶⁰

Further, Tennessee requires guilty sex offenders, determined through trial or plea, to submit to a risk assessment and treatment evaluation,⁶¹ which are considered as a factor in sentencing decisions.⁶² The decision on the specific type of assessment is delegated to the Board, which is charged with overseeing sex-offender treatment and establishing a standard assessment.⁶³ The Board currently endorses the

56. See *State v. Loomis*, 881 N.W.2d 749, 753 (Wis. 2016) (concluding that the court’s consideration of a risk assessment at sentencing was proper because it “was supported by other independent factors” and “its use was not determinative”).

57. *State v. Poole*, 945 S.W.2d 93, 96 (Tenn. 1997) (holding that if the trial court fails to comply with statutory sentencing directives, there is no presumption of correctness, and the standard of review is *de novo*).

58. *State v. Pierce*, 138 S.W.3d 820, 826–27 (Tenn. 2004) (citing *State v. Ashby*, 823 S.W.2d 166, 169 (Tenn. 1991)).

59. TENN. CODE ANN. § 41-1-412(b) (2018).

60. See TENN. DEP’T OF CORR., FY 2016 ANNUAL REPORT 11 (2016), <https://www.tn.gov/content/dam/tn/correction/documents/AnnualReport02October2016.pdf> [<https://perma.cc/H8Z7-KZX9>] (describing one of the goals of the Public Safety Act of 2016 as the “use of a single validated risk and needs assessment across the criminal justice community”).

61. See TENN. CODE ANN. § 39-13-705(b) (“Those offenders found guilty at trial or who pled guilty without an agreement as to length of sentence, probation, or alternative sentencing that are to have a presentence report prepared for submission to the court shall be required to submit to the evaluation [for treatment and risk potential].”).

62. *Public Hearing*, *supra* note 48, at 6–7 (“When the risk evaluation is part of the presentence report, it is considered as a factor in sentencing decisions.”).

63. See TENN. CODE ANN. § 39-13-704 (asserting that the Sex Offender Treatment Board shall develop standard procedures for the evaluation and treatment of sex offenders); see also *Tennessee Sex Offender Treatment Board*, TENN. DEP’T CORRECTION, <https://www.tn.gov/correction/tennessee-sex-offender-treatment-board.html> (last visited Oct. 22, 2018) [<https://perma.cc/7TXX->

standards and guidelines of the Association for the Treatment of Sexual Abusers (“ATSA”), which has its own evaluations that closely follow the recommendations of the U.S. Department of Justice.⁶⁴ The Center for Sex Offender Management (“CSOM”), established by the Department of Justice, also gives guidance to states on varying types of assessments and their applicability at different stages of the criminal process.⁶⁵ The types of information recommended for incorporation in the presentencing report include trial evidence; prior criminal records; family and social history; medical and mental-health issues; educational and employment history; financial stability; and estimated recidivism risk, including findings from psychosexual evaluations.⁶⁶ However, regardless of established procedures, trial courts are not bound to adhere to those procedures when determining what evidence is admissible at a sentencing hearing.⁶⁷

C. Psychosexual Evaluations and PPG Results as Factors in Risk Assessments

While Tennessee gives the Board discretion to use psychosexual evaluations, many states have statutes that specifically require or authorize the use of risk assessments with psychosexual evaluations.⁶⁸ The ATSA Practice Guidelines for assessing sex offenders, which the Board fully endorses, reveal the recommended use of psychosexual

AWRC] (explaining that the Sex Offender Treatment Board develops and implements standardized procedures and programs for the evaluation and treatment of sex offenders).

64. See Tennessee Sex Offender Treatment Board, *supra* note 9 (“The Board endorses ATSA Practice Standards and Guidelines for the evaluation, treatment and management of adult male sexual abusers.”); see also ATSA PRACTICE GUIDELINES, *supra* note 15, at 7 (stating that the ATSA Practice Guidelines provide recommendations and guidance regarding the assessment and treatment of adult sexual abusers).

65. The CSOM is a national clearinghouse and technical-assistance center that supports state and local jurisdictions in the effective management of sex offenders. See CTR. FOR SEX OFFENDER MGMT., <http://www.csom.org> (last visited Oct. 22, 2018) [<https://perma.cc/J4FQ-MJ9J>].

66. See *Assessments Specific to Criminal and Juvenile Justice Systems*, CTR. FOR SEX OFFENDER MGMT., http://www.csom.org/pubs/cap/2/2_1.htm (last visited Oct. 22, 2018) [<https://perma.cc/C5HY-YHXJ>] (listing example types of information that should be included in the presentencing trial report).

67. See *State v. Pierce*, 138 S.W.3d 820, 825 (Tenn. 2004) (“[T]rial courts are not bound by the Board’s standardized procedures when determining what evidence is admissible at a sentencing hearing.”).

68. See, e.g., OHIO REV. CODE ANN. § 2941.148(B) (West 2018) (allowing various types of evidence to be used to determine whether “it is likely that the person will engage in the future in one or more sexually violent offenses”); VT. STAT. ANN. tit. 28, § 204a(b)(1) (2018) (stating that a presentence investigation “[s]hall include an assessment of the offender’s risk of reoffense and a determination of whether the person is a high risk offender”); WASH. REV. CODE § 9.94A.500(1) (2018) (providing that the “court shall consider the risk assessment report and presentence reports” for sex offenders).

evaluations.⁶⁹ This includes an interview with the subject, interviews with others involved in the life of the offender or incident at issue, a review of past assessments or other official documents (e.g., other criminal records), intellectual and emotional testing, an evaluation of offense-related attitudes and interests, and documentation of other general background information.⁷⁰ The stated purposes of the evaluations are to understand the nature and extent of the offender's sexually abusive behavior, explore needs for treatment, and estimate recidivism risk.⁷¹ When evaluating risk, the most heavily weighted factors are criminal history, victim-related variables, sexual deviancy, antisocial orientation, relationship difficulties, and self-regulation problems.⁷²

Both the ATSA and the CSOM explain that the polygraph, viewing-time, and PPG tests are enhancement tools to evaluate the self-reporting portion of the psychosexual evaluation as part of an "overall assessment strategy."⁷³ Subjects are required to report their sexual preferences and history, which are confirmed through the use of the enhancement tools.⁷⁴ Polygraph tests, commonly referred to as lie-detector tests, determine the likelihood of subjects' truthfulness by measuring their physiological responses while they answer questions.⁷⁵ The polygraph test is typically an indirect assessment that either supplements the other evaluations or assesses the sexual history of the offender, not a stand-alone factor.⁷⁶ In fact, mere consideration of polygraph results within a risk-assessment instrument during sentencing is impermissible in Tennessee due to the unreliability of the test.⁷⁷ Viewing-time evaluations apply a temporal analysis of offenders'

69. See ATSA PRACTICE GUIDELINES, *supra* note 15, at 18–24; see also *supra* note 9 and accompanying text.

70. ATSA PRACTICE GUIDELINES, *supra* note 15, at 18–24.

71. *Id.* at 15.

72. *Id.* at 24–25.

73. See *id.* at 27 (stating that "the results of phallometric, viewing time, and polygraph methods are not to be used as the sole criterion" for an assessment); *Physiological Assessments of Deviant Arousal, Interests, and Preferences*, CTR. FOR SEX OFFENDER MGMT., http://www.csom.org/pubs/cap/2/2_6.htm (last visited Oct. 22, 2018) [<https://perma.cc/ME7D-Z5E3>] (urging that these tests "must be considered as part of an overall assessment strategy").

74. See ATSA PRACTICE GUIDELINES, *supra* note 15, at 75–78 (explaining enhancement tools and tests used to confirm self-reported statements by sex offenders on their sexual preferences and history).

