ABRACADABRA, HOCUS POCUS, SAME SONG, DIFFERENT CHORUS: THE NEWEST ITERATION OF THE "SCIENCE" OF LIE DETECTION

Carrie Leonetti*

PART I. INTRODUCTION

[1] The "lie detector" is the holy grail of criminal investigation and adjudication. Dissatisfied with the jury's ability to sort truth-tellers from liars, police, prosecutors, and criminologists have longed dreamed of a scientific talisman that would definitively sort the guilty from the innocent.

[2] The earliest iteration of a physiologically based lie-detector test was the polygraph, which measures galvanic skin response, blood pressure, heart and breathing rates, and perspiration as a proxy for nervous-system activity (primarily anxiety) as an (imperfect) proxy for deception.[1] The polygraph has been the primary method of "lie detection" in the criminal-justice system for a century, but it has also long been viewed with suspicion because of concerns with its reliability and the effectiveness of "countermeasures"[2] at defeating it ("beating the poly"), although it

* Associate Professor and Dean's Distinguished Faculty Scholar, University of Oregon School of Law.

[1] *See* Comm. Review Sci. Evidence Polygraph, Nat'l Research Council, THE POLYGRAPH & LIE DETECTION 11, 32 (2003) [hereinafter NRC Report], https://www.nap.edu/read/10420/chapter/1, https://perma.cc/NW97-B4YK; *see also* Cooper Ellenberg, *Lie Detection: A Changing of the Guard in the Quest for Truth in Court*, 33 L. & PSYCHOL. REV. 139, 141 (2009).

[2] *See* Paul Root Wolpe et al., *Emerging Neurotechnologies for Lie Detection: Promises and Perils*, 5 AM. J. BIOETHICS 39 (2005); *see, e.g.*, Ben Kleine, *Two Subjects of Interest in 1992 Homicide to Take Polygraph Tests*, SOUTHEAST MISSOURIAN (Jan. 15, 2017), http://www.semissourian.com/story/2377171.html, https://perma.cc/TA5B-9F95; Lisa Dayley Smith, *Man Faces up to 12 Years for Rape in Sugar City,* REXBURG STANDARD J. (Feb. 28, 2017) http://www.rexburgstandardjournal.com/news/local/man-faces-up-to-years-for-rape-in-sugar-city/article_d60931a2-fe0b-11e6-8667-f391be10461b.html, https://perma.cc/K2VZ-JT2X; Christina Sterbenz, *This Ex-Cop Thinks Lie-Detector Tests Are So Inaccurate He's Facing 100 Years in Prison for Starting a Website That Taught*

continues to be used in security screening for sensitive employment.[3] The polygraph played the starring role in *Frye v. United States*,[4] the United States Court of Appeals for the District of Columbia Circuit's seminal case establishing the necessary foundation that a proponent of scientific evidence must present as a precursor to its admissibility.[5] The *Frye* test continues to be followed by many states today.[6] For decades, almost all state courts and the federal court system, by statute and/or case law, have either prohibited the admissibility of polygraph evidence entirely or permitted its admission only with the mutual consent of both parties.[7] In 2003, the National Research Council of the National Academy of Sciences ("NRC") published

People How to Cheat Them, BUS. INSIDER, (May 18, 2015), http://www.businessinsider.com/the-crazy-story-of-an-ex-cop-who-ran-a-website-that-taught-people-how-to-cheat-polygraphs-2015-5, https://perma.cc/JMG9-YDZT.

[3] *See* Daniel D. Langleben & Jane Campbell Moriarty, *Using Brain Imaging for Lie Detection: Where Science, Law, and Policy Collide,* 19 PSYCHOL. PUB. POL'Y & L. 222, 223 (2013).

[4] Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923) (holding that the technique underlying a lie detector based on systolic blood pressure, a precursor to the modern polygraph, was inadmissible because it was too novel and had not yet gained general acceptance in its field).

[5] *See id.* at 1014.

[6] Alice B. Lustre, Annotation, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence is State Courts*, 90 A.L.R. 5th 453 (2017); Robert J. Goodwin, *Fifty Years of Frye in Alabama: The Continuing Debate Over Adopting the Test Established in Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 35 CUMB. L. REV. 231, 234 (2005).

[7] *See, e.g.,* ARK. CODE ANN., § 12-12-704 (1975); United States v. Piccinonna, 885 F.2d 1529, 1535 (11th Cir. 1989) (explaining that polygraphs were inadmissible under Florida law); Donegan v. McWherter, 676 F. Supp. 154, 157-58 (M.D. Tenn. 1987) (explaining that polygraph test results were inadmissible under Tennessee law); State v. Porter, 698 A.2d 739, 758–59 (Conn. 1997); People v. Baynes, 430 N.E.2d 1070, 1077 (Ill. 1981); State v. Kolander, 52 N.W.2d 458, 465 (Minn. 1952); People v. Leone, 255 N.E.2d 696, 697 (N.Y. 1969). *See also* DANIEL L. FAIGMAN ET AL., 5 MODERN SCIENTIFIC EVIDENCE § 38:3 (2017).

a report questioning the theoretical underpinnings and validity of the results of polygraph tests.[8]

[3] The polygraph's problem is one of correlation: the link between stress and lying is too tenuous and poorly understood for the former to be a reliable measure of the latter.[9] The NRC, therefore, called for alternatives to the polygraph that would use different measurements of deception rather than relying exclusively on emotional responses.[10] The traditional hostility of criminal courts toward the results of polygraph examinations has led to a quest by researchers to provide a more reliable lie-detection alternative. Previous pseudo-scientific lie-detection techniques, some of which nonetheless continue to be popular in law-enforcement circles, include

[8] *See* NRC Report, *supra* note 2, at 22, 24.

[9] Polygraph tests are premised on the assumption that people show stronger emotional responses to test questions they answer deceptively than those that they answer truthfully. John C. Kircher & David C. Raskin, *Laboratory and Field Research on the Ocular-motor Deception Test,* 10 EUR. POLYGRAPH 159, 161 (2016); *see also* Renee McDonald Hutchins, *You Can't Handle the Truth! Trial Juries and Credibility,* 44 SETON HALL L. REV. 505, 528 (2014) (explaining that "evidence of any one of the 'lying' emotions is not necessarily conclusive proof of dishonesty"); John O'Neil, *VITAL SIGNS: IN THE LAB; Zeroing in on a Lie's Home Base,* N.Y. TIMES (Dec. 4, 2001), http://www.nytimes.com/2001/12/04/health/vital-signs-in-the-lab-zeroing-in-on-a-lie-s-home-base.html?mcubz=3, https://perma.cc/G35Y-LUWM (last visited Aug. 19, 2017) (explaining that "[l]ie detector tests, or polygraphs, do not measure lying; they measure the fear of getting caught in a lie, by tracking things like heart rate, blood pressure and sweat, which are considered to be good reflections of anxiety levels. But not all nervous people are lying, and some liars are adept at concealing their anxiety, making polygraphs too unreliable to be accepted as evidence in most courts.").

[10] *See* NRC Report, *supra* note 2, at 212.

analyses of "precognitive" facial clues[11] and eye movements.[12] When these techniques proved to be as unreliable as the polygraph, research psychologists responded with even newer purported tests of deception.[13]

[11] *See* ALDERT VRIJ, DETECTING LIES AND DECEIT: THE PSYCHOLOGY OF LYING AND THE IMPLICATIONS FOR PROFESSIONAL PRACTICE 75 (2000) (noting that experts trained in the use of facial clues to detect deception are only barely more accurate than chance); Hutchins, *supra* note 9, at 528 (explaining that some speakers do not display the facial expressions associated with deception and that some can manipulate their facial muscles to indicate honesty when they are lying); Olin Guy Wellborn III, *Demeanor,* 76 CORNELL L. REV. 1075, 1088 (1991) (documenting how the observation of facial behavior diminished the accuracy of lie detection); *see generally* PAUL EKMAN, TELLING LIES: CLUES TO DECEIT IN THE MARKETPLACE, POLITICS, AND MARRIAGE (2009) at 129–37, 350–53 (discussing "micro expressions," which are full-face emotional expressions that are compressed in time, so quick they are usually not seen). *See generally* MALCOLM GLADWELL, BLINK: THE POWER OF THINKING WITHOUT THINKING (2005) (discussing how great decision makers have perfected the art of "thin-slicing" – filtering the very few factors that matter from an overwhelming number of variables); PHILIP HOUSTON ET AL., SPY THE LIE (2012) (explaining methods former CIA officers use to detect deception); Ursula Hess & Robert E. Kleck, *Differentiating Emotion Elicited and Deliberate Emotional Facial Expressions, in* WHAT THE FACE REVEALS: BASIC AND APPLIED STUDIES OF SPONTANEOUS EXPRESSION USING THE FACIAL ACTION CODING SYSTEM (FACS) 272 (Paul Ekman & Erika L. Rosenberg eds., 2d ed. 2005) (discussing the difference between spontaneous emotional smiles and deceptive smiles); David Matsumoto & Hyi Sung Hwang, *Evidence for Training the Ability to Read Microexpressions of Emotion,* 35 MOTIVATION & EMOTION 181 (2011) (discussing the concept of micro expressions); Stephen Porter & Leanne ten Brinke, *Reading Between the Lies: Identifying Concealed and Falsified Emotions in Universal Facial Expressions,* 19 PSYCHOL. SCI. 508, 513 (2008) (discussing the shortcomings of using microexpressions to detect deception).

[12] *See* Jeremy A. Blumenthal, *A Wipe of the Hands, A Lick of the Lips: The Validity of Demeanor Evidence in Assessing Witness Credibility,* 72 NEB. L. REV. 1157, 1194 (1993) (explaining that eye movements are highly unreliable source of deception clues); Bella M. DePaulo et al., *Cues to Deception,* 129 PSYCHOL. BULL. 74, 93 (2003) ("The 32 independent estimates of eye contact produced a combined effect that was almost exactly zero."). *See generally* Reginald B. Adams, Jr. & Robert E. Kleck, *Effects of Direct and Averted Gaze on the Perception of Facially Communicated Emotion,* 5 EMOTION 3 (2005) (discussing the link between gaze and emotional behavior); Dacher Keltner et al., *Appeasement in Human Emotion, Social Practice, and Personality,* 23 AGGRESSIVE BEHAV. 359, 362 (1997) (discussing human embarrassment and shame displays).