75. See Shneiderman, *supra* note 16, at 435–37 (describing the historical development of polygraphy).

76. See *id.* at 450–51 (commenting on the "indirect nature of the polygraph's efforts to detect lies").

77. *State v. Pierce*, 138 S.W.3d 820, 826 (Tenn. 2004) ("It is well-established in the jurisprudence of this State that polygraph evidence is inherently unreliable, and therefore irrelevant and inadmissible.").

reactions while they observe visual stimuli such as pictures of children.⁷⁸ Relative to other tests, viewing-time evaluations are newly developed and have less evidence available to support their reliability.⁷⁹

In contrast, PPG is considered the most empirically supported method of evaluating offenders' deviant sexual arousal and thus provides an additional level of accuracy when included in recidivism-risk assessments.⁸⁰ However, the primary concerns are the intrusiveness and reliability of PPG testing. While many jurisdictions require inmates to submit to treatment with PPG testing while incarcerated for rehabilitative purposes,⁸¹ the use of the results in risk assessments varies among jurisdictions, with the most common use being as a factor in parole decisions and the subsequent condition of release.⁸² The consideration of PPG testing at sentencing is either as a factor in the psychosexual evaluation or as a separate consideration of the offender's refusal to submit to testing.⁸³ While Tennessee requires psychosexual evaluations in cases where an offender requests alternative sentencing, such as probation or a suspended sentence,⁸⁴ the use of PPG elsewhere is legally questionable.⁸⁵

The ATSA guidelines—endorsed by the Board—utilize PPG (along with polygraph and viewing-time tests) to support the psychosexual-evaluation aspect of the presentencing risk assessment.⁸⁶ This entails subjective determinations by test administrators on how

78. See ATSA PRACTICE GUIDELINES, *supra* note 15, at 73–75 (explaining the procedures used in viewing-time evaluations).

79. See *id.* (noting the limitations of viewing-time evaluations).

80. See Anthony R. Beech et al., *Risk Assessment of Sex Offenders*, 34 PROF. PSYCHOL. 339, 344 (2003) (“PPG indices, especially those indicating a sexual preference for children, are predictive of sexual recidivism.”); Richard J. Howes, *Measurement of Risk of Sexual Violence Through Phallometric Testing*, 11 LEGAL MED. S368, S369 (2009) (“Predictions of level of risk to engage in acts of sexual violence . . . are rendered much more accurate by the inclusion of data from [PPG].”).

81. Odeshoo, *supra* note 8, at 3 (“PPG has . . . been upheld in many states as a component of prison treatment programs in which incarcerated sex offenders are increasingly required to participate.”).

82. See Bernstein, *supra* note 35, at 264 (“PPG testing has explicitly been ordered as a condition of supervised release in district courts within nearly all of the federal circuits.”).

83. See generally 6 WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE § 26.4(c) (4th ed. 2015) (detailing the jurisdictions that allow an inference of dangerousness from unwillingness to cooperate with psychosexual examinations).

84. See *State v. Reno*, No. M2016-01903-CCA-R3-CD, 2017 WL 3037538, at *16 (Tenn. Crim. App. July 18, 2017) (remanding to trial court for psychosexual evaluation and new sentencing hearing that considers the evaluation).

85. See *State v. Edwards*, No. E2010-01731-CCA-R3-CD, 2012 WL 1799025, at *23 (Tenn. Crim. App. May 18, 2012) (stating in dicta that the PPG test relied on by the psychosexual evaluator failed to meet the scientific-validity requirement for admissibility), *overruled on other grounds* by *State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015).

86. See *supra* notes 73–80 and accompanying text.

the PPG results affect the overall level of risk. Other risk-assessment instruments use somewhat more objective criteria when including the results of PPG testing in risk calculations. For example, the Sex Offender Risk Appraisal Guide (“SORAG”), which has also been used in Tennessee, establishes a positive point value for deviant responses and a negative point value for nondeviant responses.⁸⁷

The intrusiveness of the PPG testing presents a major threat to the rights and liberties of offenders subjected to the procedure. Still, despite the need for protections, the usefulness of PPG results in determining recidivism risk and protecting the public from dangerous offenders cautions against broad disqualification of PPG results at sentencing hearings. Furthermore, the limited constitutional remedies available at the sentencing phase establish less protection than the rules of evidence, which provide both protections from unreliable evidence and the opportunity to introduce potentially probative PPG results.

II. THE ADMISSION OF PPG RESULTS AS EVIDENCE AT SENTENCING HEARINGS

Every sentencing decision of a sex offender in Tennessee relies on a presentencing report that incorporates a psychosexual evaluation.⁸⁸ This evaluation provides an offender’s recidivism risk, which the sentencing judge weighs heavily—this strong reliance is further increased when elected judges, like Tennessee lower-court judges, factor in the detrimental effect that releasing a “high-risk” sex offender could have on their professional reputation among the electorate.⁸⁹ One restraint on judges’ discretion is the statutory obligation to apply the Tennessee Rules of Evidence at sentencing hearings.⁹⁰ This Part discusses the state of evidence law in Tennessee with respect to expert testimony and scientific evidence, which provides the framework for analyzing PPG evidence. Furthermore, it examines Tennessee’s distinctive treatment of polygraph results; in light of the scientific unreliability of this type of evidence, there is a *per se* exclusionary rule for all polygraph evidence. The similarities between

87. See GRANT T. HARRIS ET AL., *VIOLENT OFFENDERS: APPRAISING AND MANAGING RISK* app. B (3d ed. 2015) (explaining the SORAG scoring matrix).

88. TENN. CODE ANN. § 41-1-412(b) (2018) (“In preparing presentence reports . . . the department of correction shall include information identifying the defendant’s risks and needs . . .”).

89. TENN. CONST. art. VI, § 4, cl. 1 (“The Judges of the Circuit and Chancery Courts, and of other Inferior Courts, shall be elected by the qualified voters of the district or circuit to which they are to be assigned.”).

90. See *infra* note 124 and accompanying text.

PPG testing and polygraphy—in their methodology, science, and purposes—suggest that applying a strict evidentiary rule for PPG results is reasonable, especially given the outcome of the analysis under Tennessee Rules of Evidence 702 and 703. Lastly, this Part discusses the ethical and practical limitations of the PPG testing scheme in Tennessee as further support for the assertion that a higher level of scrutiny is warranted when admitting PPG evidence.

A. Evidentiary Limitations to PPG Results Introduced at Sentencing

The area of the law that likely has the most impact on the validity of PPG testing within sex-offender risk-assessment instruments is evidence law. Because the U.S. Supreme Court has held that some, but not all, individual rights continue to apply at the sentencing phase, there must be certain procedures and rules of evidence to protect such rights.⁹¹ The Tennessee legislature has declared that the rules of evidence apply at sentencing hearings, with the exception that *reliable* hearsay evidence is admissible if the opposing party has a fair opportunity to rebut the evidence and it does not violate any constitutional rights.⁹² This emphasis on the reliability of otherwise inadmissible evidence demonstrates the legislature's intent to protect offenders from arbitrary punishments.⁹³ Indeed, the Tennessee Supreme Court in *State v. Pierce* interpreted the relevant statutory provisions when proclaiming that polygraph evidence is irrelevant and inadmissible at sentencing hearings.⁹⁴ Going even

91. See *supra* notes 51–53 and accompanying text.

92. See TENN. CODE ANN. § 40-35-209(b) (“The rules of evidence shall apply, except that reliable hearsay, including, but not limited to, certified copies of convictions or documents, may be admitted if the opposing party is accorded a fair opportunity to rebut any hearsay evidence so admitted [and the evidence is not] secured in violation of the United States or Tennessee constitutions.”).

93. The “reliable hearsay” language is directed at the admission of victim-impact statements during sentencing hearings. See *id.* To avoid haphazardly sentencing an individual after hearing the victim’s statement, courts must ensure that the hearsay statements (which are otherwise inadmissible under the rules of evidence) are sufficiently reliable. See *State v. Moss*, 13 S.W.3d 374, 385 (Tenn. Crim. App. 1999) (“It is well settled in Tennessee that a trial court has statutory authority to admit trustworthy and probative evidence, including hearsay, for sentencing purposes.”). The principle of considering reliable evidence to ensure reasonable and consistent sentences is generally applicable to all sentencing decisions, regardless of the existence of a victim-impact statement.