These new tests, however, continue to share Paul Grice's model of implied signal detection as their common validation framework.[14]

[4] One new proposed technique is the ocular-motor-detection test ("ODT"), the brainchild of a group of psychologists at the University of Utah.[15] The ODT is an automated, cognitive-based test of deception that measures participants' ocular-motor responses (pupil and blink behaviors),[16] using an infrared camera that records two-dimensional gaze position, pupil diameter and positional information, and blink rates.[17] According to its developers, study subjects who were lying in a computer questionnaire about their involvement in mock crimes had smaller pass durations and, more importantly, "increased pupil responses."[18] In other words, liars had larger pupil diameters when they read the study questionnaires than truth tellers, and they gave their answers to questions related to their guilty knowledge more quickly. According to the researchers, larger pupil diameters indicated the increased cognitive effort

[13] *See* Steven J. Luck, AN INTRODUCTION TO THE EVENT-RELATED POTENTIAL TECHNIQUE 4-5 (2005).

[14] Signal detection analysis assesses responses in terms of whether a particular implied "signal" can be inferred from contextual evidence when it differs from what is actually said – in the case of lie detection, a deception signal that indicates that the speaker is implying something other than the literal words that are spoken. *See* PAUL GRICE, STUDIES IN THE WAY OF WORDS (1989).

[15] *See* Anne E. Cook et al*., Lyin' Eyes: Ocular-motor Measures of Reading Reveal Deception,* 18 J. EXP. PSYCHOL. APPL. 301–02 (2012).

[16] *See id.* at 1; *see also* Pooja Patnaik et al., *Generalizability of an Ocular-Motor Test for Deception to a Mexican Population*, 6 INTL. J. APPLIED PSYCHOL. 1 (2016).

[17] *See* Gwen Klein Kirschner, *The Tariff Classification of the EyeDetect System* (U.S. Customs Service, Treasury Dept., Jan. 31, 2014), *available at* 2014 WL 892273, at *1 [hereinafter *Tariff Classification Ruling*] (stating that the ODT device, marketed as "EyeDetect," is proprietary, and its developers have a commercial interest in its sale); *See also* Kircher & Raskin, *supra* note 9, at 159.

[18] *See* Cook, et al*., supra* note 15, at 1.

involved in lying, in comparison to telling the truth (on the theory that lying is a more difficult mental task for humans than telling the truth).[19] Another technique involves recording brain waves to detect "event related potentials" or "ERPs."[20]

[5] This technique is sometimes known more colloquially as "brain fingerprinting."[21] The ERP technique relies on an electroencephalogram ("EEG"), the electrodes of which detect electrical activity in the human brain.[22] Over time, scientists have mapped this electrical activity, using statistical averages and time-frequency analyses of the waveforms, to identify neural responses associated with specific sensory, cognitive, and motor events.[23] ERPs are these specific responses.[24] Researchers have proposed using ERPs to detect deception, as well as serious mental illnesses like schizophrenia.[25] The academic developers of both ODT and ERP have formed private companies that market them for use as forensic lie

[19] *See, e.g.,* United States v. Semrau, 693 F.3d 510 (6th Cir. 2012) (holding, as a matter of first impression, that fMRI testing indicating truthfulness was inadmissible under the Federal Rules of Evidence). This "cognitive effort" hypothesis is consistent with functional Magnetic Resonance Imagery ("fMRI") data, which suggest that increased brain activity is needed to inhibit truthful answers as a precursor to lying. *See, e.g.*, Hakun*, infra* note 20, at 518. Of course, fMRI itself has struggled to gain judicial acceptance as a reliable lie detector.

[20] *See* J. G. Hakun et al., *Towards Clinical Trials of Lie Detection with fMRI,* 4 SOC. NEUROSCI. 518, at 520–21 (2009).

[21] *See, e.g.*, BRAIN FINGERPRINTING TECHNOLOGY, http://www.brainwavescience.com (last visited Aug. 19, 2017).

[22] *See* Steven J. Luck, AN INTRODUCTION TO THE EVENT-RELATED POTENTIAL TECHNIQUE 3–4 (2005).

[23] *See id.* at 4.

[24] *See id.*

[25] *See* BRAIN FINGERPRINTING, *supra* note 21.

detectors.[26] Unfortunately, like their predecessors, these new techniques are insufficiently developed to be reliable enough for use in criminal investigations and trials for the purpose of determining the veracity of the statements made by suspects, witnesses, and other relevant parties.[27] First, these techniques have not yet been sufficiently validated scientifically, even at a theoretical level. Second, even assuming they are scientifically valid as laboratory methodologies, there are significant gaps between their laboratory use and their proposed criminal-justice uses that render them unreliable in this application.

[6] This article argues that these new techniques, like the polygraph tests that came before them, are not, and likely never will be, ready to be used in criminal investigations and trials to reliably sort truth-tellers from liars. Section II argues that these techniques lack sufficient classification accuracy because the underlying mechanisms behind the measured physiological changes that they record are not understood. This means that the causal inference of deception that is drawn from the measured physiological changes is not scientifically valid. It explains how these tests rely upon a series of judgment calls relating to the subjective cutoffs that test administrators use to distinguish "normal" (truthful) responses from "abnormal" (deceptive) ones, which inject an unacceptable level of human error into the interpretation of test results.[28] It notes that the experimental

[26] *See*, *e.g.*, EYEDETECT, http://converus.com/eyedetect/, https://perma.cc/L38Q-QHK2 (last visited Oct. 27, 2017) (marketing an ODT test as "EyeDetect" by a company called Converus); Joseph Neighbor, *A Startup Wants to Use Eye Tracking to Detect If Syrian Refugees Are Terrorists,* VICE (Feb. 9, 2016), https://motherboard.vice.com/en_us/article/kb7wxw/eyedetect-converus-eye-tracking-lie-detector-syrian-refugees, https://perma.cc/9FB9-3NSP (describing Converus' goal to use the software for refugee vetting).

[27] *See generally* Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993) (describing the requirement that technology must be sufficiently developed in order to be reliable and thus, admissible).

[28] *See* Marc Green, *Human Factors in Forensic Evidence* (2013), http://www.visualexpert.com/Resources/forensics.html, https://perma.cc/JH2L-9L6E (last visited Oct. 27, 2017).

laboratory results that purport to validate the classification accuracy of these techniques fail to account for questioner expectancy and subject effects and produce results that are not replicable under real-world conditions, a necessary precondition for external validity and generalizeability.[29] It also notes that the developers of these new techniques have not tested whether they can be defeated with counter-measures like other, older purported lie-detection techniques.[30]

[7] Section III argues that these validity concerns, in turn, give rise to concerns regarding the reliability of generalizing from these group-level studies to the brain function of any given individual. It argues that, even if the group-level results were scientifically valid, they would nonetheless lack reliable predictive value at the individual level, which would require a differential approach tailored to the particular individual whose veracity was at issue. It also argues that, because the tests deal in aggregate averages, they are not going to be accurate for "outliers," people whose pupils or brains behave in ways that deviate significantly from the mean of the population used to develop the tests.

[8] The conclusion argues that these validity concerns are particularly acute in the context of the criminal-justice system. It notes that criminal law has been quick to adopt under-validated forensic techniques in the past, and that this has resulted in wrongful convictions. It concludes that the criminal-justice system should be the last arena in which these types of techniques are deployed, rather than the first.

[29] *See* Paul C. Giannelli, *Polygraph Evidence: Part I*, *Faculty Publications* at 270–71 (1994), http://scholarlycommons.law.case.edu/faculty_publications/339, https://perma.cc/2VUG-CR64 (last visited Oct. 27, 2017).

[30] *See* Paul Root Wolpe et al., *Emerging Neurotechnologies for Lie-Detection: Promises and Perils*, 5 AM. J. OF BIOETHICS 39 (Aug. 19, 2006).

PART II. MEASUREMENT VALIDITY & CLASSIFICATION ACCURACY

[9] Studies purporting to validate the forensic use of these tests for lie-detection purposes have reported relatively high levels of discriminant validity.[31] For example, proponents of the ODT test claim to have success rates as high as 85% in separating the guilty from the innocent.[32] These test results, however, depend on a series of methodological assumptions, the failure of any one of which renders their conclusions invalid.[33]

A. Correlations & A Reverse Inference

[10] The first and most fundamental validity concern with these new purported lie-detection tests is the same correlation/causation dilemma that existed with other, earlier purported lie-detector tests. Like all lie-detection techniques currently on offer (polygraphy, EEG, functional Magnetic Resonance Imaging ("fMRI")[34], these techniques do not directly measure deception, but rather measure indirect observable phenomena thought to *correlate with* deception.[35]

[11] The underlying assumption of all existing lie-detection techniques is that a hidden mental state (deception) has a detectable physiological counterpart.[36] The proposed correlations between the observed pupil

[31] *See* Kircher & Raskin, *supra* note 9, at 164.

[32] *See id.* at 166; *see also* Cook et al., *supra* note 15.

[33] *See id.* at 160.

[34] FMRI is a hemodynamic measure. It detects a blood oxygen level dependent signal, which reflects a delayed, secondary consequence of neural activity. *See* LUCK, *supra* note 22, at 12–13, 24. For this reason, ERP proponents claim that it is a more direct measure of the underlying neural activity than fMRI. *See id.*

[35] Kircher & Raskin, *supra* note 9, at 165.

[36] *See* Kedar Nath Sahu et.al., *Radar Based Lie Detection Technique*, 14 GLOBAL J. OF RESEARCHES IN ENGINEERING: F ELECTRICAL AND ELECTRONICS ENGINEERING 1 (2014).

widening or the changed electrical activity in the brain, and the unobservable hypothesized deception, are based on aggregate data about eye and brain behavior in test subjects, and there is a missing, invisible step in the purported causal chain.[37] The theory behind all lie detection follows the same logic: lie → unobserved (hypothesized) mental process → observed physiological reflex that appears, in the aggregate, to correlate with lying.[38] The assumption is that measured differences in the observed physiological reflex (the target variables) between two behavioral conditions (truth telling and lying) are identical except for the controlled variable (truth versus lie).[39] This in turn leads to a reverse inference: the observed physiological reflex, therefore, is evidence of the unobservable mental process of formulating the lie.[40] In laboratory studies of these measuring techniques, the postulated mental state (deception/concealment) is a dependent variable that is measured by manipulation of a task (directed lying).[41] The validation experiments typically have the same basic design: the observable phenomena, like pupil dilation and electrical activity in the brain, are recorded during different experimental conditions (*e.g.* truth telling and lying or revelation and concealment of intimate knowledge of a guilty subject matter).[42] Researchers ask study subjects to answer some questions truthfully and others dishonestly, and then evaluate the measurable phenomenon (pupil size, brain activity) in each condition, in order to identify the physiological conditions that correlate with lying.[43] The responses are then evaluated in contrasting terms across those conditions;

[37] *See id*. at 10.