94. See 138 S.W.3d 820, 826 (Tenn. 2004) (“Tennessee Rule of Evidence 402 states that ‘[e]vidence which is not relevant is not admissible.’ It is well-established in the jurisprudence of this State that polygraph evidence is inherently unreliable, and therefore irrelevant and inadmissible.” (alteration in original)).

further, the court directed trial courts to not consider any portion of a risk-assessment report that relies on polygraph results.⁹⁵

PPG testing might be distinguishable from *Pierce*'s treatment of polygraphy because PPG is a different test with different procedures and purposes, especially in light of the established polygraph jurisprudence but lack of PPG discussion.⁹⁶ However, in *State v. Edwards*, the Tennessee Court of Criminal Appeals extended the court's analysis in *Pierce* to PPG testing.⁹⁷ Unfortunately, it is the only criminal case in Tennessee to discuss the admissibility of PPG testing, and the discussion is merely dicta because the court dismissed the case for insufficient evidence before moving to its discussion of the presentencing report.⁹⁸ Regardless, the opinion's analysis of PPG evidence at sentencing sheds light on the evidentiary limitations of the evaluation. Before examining *Edwards* and the admissibility of PPG testing or polygraphy, it is helpful to understand the state of evidence law in Tennessee with respect to scientific evidence after the *Daubert v. Merrell Dow Pharmaceuticals, Inc.*⁹⁹ decision, which changed the landscape of the admissibility of expert testimony.

1. Rules Applicable to Scientific Evidence

The Tennessee Supreme Court in *McDaniel v. CSX Transportation, Inc.* defined the standard for admissibility of scientific evidence in Tennessee courts under Rules 702 and 703, regarding expert testimony.¹⁰⁰ While the newly articulated interpretation of Federal Rule of Evidence 702 in *Daubert*¹⁰¹ was informative to the analysis of scientific evidence in Tennessee (specifically the factors that

95. *See id.* (“[P]olygraph examination results, testimony on such results, or testimony regarding a defendant’s willingness or refusal to submit to a polygraph examination is not admissible during capital or non-capital sentencing hearings.”).

96. *See id.* (citing the line of cases in Tennessee that hold polygraph evidence unreliable and therefore inadmissible).

97. *See No. E2010-01731-CCA-R3-CD*, 2012 WL 1799025, at *22–24 (Tenn. Crim. App. May 18, 2012) (“A trial court is not required to consider a psychosexual evaluation that is based upon otherwise inadmissible evidence.”), *overruled on other grounds by State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015); *see also infra* notes 150–167 and accompanying text.

98. *See Edwards*, 2012 WL 1799025, at *23 (finding that no evidence had been presented on the scientific reliability of PPG testing).

99. 509 U.S. 579 (1993).

100. 955 S.W.2d 257, 263–65 (Tenn. 1997).

101. The Court in *Daubert* held that the *Frye* “general acceptance” test that had been used with scientific evidence for decades was not intended to be incorporated into the new Federal Rules of Evidence. 509 U.S. at 588. Instead of deferring solely to the acceptance of the general scientific community to determine reliability, courts are now required to play a “gatekeeping role” and make their own assessment of such criteria. *See id.* at 597 (assigning the trial judge the task of ensuring that scientific evidence “rests on a reliable foundation”).

determine reliability), it was not adopted as a matter of law. Textually, the court distinguished the federal rule from the Tennessee rule by recognizing that “Tenn. R. Evid. 702 requires that the scientific evidence ‘*substantially* assist the trier of fact,’ while its federal counterpart requires only that the evidence ‘assist the trier of fact.’”¹⁰² Thus, the “probative force of the testimony must be stronger before it is admitted in Tennessee.”¹⁰³ Additionally, the court identified that the language of Tennessee Rule 703, as opposed to the federal rule, was “designed to encourage trial courts to take a more active role in evaluating the reasonableness of the expert’s reliance upon the particular basis for his or her testimony,” even though the underlying data need not be admissible.¹⁰⁴

Relying on the *McDaniel* decision, Tennessee courts have established that expert evidence and testimony must first meet the heightened relevancy standard.¹⁰⁵ Further, courts must find the evidence to be reliable, and they may consider the following factors: (1) whether and with what methodology the scientific evidence has been tested; (2) whether the evidence has been subjected to peer review or publication; (3) whether the potential rate of error is known; (4) whether the evidence is generally accepted in the scientific community; (5) whether the expert’s research in the field has been conducted independent of litigation;¹⁰⁶ (6) the expert’s qualifications for testifying on the subject at issue; and (7) the straightforward connection between the expert’s knowledge and the basis for the opinion such that no “analytical gap” exists between the data and the opinion offered.¹⁰⁷ Finally, the underlying foundation on which the expert’s testimony relies must be reasonable and trustworthy.¹⁰⁸

102. *McDaniel*, 955 S.W.2d at 264. Compare TENN. R. EVID. 702 (“If scientific, technical, or other specialized knowledge will substantially 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 . . .”), with FED. R. EVID. 702:

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

103. *McDaniel*, 955 S.W.2d at 264.

104. *Id.* at 265 (citing Robert Banks, Jr., *Some Comparisons Between the New Tennessee Rules of Evidence and the Federal Rules of Evidence Part II*, 20 MEM. ST. U. L. REV. 499, 559 (1990)).

105. *State v. Stevens*, 78 S.W.3d 817, 832–33 (Tenn. 2002) (noting that both scientific and nonscientific evidence must meet the requirements of relevance and reliability).

106. TENN. R. EVID. 702 advisory commission’s 2001 comment; *McDaniel*, 955 S.W.2d at 265.

107. In addition to the *McDaniel* factors, the court in *Stevens* offered the last two factors as additional considerations when an expert’s reliability is challenged. 78 S.W.3d at 834–35.

108. See TENN. R. EVID. 703 (“The court shall disallow testimony in the form of an opinion or inference if the underlying facts or data indicate lack of trustworthiness.”).

2. Admission of Polygraph Evidence

Evidence law in the United States has a long history of excluding polygraph results. Indeed, *Frye v. United States* excluded evidence from a rudimentary version of polygraphy because the scientific community deemed polygraph science unacceptable, establishing the well-known *Frye* test of general acceptance for expert testimony.¹⁰⁹ To understand why courts almost always exclude polygraph evidence, one must first ask what the test's purpose is with respect to proof. Is it to establish credibility or, more generally, guilt? Those are certainly the reasons for introducing polygraph evidence at trial, and without hesitation, Tennessee courts exclude it by simply citing precedent that finds polygraph results unreliable but without engaging in much independent analysis.¹¹⁰ The genesis of this per se exclusionary rule seems to be *Marable v. State*, which asserted that the "unquestioned and unanimous weight of authority and general rule is that the results of a lie detector test are inadmissible in evidence."¹¹¹ In *Marable*, it is unclear whether the reason for exclusion was a classic application of the *Frye* principle of general acceptance or a judicial determination of reliability.

Even though there is a lack of judicial analysis of the scientific aspect of polygraphy, the current expert testimony rules set forth in *McDaniel* would likely exclude the results anyway. The justifications for exclusion of polygraph evidence are the unacceptably high or unknown error rate, the lack of appropriate validation studies, and the absence of standardization throughout the scientific community.¹¹² Polygraph exclusion is also supported by the National Academy of Sciences' report on the scientific research conducted regarding polygraphy. The report found that there is little basis for the expectation that the test could have a high level of accuracy because the "physiological responses measured . . . are not uniquely related to deception" and "uncontrolled variation[s] in test administration . . . can

109. 293 F. 1013 (D.C. Cir. 1923). The *Frye* test deferred to the scientific community to decide whether the science was acceptable. *See id.* at 1014. In contrast, the standard set forth in *Daubert* placed the decisionmaking power in the hands of the judiciary. *See* Edward K. Cheng & Albert H. Yoon, *Does Frye or Daubert Matter? A Study of Scientific Admissibility Standards*, 91 VA. L. REV. 471, 471–72 (2005) (pointing out that legal commentators have debated which institution should hold the authority to make admissibility decisions).

110. *See, e.g.*, *State v. Campbell*, 904 S.W.2d 608, 614 (Tenn. Crim. App. 1995) ("It has long been established that the results of a polygraph examination are not admissible as evidence in a criminal prosecution.").

111. 313 S.W.2d 451, 458 (Tenn. 1958). This is the first case in Tennessee to address the admissibility of polygraph evidence.