[38] *See* Jerome H. Skolnick, *Scientific Theory and Scientific Evidence: An Analysis of Lie-Detection* 70 YALE L.J. 694, 699–700 (1961).

[39] *See id*. at 705.

[40] *See id*. at 727.

[41] NORMAN ANSLEY, LEGAL ADMISSIBILITY OF THE POLYGRAPH 17 (1973).

[42] *See* NATIONAL RESEARCH COUNCIL, THE POLYGRAPH AND LIE DETECTION 76 (2003).

[43] *See id.* at 71.

larger pupils when lying; smaller pupils when telling the truth.[44] The reverse inference that is drawn from these aggregate studies (*e.g.* that an individual's pupils are larger when lying than when telling the truth) is only deductively valid if the change in pupil size *only* occurs when the individual is lying, but this fundamental assumption has not only never been tested, it cannot be.[45]

[12] Put in simple terms, even if one accepts the most optimistic estimates of ODT and ERP accuracy, the detected physiological state *may* indicate deception, it does not *necessarily* do so.[46] There is no way to disambiguate false-positive results during test administration and determine conclusively whether the physiological condition being observed (larger pupil size, changed brain activity) is actually evidence of deception, rather than some distinct mental process that also correlates with the observed change in condition.[47]

[13] Furthermore, the experimental conditions themselves can vary in infinite ways. For example, lies vary by the significance of the consequence of their detection; the amount of the lie (entirely false versus partially false); the level of concealment (answering "no" to a question whose truthful answer is "yes," as opposed to failing to volunteer guilty knowledge when not directly asked for it); whether the trigger itself is truthful, deceptive, accurate, or misleading.[48] There are also likely variations across test

[44] *See id.* at 70.

[45] *See* David C. Raskin et al., CREDIBILITY ASSESSMENT: SCIENTIFIC RESEARCH AND APPLICATIONS 169 (1st ed. 2014).

[46] *Cf.* Dov Fox, *The Right to Silence Protects Mental Control*, 42 AKRON L. REV. 763, 764 (2009) (arguing that a correlation between deception and prefrontal-cortex activation does not permit a reverse inference that all such activation indicates deception).

[47] *See* Cook et al., *supra* note 15.

[48] This latter variation is particularly salient in the context of police-conducted interrogations, in which a commonly trained and practiced technique is to "catch" liars by tripping them up with counter-lies. *See* Gohara, *infra* note 108; *cf.* Griffin, *infra* note 50,

administration in terms of the expectancy of the administrator, particularly if that expectancy is communicated, even implicitly, to the test subject through tone of voice, question wording, facial expressions, or responses to answers.[49] If different studies (or the same study across different subjects) use experimental "lies" and "truths" that vary in terms of their mental effort or moral significance for the study subjects, or are administered by administrators giving different feedback to the subjects, the difference in those tasks might correlate with different, confounding mental processes, creating a mismatch between test conditions and results that would frustrate the physiological deception mapping that these tests purport to accomplish. Similarly, a mismatch in these same experimental conditions across the experimental and control groups of an individual study could also produce dangerously unreliable and misleading results.

[14] This mismatch danger is particularly salient in the context of the detection of "lies," as opposed to other mental-process correlates, since the entire theory of deception detection depends on an underlying assumption that there is such a thing as an objectively classifiable, capital-L "lie," an assumption that is belied by an entire school of modern philosophical theory.[50]

---

at 1517 (discussing the variable nature of social norms regarding "self-protective perjury").

[49] *Cf.* Par Anders Granhag & Aldhert Vrig, *Deception Detection,* PSYCHOLOGY AND LAW: AN EMPIRICAL PERSPECTIVE 43, 65 (Neil Brewer & Kipling D. Williams eds., 2005) ("[L]ie catchers who attributed positive trait characteristics (dispositional) to the person they were judging also tended to judge this person as truthful in a given situation (state).").

[50] *See* Lisa Kern Griffin, *Criminal Lying, Prosecutorial Power, and Social Meaning*, 97 CAL. L. REV. 1515, 1517 (2009) (noting "the lack of consensus among moral philosophers about defensive falsehoods that merely mislead"); *see, e.g.*, MANUEL GARCÍA-CARPINTERO & MAX KÖLBEL, RELATIVE TRUTH 1 (2008).