112. *See* Shniderman, *supra* note 16, at 469 (claiming that judges treat polygraphy with more hostility than with similar forensic sciences).

be expected to . . . limit the level of accuracy.”¹¹³ Yet because the individualized factors discussed in *McDaniel* allow judges to consider specific cases individually, polygraph evidence should not be summarily dismissed as unreliable under Rule 702.

The Tennessee Supreme Court has also extended the per se exclusionary rule to other contexts, such as defendants’ attempts to establish residual doubt of guilt (essentially a credibility determination based on the defendant’s claim of innocence) at sentencing.¹¹⁴ Thus, rather than adjust the analysis under Rule 702 in that context, the court extended the same per se exclusion based on the *purpose* of introducing the polygraph at trial. This outcome is at odds with the National Academy of Sciences’ understanding of the assessment of validity. In its report, it points out that the usefulness or reliability of the evidence depends on the context—“[v]alidity is not something that courts can assess in a vacuum.”¹¹⁵ Therefore, it seems the Tennessee courts are foregoing the proper analysis established by the legislature in Rule 702 and the Tennessee Supreme Court in *McDaniel* and continuing to adhere to the per se exclusionary rule, even though it was pronounced in the era before the new rules of evidence were established. To be clear, a per se exclusionary rule might very well be warranted, yet analysis under Rule 702—namely, the articulation of the unreliability of the procedure—is the ignored element that is vital to maintaining consistency with the rule.

The purpose of using polygraph results in the context of sex-offender risk assessments at sentencing is to establish the credibility of the subject.¹¹⁶ The sex-offender risk assessments endorsed by the ATSA include the use of polygraphy to clarify discrepancies, facilitate disclosure of sexual history, and explore the subject’s involvement in prohibited behavior.¹¹⁷ Before 2004, polygraph results were permitted in risk assessments for consideration by sentencing judges in Tennessee.¹¹⁸ In broad terms, judges and researchers believed

113. NAT’L RES. COUNCIL, THE POLYGRAPH AND LIE DETECTION 212–13 (2013), <https://www.nap.edu/read/10420/chapter/1> [<https://perma.cc/TE8M-7J49>] [hereinafter THE POLYGRAPH AND LIE DETECTION].

114. *See State v. Hartman*, 42 S.W.3d 44, 60 (Tenn. 2001) (holding polygraph evidence was not admissible to establish residual doubt regarding the defendant’s guilt).

115. THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 204.

116. *See ATSA PRACTICE GUIDELINES*, *supra* note 15, at 77 (explaining that the primary type of polygraph test performed under the guidelines is designed to “clarify discrepancies” between what the subject reports and the official description of the convicted offense).

117. *See id.* at 75–76 (listing the objectives to be served by postconviction polygraph testing).

118. The Tennessee Supreme Court’s decision in *State v. Pierce* ended the use of polygraph evidence in presentencing reports. *See* 138 S.W.3d 820, 826 (Tenn. 2004) (“It is well-established in the jurisprudence of this State that polygraph evidence is inherently unreliable, and therefore irrelevant and inadmissible.”).

polygraph results could corroborate the risk levels assigned by the assessment tools and could be “interpreted in conjunction with other relevant information to inform decision making.”¹¹⁹ The consideration of polygraph evidence in this context was challenged in *State v. Pierce*.¹²⁰ Again relying on the per se exclusionary rule for polygraph evidence, the Tennessee Supreme Court declared that consideration of such evidence when imposing sentences was improper.¹²¹ Going even further, the court invalidated any portion of a risk assessment that relies on polygraph results.¹²² There, however, the court had to deal with the broad discretion afforded to both the sentencing judge and the Board.

Sifting through the various statutes that govern presentencing risk assessments of sex offenders, the court recognized that while the legislature gave the Board authority to develop risk-assessment procedures and mandated that the presentencing reports be prepared in accordance with those procedures, there were no direct instructions requiring the use of polygraphy specifically within any of those instruments.¹²³ Moreover, the court distinguished the Board’s discretion to prescribe standards for the reports from the trial judge’s discretion to determine what is admissible at a sentencing hearing. Indeed, the judge must screen the evidence used at a sentencing hearing to ensure compliance with the rules of evidence, which the legislature has determined apply at sentencing hearings.¹²⁴ Thus, without discussing Rule 702 and instead relying on the per se exclusionary rule of polygraph evidence, the *Pierce* court invalidated the risk assessment used to evaluate *Pierce* because it relied on polygraph results.

However, the court did articulate that, given the “well-established” jurisprudence regarding the unreliability of polygraphy, results from such tests are irrelevant and thus excluded under Rule 402—the general rule allowing only relevant evidence to be admitted.¹²⁵ Two factors might be at play in the *Pierce* court’s decision to use Rule

119. ATSA PRACTICE GUIDELINES, *supra* note 15, at 78.

120. 138 S.W.3d at 825–26.

121. *See id.* at 826 (“[W]e conclude that the trial court erred by considering the results of *Pierce*’s polygraph examination when determining his sentence.”).

122. *Id.*

123. *See id.* at 824–25.

124. *See id.* at 825 (“[T]rial courts are not bound by the Board’s standardized procedures when determining what evidence is admissible at a sentencing hearing.”); *see also* TENN. CODE ANN. § 40-35-209(b) (2018) (describing the procedures for admission of evidence at sentencing hearings).

125. *Pierce*, 138 S.W.3d at 826; *see also* TENN. R. EVID. 402 (“All relevant evidence is admissible except as provided by the Constitution of the United States, the Constitution of Tennessee, these rules, or other rules or laws of general application in the courts of Tennessee.”).

402 as opposed to Rule 702. First, no expert testimony was offered in support of or opposition to introducing the polygraph evidence—both parties stipulated that polygraph evidence was inadmissible.¹²⁶ Therefore, the court did not conduct a full *McDaniel* analysis. Second, the court may have felt a need to ground its common law exclusionary rule in *some* codified rule of evidence, and the categorical nature of Rule 402¹²⁷ was a better fit than the balancing approach of Rule 702.¹²⁸ While intuitive, this analytical step had not been articulated previously in any detail. As it stands, the Tennessee Supreme Court has adhered to the per se exclusionary rule for polygraph evidence, even in the context of its use as a component of presentence sex-offender risk assessments.

3. Admission of PPG Evidence

The purpose for which a party attempts to introduce polygraph evidence sheds light on the admissibility of PPG testing. Evidence of PPG testing and polygraphy are introduced for strikingly similar purposes. The primary purpose of polygraphy is credibility—the objective is to determine if someone is being deceptive.¹²⁹ While the immediate purpose of PPG testing is to determine a subject’s level of deviant arousal, it is introduced into evidence for the secondary purpose of credibility—the objective is to determine if the subject was truthful in self-reporting his deviant arousal.¹³⁰ If it was proffered for its immediate purpose, it would almost always be excluded. Applying the federal *Daubert* standard, the Fourth Circuit in *United States v. Powers*¹³¹ denied a defendant’s attempt to introduce PPG results at trial to show he did not exhibit the characteristics (i.e., deviant arousal) of a fixated pedophile. The scientific literature presented to the court acknowledged the lack of accepted standards for PPG as a diagnostic

126. See Brief of Appellant at 14, *State v. Pierce*, 138 S.W.3d 820 (Tenn. 2004) (No. E2001-01734-SC-R11-CD) (asserting the per se exclusion rule for polygraph evidence); Brief of the State of Tennessee at 13, *State v. Pierce*, 138 S.W.3d 820 (Tenn. 2004) (No. E2001-01734-SC-R11-CD) (“The State does not dispute the defendant’s assertion that the polygraph results were inadmissible as evidence.”).

127. See TENN. R. EVID. 402 (“Evidence which is not relevant is not admissible.”).

128. See *supra* Section II.A.1.

129. THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 1 (discussing that polygraph testing was developed to detect deception through psychophysiological observations).

130. The first objective of PPG testing mentioned by the ATSA is “obtaining objective behavioral data about the client that may not be readily established through other assessment means.” ATSA PRACTICE GUIDELINES, *supra* note 15, at 26. The second objective is “exploring the reliability of client self-report.” *Id.*

131. 59 F.3d 1460, 1471 (4th Cir. 1995).

tool and led the court to conclude that the test is only reliable as part of treatment.¹³²

While no consensus throughout the scientific community exists two decades after the *Powers* decision, there has been an increasing understanding that the use of phallometric evaluations can produce insight into a subject's level of deviant arousal.¹³³ What is most unclear is the relationship between the stimuli provided and the erectile response.¹³⁴ Due to each individual's idiosyncratic preferences, it is difficult to determine if the deviant aspect of a particular stimulus is in fact the aspect that produces the erectile response.¹³⁵ Due to these shortcomings, some researchers have pointed out that producing a standardized test that identifies deviance may be impossible.¹³⁶ Furthermore, recall the National Academy of Sciences' reasons for finding that polygraphy lacks reliability—the bodily responses measured are not distinctively associated with lying, and variations in testing procedures limit accuracy.¹³⁷ Both of these problems are present when analyzing PPG testing. The physiological response of an erection is not *uniquely* related to deviancy because of idiosyncratic sexual preferences,¹³⁸ and variations in test administration also exist with PPG testing.¹³⁹ Indeed, the ability to fake physiological action, whether it be avoiding arousal during a PPG test or adjusting heart rate during a polygraph test, affects reliability.¹⁴⁰ Since PPG testing for deviant arousal has the same characteristics that make polygraph evaluations unreliable, the admission of PPG results must be heavily scrutinized.

132. See *id.* (“[A]lthough useful for treatment of sex offenders, it has no accepted standards in the scientific community.”). The scientific literature used by the court resembled the discussion *supra* Section I.A.

133. See *supra* notes 32–37 and accompanying text.

134. See *supra* notes 32–37 and accompanying text.

135. See Marshall, *supra* note 28 (“[I]f all sexual offenders prove to be . . . idiosyncratically aroused, then producing a standardized test that identifies deviance, may be impossible.”).

136. See *id.* (“[T]here have been no empirical determinations of what constitutes the appropriate content of stimuli for preference testing.”).

137. See THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 212–13 (“[T]he responses measured by the polygraph do not all reflect a single underlying process . . .”).

138. See Marshall, *supra* note 28 (“[T]he resultant arousal patterns may differ not because the stimuli depict normative or deviant events but simply because they depict differing sexual elements.”).

139. See THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 13 (reporting that a variety of technological responses have been developed that use physiological responses to make inferences about deceptiveness); Odeshoo, *supra* note 8, at 12–13 (“[V]ariables [in PPG testing] may differ, including the type of stimuli used, the content, duration, and interval between presentations, the types of instructions given to subjects, the type of equipment used, as well as how responses are counted.”).

140. See Becker & Murphy, *supra* note 36, at 122–23 (identifying the faking of erectile responses as a threat to the validity of PPG testing); Odeshoo, *supra* note 8, at 12 (“[I]t is possible for at least some subjects to fake PPG test results.”).

For the foregoing reasons, attempting to admit PPG results under Rules 702 and 703 for the purpose of proving deviant arousal implicates many of the *McDaniel* factors in a way that would favor exclusion. The extensive scientific research on the physiological response to stimuli and connection to recidivism suggests some general acceptance of such evidence in the scientific community, the fourth *McDaniel* factor.¹⁴¹ However, major concerns arise when considering the methodological testing, rate of error, and existence of an analytical gap between the data and the opinion provided. Any argument that methodological testing has been conducted is undermined by the lack of standardization across testing centers, because the different PPG procedures can have an impact on the results. For example, the different levels of precision between volumetric and circumferential methods of testing could produce varied results.¹⁴² What is more, the lack of standardized PPG testing presents problems in assessing PPG's rate of error. Finally, there exists an analytical gap between arousal data and opinions on likely recidivism because, as discussed earlier, one can consciously decide not to act on deviant thoughts.¹⁴³

At the sentencing phase, courts would likely be more willing to admit PPG evidence for its secondary purpose—credibility. Due to the shift in sentencing theory to evidence-based approaches, and because of the relationship between deviant arousal and recidivism, the accuracy of the self-report is of great significance.¹⁴⁴ A recent survey of studies concerning the connection of arousal to recidivism suggests that “phallometric measures, particularly of sexual interests in children, appear to have promise as somewhat weak predictors of reoffending, although it is also clear that these indices function best as part of a more comprehensive prediction package.”¹⁴⁵ Thus, the reliability issues associated with PPG testing are mitigated when it is used as a supplement to the self-report. This approach is consistent with the ATSA Practice Guidelines, used by the Board, which instruct that PPG test results are to be “interpreted in conjunction with other relevant information (such as, the individual's offending behavior, use of fantasy, and pattern of masturbation) to determine risk.”¹⁴⁶

Much of the reason that PPG testing is utilized in Tennessee sentencing decisions is the Board has endorsed the use of this procedure

141. *See supra* Section I.A.

142. *See supra* note 28 and accompanying text.

143. It should be noted that this analysis is based on the general science and methodology behind PPG testing; each inquiry under *McDaniel* will depend on the facts of a specific case.

144. *See supra* notes 38–40 and accompanying text.

145. Marshall, *supra* note 28.

146. ATSA PRACTICE GUIDELINES, *supra* note 15, at 70–71.

in psychosexual evaluations that are included in the presentencing reports.¹⁴⁷ Furthermore, defendants rarely challenge the use of the evaluation on appeal.¹⁴⁸ One of those rare challenges was in *State v. Edwards*, where the defendant challenged the use of PPG testing in a risk assessment considered at sentencing.¹⁴⁹ While the Tennessee Supreme Court later overruled *Edwards*'s holding on the substantive criminal offense at issue,¹⁵⁰ the evidentiary analysis in dicta should be considered. Additionally, the background facts that gave rise to the issues surrounding the admissibility of PPG evidence are helpful to understanding PPG testing in practice.

Edwards was initially subjected to a psychosexual evaluation and PPG test administered by Counseling and Consultation Services, Inc. (“CCS”).¹⁵¹ He was given a “Sexual Scenario Rating Scale” (“SSRS”) test created by CCS’s president as a type of self-report of arousal.¹⁵² The court described the other portion of the PPG test: “A monitoring device was placed on the Defendant’s penis to monitor ‘his penile response’ to several audio and visual stimuli. Devices were also attached to the Defendant to monitor his ‘Galvanic Skin Response (GSR) and respiration . . . to assist in detecting faking.’”¹⁵³ Essentially, Edwards was subjected to a PPG test and polygraph simultaneously.¹⁵⁴ The report by CCS determined that Edwards’s arousal was deviant due to significant responses to four stimuli during the PPG test and that his

147. See Tennessee Sex Offender Treatment Board, *supra* note 9.

148. For example, the defendant in *State v. Pierce* chose to not challenge the PPG test and only focus on the polygraph’s unreliability. 138 S.W.3d 820, 822 n.2 (Tenn. 2004) (“Pierce has not challenged the use of this particular test, and we express no opinion on the propriety of its use in this case.”).

149. No. E2010-01731-CCA-R3-CD, 2012 WL 1799025, at *23 (Tenn. Crim. App. May 18, 2012) (finding that there was no evidence of the scientific reliability of the PPG test), *overruled on other grounds* by *State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015).

150. The Tennessee Supreme Court rejected the Court of Criminal Appeals’ holding in *Edwards* that a jury instruction for criminal attempt was inappropriate when only two possible interpretations of the facts existed—that the offense was either completed or not even attempted. *Thorpe*, 463 S.W.3d at 861–63. Instead, the Supreme Court held that the lesser-included offense of attempt is available when the proof has fairly raised the completed offense and the charged offense has a requisite intent element. *Id.* The Court did not address the sentencing decision of the *Edwards* opinion. *Id.* at 863 n.5 (“We expressly overrule [*Edwards*] to the extent that it conflicts with this holding.”). Furthermore, the court in *Edwards* reversed and dismissed the lower court’s judgment for insufficient evidence; it reached the sentencing-hearing evidentiary issue only for the purpose of addressing all issues in the defendant’s appeal. 2012 WL 1799025, at *23 n.8 (“[W]e are only addressing this issue so as not to preterm it.”).