[15] These studies artificially construct a binary environment in which there are only two options: lying or telling the truth.[51] In reality, there is a great deal of real estate that falls between these two poles. One can imagine a criminal investigation or trial that seeks to resolve whether an individual killed another person in self-defense. The hypothetical suspect/defendant claims that the victim was imminently about to use deadly force, so the killing was justified. It is well documented, in other contexts, that individuals faced with threats tend to overestimate the severity of those threats.[52] For example, eyewitnesses tend to overestimate the length of traumatic criminal episodes and the size of weapons when recounting their perception of violent crimes to which they were witnesses.[53] These witnesses are neither lying nor telling the truth. They are subjectively sincerely, but objectively inaccurately, misremembering the events that they saw, and are recounting them in a way that is driven by underlying cognitive biases, which can include racial stereotyping.[54] The entire concept of mistaken self-defense is built upon the premise that an individual could

[51] *See* Vaughan Bell, *Vaughan Bell: The Truth About Lie Detectors*, THE GUARDIAN (Apr. 21, 2012), https://www.theguardian.com/science/2012/apr/22/lie-detector-fallibility-criminal-psychology, https://perma.cc/ERG9-NFHQ.

[52] *See* Paul Leinwand & Cesare Mainardi, *The Fear of Disruption Can Be More Damaging than Actual Disruption,* STRATEGY & BUSINESS (Sep. 27, 2017), https://www.strategy-business.com/article/The-Fear-of-Disruption-Can-Be-More-Damaging-than-Actual-Disruption?gko=b4a17, https://perma.cc/9LNG-K6AL.

[53] *See, e.g.*, Saul M. Kassin & Lawrence S. Wrightsman, THE AMERICAN JURY ON TRIAL 82 (1988); *see also* John C. Brigham & Robert K. Bothwell, *The Ability of Prospective Jurors to Estimate the Accuracy of Eyewitness Identification*, 7 L. & HUM. BEHAV. 19, 21 (1983); Robert J. Hallisey, *Experts on Eyewitness Testimony in Court – a Short Historical Perspective*, 39 HOWARD L.J. 237, 256 (1995). The effect is greater when the witness and the perpetrator are of different races. *See* ELIZABETH F. LOFTUS, EYEWITNESS TESTIMONY 136–39 (1979).

[54] *See generally* Loftus*, supra* note 53, at 21–22, 36–51 (explaining the selective and malleable nature of human memory and that cognitive biases can color a witness's memory of an event); BARRY SCHECK, ET AL., ACTUAL INNOCENCE 55 (2003) ("What happens in front of the eyes is transformed inside the head, and is refined, revisited, restored, and embellished in a process as perpetual as life itself.").

genuinely, but mistakenly, believe that the use of deadly force was necessary in self-defense.[55] Furthermore, the definition of a "lie" or of "guilty knowledge" is itself socially constructed and varies inter-culturally.[56] Even a study that dealt with a very large group of subjects (large enough to eliminate the likelihood of sampling error) would still rely on an objective identification of what a "lie" is.[57] This identification seems particularly unachievable in a forensic application, since credibility disputes in criminal investigations and prosecutions typically arise in situations in which there is no conclusive extrinsic evidence of "truth."[58] The recent

[55] *See, e.g.*, In re: Christian S., 872 P.2d 574, 575 (Cal. 1994) ("Under the doctrine of imperfect self-defense, when the trier of fact finds that a defendant killed another person because the defendant *actually* but unreasonably believed he was in imminent danger of death or great bodily injury, the defendant is deemed to have acted without malice and thus can be convicted of no crime greater than voluntary manslaughter.") (emphasis in original); The trial of George Zimmerman is a good example of this phenomenon. Many believe that his killing of Trayvon Martin was driven, at least in part, by racial stereotyping. *See, e.g.,* Tom Foreman, *Analysis: The Race Factor in George Zimmerman's Trial*, CNN (July 15, 2013), http://www.cnn.com/2013/07/14/justice/zimmerman-race-factor/index.html, https://perma.cc/J935-3CHB (last visited Sept. 11, 2017). If this is true, that fact would make his subjective claim to self-defense more, rather than less, "true." Racial stereotypes would have caused him to overestimate the threat that Martin posed creating a genuine but mistaken belief that his life was in danger. His description of that would be neither a lie nor the truth, but rather a warped perception of reality.

[56] *See* Joseph W. Rand, *The Demeanor Gap: Race, Lie Detection, and the Jury*, 33 CONN L. REV. 1, 4 (2000) ("[P]eople of different races might have contradictory cultural cues to indicate deception, such that an African-American witness might attempt to connote sincerity through indirect, non-assertive speech patterns, but the Caucasian juror might misread these cues to indicate dissembling."). *See, e.g.*, Sara Bernal, *Bullshit and Personality,* BULLSHIT AND PHILOSOPHY 82 (Gary L. Hardcastle & George A. Reisch eds., 2006) (describing the salient differences between "bullshitting" and outright deception).

[57] *See generally* United States v. Semrau, 693 F.3d 510, 521–22 (6th Cir. 2012) (holding that expert testimony regarding lie detection through fMRI still lacked general acceptance).

[58] *See, e.g.*, Saul M. Kassin et al., *Confessions That Corrupt: Evidence from the DNA Exoneration Case Files*, 23 PSYCHOL. SCI. 41, 42, 44 (2012); Saul M. Kassin et al., *"I'd*

spate of DNA exonerations in the United States is compelling circumstantial evidence that the criminal-justice system is not very accurate at distinguishing truth from lies in critical situations,[59] so verifying, even in hindsight, whether an individual was "lying" (as opposed to telling the truth but being extraordinarily unlucky with regard to circumstances) would be an educated guess, at best.