151. *Edwards*, 2012 WL 1799025, at *20.

152. *Id.*

153. *Id.* (omission in original).

154. Tools that measure the physiological changes associated with lying or faking are functionally identical to polygraphs. See THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 12–13 (stating the equipment used during polygraph examinations includes instruments to record respirations, heart rate, and electrical conductance at the skin surface).

reactions to the SSRS were typical of someone attempting to mislead the evaluator, because his answers were ambiguous.¹⁵⁵ Overall, CCS did not recommend alternative sentencing and found that he was a moderate risk to reoffend.¹⁵⁶

Edwards filed an objection to the CCS report and requested a second evaluation, which the court granted. The second test was conducted by Sex Offender Solutions (“SOS”).¹⁵⁷ The report by SOS also found that Edwards was a moderate risk to reoffend but stated there was no clinical reason that he should not participate in outpatient treatment via alternative sentencing.¹⁵⁸ At the sentencing hearing, Edwards presented the expert testimony of SOS’s vice president to refute the CCS report.¹⁵⁹ The expert testified that the SSRS self-report was not used or approved by any recognized clinical board and had not been peer reviewed.¹⁶⁰ He further testified that the PPG test was unreliable due to the lack of established research on the relation between erectile response and deviant arousal¹⁶¹—effectively echoing many of Professor Marshall’s concerns about the state of PPG research.¹⁶² Despite this testimony, the trial court considered both the CCS and SOS reports in its sentencing determination.¹⁶³ The court denied alternative sentencing because the two reports “indicate[d] basically the same conclusion, the Defendant has a moderate probability of re-offending,” and the CCS recommendation specifically cautioned against sending him back into the community.¹⁶⁴

The Tennessee Court of Criminal Appeals, in extending the polygraph holding of *Pierce* to PPG evidence, stated that a trial court is not required to consider a psychosexual evaluation based on otherwise inadmissible evidence.¹⁶⁵ In fact, as the gatekeeper with respect to expert evidence, the trial court abused its discretion by admitting and considering the CCS report.¹⁶⁶ Using the *McDaniel* framework, the court viewed the testimony of the SOS expert and the lack of evidence

155. *Edwards*, 2012 WL 1799025, at *20.

156. *Id.*

157. *Id.* at *21.

158. *Id.*

159. *Id.*

160. *Id.*

161. *Id.*

162. *See supra* notes 33–37 and accompanying text.

163. *Edwards*, 2012 WL 1799025, at *22.

164. *Id.*

165. *Id.*

166. *Id.* at *23.

supporting the reliability of the CCS report as demonstrating that the PPG and SSRS tests by CCS failed to meet the *McDaniel* standard.¹⁶⁷

While the court in *Edwards* referenced the *Pierce* decision, it declined to extend the per se exclusionary rule to PPG results, instead analyzing the admissibility of PPG evidence under Rule 702 and the *McDaniel* standard. The court's two options were to declare the PPG evidence wholly unreliable and therefore irrelevant under Rule 402, in the same fashion that the *Pierce* court dealt with polygraph evidence, or to conduct a Rule 702 balancing inquiry into the scientific evidence.¹⁶⁸ Given the *Pierce* decision to root the per se exclusionary rule for polygraph evidence in Rule 402,¹⁶⁹ the door was open to exclude PPG evidence as irrelevant as well, especially because of the procedures' base-level similarities (i.e., evaluation of physiological responses and lack of standardization among testing facilities)¹⁷⁰ and the purpose for which the results are used in court.

In addressing the question of what PPG testing is used to prove, the primary answer will likely always be proof of deviant arousal and thus recidivism risk. But this purpose runs counter to the use of PPG testing recommended by the scientific community. As previously noted, the majority of research indicates that PPG testing should not be used exclusively to determine deviant arousal. The problem of unreliability is much more significant when PPG serves this purpose due to the lack of consensus on the validity of the results¹⁷¹—therefore, the outcome under *McDaniel* will almost always be exclusion. However, consider again the National Academy of Sciences' assertion that “[v]alidity is not something that courts can assess in a vacuum.”¹⁷² Proponents of PPG testing will argue that the results are not used exclusively to determine deviant arousal, which is true if the ATSA guidelines are followed. Therefore, an alternative, and more viable, purpose for introducing PPG testing is as proof that the information given by the offender for the presentencing report is accurate and credible. While not perfectly comparable, this purpose is much more analogous to polygraphy, and given the similarities between the theoretical and practical application

167. *Id.*

168. *See* *State v. Pierce*, 138 S.W.3d 820, 826 (Tenn. 2004) (confirming Tennessee's rule that polygraph evidence is inherently unreliable and therefore irrelevant and inadmissible).

169. *See supra* notes 125–126 and accompanying text.

170. *Compare* THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 212 (describing the problems with the measurement of physiological responses when conducting polygraph examinations), *with* Merdian & Jones, *supra* note 23, at 142 (explaining how PPG testing measures the physiological response of arousal).

171. *See supra* notes 42–43 and accompanying text.

172. THE POLYGRAPH AND LIE DETECTION, *supra* note 113, at 204.

of the evaluations, the evidence should be analyzed the same in the context of sentencing decisions.

In sum, evidence law in Tennessee—specifically concerning the treatment given to inherently unreliable polygraph evidence—raises the question why the rules do not treat PPG results with the same rigor. The expert testimony analysis of PPG testing under *McDaniel* presents a showing of unreliability that should, more often than not, lead to exclusion under Rule 702. The similar purposes of polygraphy and PPG testing imply that the inherent unreliability is present in both types of evidence and at least suggest that a strict analysis of admissibility is reasonable.

B. The Ethical and Practical Limitations of PPG Testing

In addition to the purely evidentiary reasons that PPG testing be limited in the sentencing context, various ethical and practical concerns about the test's current use arise in the sex-offender sentencing system in Tennessee. These concerns are not contemplated by evidence law but should nonetheless be considered as reasons for reform due to their use in highly consequential sentencing decisions. One major ethical issue is the concept of punishing an individual on the basis of his physiological response to sexual stimuli, regardless of deviancy. The connection between arousal and recidivism risk implies that those who are aroused by deviant material are so dangerous that they must be incarcerated. Yet there are various issues with using only arousal as a standard to measure dangerousness. The U.S. Supreme Court recently stated that a sentence based on a determination of dangerousness that is informed, in whole or in part, by an immutable characteristic threatens to depart from “a basic premise of our criminal justice system: Our law punishes people for what they do, not who they are.”¹⁷³ Many scholars have critiqued the use of measurements that suggest dangerousness within the criminal justice system and have concluded that limitations should be created that only allow its use when it can be shown that “the individual will engage in harmful activity in the absence of the intervention.”¹⁷⁴ The U.S. Supreme Court has maneuvered around this obstacle by distinguishing between the criminal and civil contexts of sexually violent–predator laws to allow

173. *Buck v. Davis*, 137 S. Ct. 759, 778 (2017) (“Buck may have been sentenced to death in part because of his race. As an initial matter, this is a disturbing departure from a basic premise of our criminal justice system: Our law punishes people for what they do, not who they are.”).

174. Christopher Slobogin, *Dangerousness as a Criterion in the Criminal Process*, in *LAW, MENTAL HEALTH, AND MENTAL DISORDER* 360, 379 (Bruce D. Sales & Daniel W. Shuman eds., 1996).

the civil commitment of individuals designated as sexually dangerous.¹⁷⁵

Punishment based on an individual's perceived dangerousness can be viewed as a broad argument against risk assessments in general. More specifically, using PPG testing as a measurement of dangerousness is an example of the state scrutinizing an individual because of his thoughts and feelings, even though he may be capable of controlling his arousal and acting in a legal manner. The U.S. Constitution protects the freedom of mind in a number of ways—courts have used the First, Fourth, and Fifth Amendments to establish a general freedom of mind through the rights to speak, remain silent, and enjoy privacy.¹⁷⁶ Nevertheless, it is likely that a federal constitutional challenge to PPG results used at sentencing would be unsuccessful because of both the diminished rights of those convicted of crimes and the broad discretion afforded to sentencing judges.

As an ethical matter, when deciding an offender's risk level or dangerousness, PPG clinicians should account for the ability of an individual to choose his actions and consciously decide not to act on deviant thoughts. Interestingly, the ethical rules put forth by the ATSA and adopted by the Board emphasize that informed consent to do PPG testing is required.¹⁷⁷ Yet some jurisdictions allow an inference of dangerousness from the failure to subject oneself to the test.¹⁷⁸ But one could imagine that many offenders would not volunteer for the test either because of the possible negative results or because of the intrusiveness of the procedure. Those who do volunteer would likely be confident in their ability to fake the test or in their own lack of deviancy, which would present less assistance in identifying high-risk individuals.

Nevertheless, it may still be useful to distinguish between those who can choose to suppress their arousal, suggesting they can choose not to act on their deviant thoughts, and those who cannot control their arousal, suggesting they are susceptible to succumbing to their desires. From a free-will perspective, this information is helpful in diverting those who can choose their actions to alternative sentencing and those

175. See John M. Fabian, *Kansas v. Hendricks, Crane and Beyond: "Mental Abnormality," and "Sexual Dangerousness": Volitional vs. Emotional Abnormality and the Debate Between Community Safety and Civil Liberties*, 29 WM. MITCHELL L. REV. 1367, 1383–84 (2003) (explaining the U.S. Supreme Court's distinction between civil and criminal proceedings in *Kansas v. Hendricks*, 521 U.S. 346, 369 (1997)).

176. See Odeshoo, *supra* note 8, at 28–31 (focusing on the First Amendment's protection of free speech in discussing constitutional protections of the freedom of mind).

177. ATSA PRACTICE GUIDELINES, *supra* note 15, at 26 (recommending that clinicians obtain specific informed consent from subjects prior to conducting PPG testing).

178. See *supra* note 83 and accompanying text.

who do not or cannot to incarceration or specified rehabilitative centers. Unfortunately, from a practical perspective, the ethical recommendations of voluntary consent seem to diminish the effectiveness of making such determinations. This should be considered by sentencing judges when deciding to rely on PPG results within sex-offender risk assessments, given their broad discretion to make sentencing decisions and their duty to take an active role in evaluating the reliability of expert evidence.¹⁷⁹

Finally, one major practical limitation of considering PPG results at sentencing hearings is the financial cost. The Tennessee legislature has dictated that the cost of risk assessments and treatment are to be incurred by the offender.¹⁸⁰ Furthermore, if an offender desires probation or alternative sentencing, that offender is required to submit to an evaluation to determine risk level.¹⁸¹ Regardless, the same statute dictates that every sex offender found guilty at trial or who pleads guilty is required to submit to an evaluation.¹⁸² Virtually every convicted offender is required to pay for their own risk assessment unless deemed indigent by the court.¹⁸³ If an offender is indigent, the Board will pay treatment professionals up to \$850, depending on the type of psychosexual evaluation completed.¹⁸⁴ The amount paid for a PPG test specifically is unclear; objective testing is allocated \$175 and includes both polygraph and PPG.¹⁸⁵ These amounts are startling because the test can cost up to \$1,000 per session.¹⁸⁶ The level of expertise and detail taken in the vast majority of cases is unsurprisingly low¹⁸⁷—a predictable outcome, considering the low amount of funds available to indigent offenders, nonindigent defendants' desire to go with the cheapest provider available, and the lack of providers with clinical

179. See *supra* note 104 and accompanying text.

180. See TENN. CODE ANN. §§ 39-13-705(c), -706 (2018).

181. See *id.* § 39-13-705(b).

182. *Id.*

183. See *Public Hearing*, *supra* note 48, at 7–8 (reporting that “offenders are required to pay for all costs associated with treatment and evaluations unless they . . . have been deemed indigent by the Court”).

184. See *Reimbursement Management*, *supra* note 48, at 1–2 (breaking down reimbursement rates for evaluations of sex offenders).

185. *Id.*

186. Odeshoo, *supra* note 8, at 8.

187. For example, the expert witness in *State v. Edwards* specifically called into question the expertise and detail of CCS, which the court weighed heavily in concluding that the results from the CCS report were unreliable. See No. E2010-01731-CCA-R3-CD, 2012 WL 1799025, at *23 (Tenn. Crim. App. May 18, 2012), *overruled on other grounds* by *State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015). CCS is commonly used as an assessor of sex offenders for presentence reports and was also the organization that conducted the report in *State v. Pierce*, which was similarly considered unreliable. See Brief of Appellant at 13, *State v. Pierce*, 138 S.W.3d 820 (Tenn. 2004) (No. E2001-01734-SC-R11-CD).

competency to conduct the assessments.¹⁸⁸ Given this fact, sentencing judges should take into account the economic incentives of treatment providers and recognize that many offenders are financially forced into poorly administered evaluations.

III. A RECOMMENDED RULE FOR ADMISSION OF PPG RESULTS AT SENTENCING

The facts in *Edwards* demonstrate that conclusions regarding a sex-offender's risk level do not always change with the addition of PPG results, showing that the assessment of risk can be at least as accurate without the test.¹⁸⁹ However, it is also clear that a properly conducted PPG test can reveal valuable information that could impact whether an offender should receive alternative sentencing. Making the presentencing risk assessments more accurate can reduce prison populations and reroute individuals to proper treatment or rehabilitative institutions. PPG testing's usefulness as a rehabilitative tool is much more relevant in the context of sexual crimes, which commonly involve mentally ill individuals.¹⁹⁰ At the same time, public-safety goals are served when individuals who cannot control their actions based on deviant arousal are incapacitated. The Tennessee Sentencing Commission considers incapacitation as one of the theoretical foundations of the Tennessee Criminal Sentencing Reform Act of 1989.¹⁹¹ The purpose of the Act recognizes that, in all cases concerning the most serious offenses or involving offenders with criminal histories that show a "clear disregard for the laws and morals of society" or "failure of past efforts at rehabilitation," incarceration should be the first priority.¹⁹² Therefore, PPG testing seems to have a place at the sentencing phase in Tennessee.

For these reasons, a per se exclusionary rule, like that for polygraph evidence, would be unwise. Yet the similarities in the science behind both PPG testing and polygraphy, as well as the purpose for which both types of evidence are introduced, seem to warrant similar treatment. One of the justifications for the per se exclusionary rule for

188. See *Public Hearing*, *supra* note 48, at 7 (describing the issues contributing to the backlog of assessments).

189. While CCS used PPG testing and SOS did not, both reports found a moderate level of risk. See *Edwards*, 2012 WL 1799025, at *20–21.

190. See Fabian, *supra* note 175, at 1369 (describing the relation between mental abnormality and sexual dangerousness).

191. TENN. CODE ANN. tit. 40, ch. 35 (2018).

192. *Id.* § 40-35-102(5).

polygraph evidence is avoiding the usurpation of the fact finder's role.¹⁹³ That is, the jury (or judge) should consider the credibility of the witnesses instead of relying on the results of a machine. Such justification has little significance in the sentencing context, given that the judge making the decision has the discretion to take into account a variety of considerations and likely will be influenced by the fact that the offender has already been found guilty, in some form or another, of a sex offense. Furthermore, the per se exclusionary rule's lack of an articulation of the reliability of PPG results does little to protect the essence of Rule 702.

While courts could continue to rely on a Rule 702 analysis using *McDaniel*, such analysis is generally only discussed in any detail at the sentencing hearing when one of the parties challenges the PPG results. Given the ethical, practical, and constitutional implications of the testing that are not within the purview of Rule 702, along with the pervasive reliability problems relevant to the analysis in every case, the law should require a more stringent inquiry. Accordingly, the rule for admission of polygraph evidence in New Mexico provides a framework that accomplishes an appropriate level of scrutiny by offering an intermediate option between per se exclusion and a case-by-case judicial determination of the reliability of the science.