[16] Even if one accepts that a lie can be meaningfully defined and that there is a well-established *association* between lying and the physiological reflex being measured (pupil size or neurotransmitter-mediated neural activity), the nature of the association – the brain mechanism of deception that is hypothesized to underlie the observed characteristic – is not fully understood, and so there is no way to test the causal deduction from the observed correlation.[60]

[17] In other words, the observable phenomenon (pupil dilation) is a surrogate for a theoretical underlying brain mechanism and, like with other observed physiological phenomena thought to correlate with deception, the fundamental question of *why* the pupil seems to open more when people lie remains unanswered.[61] Because of this, researchers cannot validate any causal connection between the observed characteristics and the underlying

Know a False Confession If I Saw One": A Comparative Study of College Students and Police Investigators, 29 L. & HUMAN. BEHAV. 211, 221–22 (2005). *See generally* Steven A. Drizin & Richard A. Leo, *The Problem of False Confessions in the Post-DNA World*, 82 N.C. L. REV. 891, 948 (2004) (discussing credibility disputes in criminal proceedings where suspect questioning was not memorialized in a form jury can perceive).

[59] *See* sources cited *supra* note 58.

[60] *See* Dov Fox, *supra* note 46.

[61] *See* Adina L. Roskies, *Neuroimaging and Inferential Distance*, 1 NEUROETHICS 19, 24 (Feb. 7, 2008), http://www.dartmouth.edu/~adinar/CV_files/neuroethics%20inferential%20distance.pdf, https://perma.cc/9E45-UAUX.

mental state or process that they are thought to represent.[62] One simply cannot know whether the mental state or process (lying or withholding or concealing guilty knowledge) is a necessary determinant of the observed phenomenon associated with it solely through laboratory testing that establishes the association.[63] As one proponent of ERP noted: "ERPology experiments do not directly tell us anything important about the mind or brain . . . ."[64] Because this intermediate step remains a mystery, there is no way to control for any residual confounding variables with which it might correlate, other than deception.[65]

[18] Obvious examples of potential confounding variables are stress, embarrassment, caution, deliberation, and autobiographical memory.[66] For example, if lying causes the pupil to dilate because lying is stressful and stress causes pupil dilation, then the ODT accomplishes nothing different than the polygraph.

[19] Another obvious potential confounding variable is circumspection.[67] FMRI research suggests that increased brain activity occurs during deception because truth is a natural reflex and it takes

[62] *See* Henry T. Greely, *Neuroscience, Mindreading, and the Courts: The Example of Pain*, 18 J. HEALTH CARE L. & POL'Y 171, 195 (2015).

[63] *See* Teneille Brown & Emily Murphy, *Through a Scanner Darkly: Functional Neuroimaging as Evidence of a Criminal Defendant's Past Mental States*, 62 STAN. L. REV. 1119, 1162 (2010).

[64] Luck, *supra* note 22, at 5.

[65] *See* Roskies, s*upra* note 61, at 24.

[66] *See id.*

[67] *Cf.* Scott T. Grafton et al., *Brain Scans Go Legal*, 17 SCI. AMER. MIND 30, 36–37 (2006) (arguing that the correlation between prefrontal-cortex activation and deception does not make such activation in individuals a sufficiently reliable of their veracity as witnesses at trial).

additional effort to "suppress" the truth.[68] However, it also takes additional effort to carefully consider and word a truthful and persuasive answer, a function that one would expect any critical witness, under oath during a police interview or while testifying at trial, to perform.[69] In other words, even a truthful witness will not blurt out the truth in a grave circumstance in which an answer deemed untruthful could have significant consequences, but rather the witness will cautiously answer questions when being interviewed by the police, deposed by a lawyer, or cross-examined by an adverse party's attorney. So, even if one accepts the "cognitive effort" theory, deception is not the only verbal communication process that requires more cognitive effort and inhibition than hastily blurting out truthful information.[70] This generates a form of selection bias called non-independence (or circularity) error.[71] Non-independence is a particular problem in neuroscience-based tests like the ERP.[72] Non-independence occurs when a test uses an input that biases later results toward significance, creating and measuring data that contain less relevant signal information than test results seem to indicate.[73] Non-independence produces distorted correlation patterns.[74] This leads to false-positive test results: meaningless

[68] *See* D. D. Langleben et al., *Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study*, 15 NEUROIMAGE 727 (2002).

[69] *See* Grafton et al., *supra* note 67, at 36.

[70] *See id.*

[71] *See* Edward Vul et al., *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition,* 4 PERSP. ON PSYCHOL. SCI. 274, 279 (2009).

[72] *See id.* at 280.

[73] *See* Edward Vul & Nancy Kanwisher, *Begging the Question: The Nonindependence Error in fMRI Data Analysis,* FOUNDATIONAL ISSUES IN HUMAN BRAIN MAPPING 71, 72–73 (Stephen Jose Hansen & Martin Bunzl, eds., 2010).