New Mexico is the only state that allows for the routine admission of polygraph evidence,¹⁹⁴ adopting the position that a categorical exclusion of polygraph results would be unwise.¹⁹⁵ The need for a standardized evaluation of the evidence for admission was tantamount to ensuring the reliability of polygraph evidence.¹⁹⁶ New Mexico Rule of Evidence 11-707(C), governing the admissibility of polygraph results, states:

C. Admissibility of results. A polygraph examiner's opinion as to the truthfulness of a person's answers in a polygraph examination may be admitted if:

- (1) the polygraph examination was administered by a qualified polygraph examiner;
- (2) the polygraph examination was quantitatively scored in a manner that is generally accepted as reliable by polygraph experts;

193. See Shniderman, *supra* note 16, at 443 (noting that usurpation of the jury function is one of two primary arguments in favor of excluding polygraph evidence).

194. See *id.* at 442 ("Only New Mexico allows for the routine admission of polygraph evidence.").

195. See *Lee v. Martinez*, 96 P.3d 291, 306 (N.M. 2004) (holding that polygraph evidence is admissible in certain situations).

196. *Tafoya v. Baca*, 702 P.2d 1001, 1004–05 (N.M. 1985) ("This Court adopted Rule 707 in an attempt to standardize the admission of such test results . . . and assure the reliability of polygraph test evidence that was to be admitted.").

- (3) the polygraph examiner was informed as to the examinee's background, health, education, and other relevant information prior to conducting the polygraph examination;
- (4) at least two (2) relevant questions were asked during the examination;
- (5) at least three (3) charts were taken of the examinee; and
- (6) the entire examination was recorded in full on an audio or video recording device, including the pretest interview and, if conducted, the post-test interview.¹⁹⁷

If a party plans to introduce polygraph evidence, they must provide written notice, with a copy of all relevant clinical documentation, no less than thirty days prior to the judicial proceeding.¹⁹⁸ Furthermore, the rule dictates that no witness be forced to take a polygraph examination and that the court may compel a second examination if the party challenging the admission shows good cause.¹⁹⁹ In that case, the results from the subsequent test are admissible if conducted in accordance with the rule; if the individual refuses to take the second test, the results of the first are inadmissible as well.²⁰⁰ The prerequisites for the admission of polygraph evidence, designed to ensure consistently reliable results, present a nexus with the expert evidence rules that would be nonexistent with a per se exclusionary rule.

This same structure can allow for admission of PPG results at sentencing while protecting against its negative implications. The rule would require the party admitting the PPG results to lay the proper foundation that would be traditionally required at trial under Rule 702. The last three requirements under New Mexico Rule 11-707(C) can be adjusted to reflect the technicalities of PPG procedures. For example, in place of the Rule 11-707(C)(4) requirement, there should be a condition that at least two relevant stimuli be presented to the subject, with relevance determined by the offense for which the subject was convicted. Additionally, in the same vein as the Rule 11-707(C)(6) requirement, extensive written records must be kept regarding the procedure, self-reporting, and interviews; audio or video recordings of PPG testing would not necessarily be warranted given the intrusiveness of the procedure. Finally, the same procedural protections regarding notice, challenge, and a second evaluation would guard against the situation at issue in *Edwards*, where the offender's first evaluation was improperly conducted.²⁰¹ The foregoing requirements

197. N.M. R. EVID. 11-707(C).

198. N.M. R. EVID. 11-707(D).

199. N.M. R. EVID. 11-707(F).

200. *Id.*

201. See *State v. Edwards*, No. E2010-01731-CCA-R3-CD, 2012 WL 1799025, at *23 (Tenn. Crim. App. May 18, 2012) (holding that the trial court abused its discretion in admitting the CCS

create a forcing function for each case that would produce consistent use of reliable PPG results, as opposed to separately evaluating the varied methods each time the test is challenged at the sentencing phase.

In practice, assuming that the offender consents to a psychosexual evaluation that includes a PPG test, challenges will inevitably arise when one party disagrees with the findings of the risk assessment and requests a preliminary hearing to argue the challenge. When the results indicate that the offender has a higher risk of reoffending, the offender will likely challenge the validity of the test under this proposed rule. Conversely, when the results suggest a low risk, the State will likely challenge in cases where incarceration is the goal. If the challenger cannot show cause to warrant a second evaluation, then presumably none is needed, and the results are admissible. A good-cause showing would look something like the attack of the CCS report in *Edwards*,²⁰² while a showing by the opposing party that the evaluation was completed in accordance with the rule and is thus reliable (essentially a *McDaniel* analysis) would foreclose a showing of good cause. This preliminary hearing to determine good cause will essentially perform the function of forcing the proponent to lay the proper foundation. Nevertheless, even if there is no challenge, the party seeking to introduce the results at the sentencing hearing still must show adherence to the rule and, essentially, the requisite level of reliability. Even though the opposing party has the burden of showing good cause for a second evaluation, the burden should not shift to the opposing party to show reliability for admission at the sentencing hearing—it should remain with the party introducing the evidence and be met prior to any challenges.²⁰³

Providing the chance for a second evaluation if the challenger shows good cause will give the opposing party an opportunity to provide the sentencing judge with multiple data points instead of just one. If an offender displays no deviant arousal and the State shows good cause to require a second evaluation, the offender can refuse, and the first evaluation will simply be inadmissible. A major reform that the legislature must address along with adoption of this rule would be the

report during its sentencing hearing), *overruled on other grounds* by *State v. Thorpe*, 463 S.W.3d 851 (Tenn. 2015).

202. *See id.* (recounting expert testimony that criticized the CCS report).

203. Shifting the burden to the opposing party would be inconsistent with the requirement for the proponent to lay the proper foundation for admission. *See, e.g., Billips v. Commonwealth*, 652 S.E.2d 99, 101–02 (Va. 2007) (holding PPG results inadmissible at sentencing hearing because of lack of foundation). The Supreme Court of Virginia in *Billips* overturned the lower court's decision to place the initial burden on the opponent of the evidence to show unreliability; the high court declared that the party introducing the evidence must show reliability. *See id.*

financial structure of the testing. Currently, the offender must pay for all evaluations, unless he is indigent. With the addition of a potential second PPG test, a nonindigent offender may be forced to pay twice. If the offender shows good cause and takes a second test, the financial burden can remain with the offender. But if the State shows good cause and the offender is stuck between paying for a second evaluation or keeping out potentially helpful evidence, the financial burden should shift to the State. In the case of indigent offenders, the State should maintain the current reimbursement structure for all tests, hopefully incentivizing investments in reliable PPG testing facilities to avoid paying for a second test.

CONCLUSION

While this Note advocates for a rule specifically for admission of PPG evidence at sentencing hearings, there are further potential interactions with other evidence, such as the polygraph or viewing-time evaluations, that could assist in the reliability of PPG testing when used to assess the arousal of convicted sex offenders. The *per se* exclusion of polygraph evidence could be unnecessary when used in conjunction with PPG testing to achieve the same purpose—evaluating both results together would reduce the reliability concerns of each test in isolation. While attacking the *per se* exclusionary rule of polygraph evidence is beyond the scope of this Note, the foundational rule presented in Part III could easily be applied to polygraph evidence for the limited purpose of admission at sentencing hearings.

PPG testing has improved dramatically since its origin, discussed at the opening of this Note. Using the results to determine whether to divert convicted offenders to alternative sentencing assists in the rehabilitation of those offenders who are most likely to respond to treatment and protects the public from offenders whose risk to reoffend is substantially high. A rule that allows for consensual participation in the testing and presents adversarial opportunities to question its reliability helps to both protect the rights and liberties of the offenders and avoid abuse of the evidence by individuals who might attempt to misrepresent their responses. Further advances in the science and understanding of psychosexual traits will present the opportunity for increased reliability—the Board should continue to

focus on standardization of PPG testing, which will undoubtedly aid in advances in this area of psychology.

*Steven Poland**

* J.D. Candidate, 2019, Vanderbilt University Law School; B.A., 2012, Western Kentucky University. I would like to thank Professors Edward Cheng, Nancy King, and Christopher Slobogin for their advice, guidance, and insight, and the editors and staff of the *Vanderbilt Law Review* for their diligent work in editing and refining this Note. Last but certainly not least, I would like express my gratitude to my family, especially my wife, Courtney, for their unwavering love and support in the pursuit of my dreams.