[74] *See* Vul et al., *supra* note 71, at 281.

effect sizes (correlations) that create unreliable conclusions out of nothing, "findings" from noise.[75]

B. Sensitivity & Specificity

[20] Even assuming that increased pupil size or heightened brain activity indicate deception or concealment of guilty knowledge, researchers have not yet validated objective, standardized classification rules governing how large a pupil must be, or how much additional brain activity there must be, to reach a reliable conclusion that a particular individual is not being truthful – *i.e.,* the sensitivity and specificity of the purported physiological measurements in detecting deception.[76] Sensitivity and specificity are essentially measurements of accuracy – two crucial types of error rates, in *Daubert* terminology.[77] In a clinical context, "sensitivity" describes how often the test correctly identifies an individual who is lying as a proportion of overall liars to whom the test is administered.[78] "Specificity" describes how often the test exonerates one who is telling the truth as a proportion of overall truth-tellers to whom the test is administered.[79] A perfect test would have perfect sensitivity and specificity (100%).[80] The lower a test's sensitivity and specificity, the more likely it is to misclassify.[81] False

[75] *See generally id.* at 281 (discussing implications of non-independent analysis on correlation).

[76] *See* Langleben & Moriarty, *supra* note 3, at 12–14.

[77] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 594 (1993).

[78] *See* Christopher P. Guzelian et al., *A Quantitative Methodology for Determining the Need for Exposure-Prompted Medical Monitoring*, 79 IND. L.J. 57, 81–82 (2004).

[79] *See id.*

[80] *See* Richard L. Elliott, *Neuropsychiatry in the Courtroom*, 62 MERCER L. REV. 933, 943 (2011); Guzelian et. al., *supra* note 78, at 81.

[81] *See* Elliott, *supra* note 80, at 942.

positives occur because of innocent but undetected confounding variables (*e.g.* stress or introspection) in a test with low specificity.[82]

[21] Ordinarily, in order to determine the specificity and sensitivity of a new test, researchers compare the positive and negative results of the new test to a known clinical standard.[83] A researcher can only estimate the sensitivity and specificity of a test with an answer key – if the truth or lying of the test subject can be determined conclusively in another way.[84] The problem, in the context of lie detection, is that there is no clinical standard for lying or telling the truth. Because the researchers who seek to validate these new tests cannot compare them to a known, perfect standard (a perfectly accurate pre-existing test of deception), test administrators instead have to utilize subjective cutoffs – unscientific judgment calls about the line between false and true positives (or false and true negatives).[85]

---

[82] *See* Jennifer Vogel & Madeleine Baran, *Inconclusive: The Truth About Lie Detector Tests*, APM REPORTS (Sept. 10, 2016) ("Littlefield found that a person's body can trigger similar test results when undergoing "stressful truth telling" as when lying), https://www.apmreports.org/story/2016/09/20/inconclusive-lie-detector-tests; *see also* Robert Steinbrook, *The Polygraph Test — A Flawed Diagnostic Method*, 322 NEW ENG. J. MED. 2, 122–23 (1992), http://www.nejm.org/doi/pdf/10.1056/NEJM199207093270212, https://perma.cc/63RN-LVUY.

[83] *See* NAT'L RESEARCH COUNCIL, THE POLYGRAPH AND LIE DETECTION 66 (The Nat'l Acad. Press, 2003) https://www.nap.edu/download/10420#, https://perma.cc/4M4Q-DWKN.

[84] *Compare* AMERICAN PSYCHOLOGICAL ASSOCIATION, *The Truth About Lie Detectors (aka Polygraph Tests)* (explaining the methods used to determine true and false responses in lie detector tests and further discusses the scientific shortfalls of these methods), http://www.apa.org/research/action/polygraph.aspx (last visited Oct. 25, 2017); *and* Christopher J. Mattocks et al., *A Standardized Framework for the Validation and Verification of Clinical Molecular Genetic Tests*, 18 Eur. J. Hum. Genetics 1276–288 (2010) (explaining the implementation of an elaborate framework for validation of a scientific testing technique).

[85] *See generally* Aleksandra Slavkovic, *Evaluating Polygraph Data*, CARNEGIE MELLON UNIV. (discussing the standards under which polygraphs are analyzed and drawing conclusions regarding the problems with current standards of analysis),

[22] The potential for variation in the way that these judgment calls are made increases the potential variability of the test results and undercuts the validity and reliability of these techniques as conclusive tests of deception. In the absence of a clinical standard for comparison, sensitivity and specificity estimates will be biased, either in favor or against detection (toward false positives or false negatives), and the direction of their bias is usually undetectable.[86]

C. Algorithms

[23] Both ODT and ERP tests rely on computers to measure and interpret test results.[87] While this automation removes some of the human subjectivity, bias, and examiner error of other purported lie-detection tests (like interpreting facial cues and eye movements), it is not without its human judgment calls.[88] In the case of automated measurements like ODT and

http://www.stat.cmu.edu/tr/tr766/tr766.pdf, https://perma.cc/HT7Y-5GQP (last visited Nov. 7, 2017).

[86] *See* Ewout H. Meijer et. al., *Deception Detection with Behavioral, Autonomic, and Neural Measures: Conceptual and Methodological Considerations That Warrant Modesty*, 53 PSYCHOPHYSIOLOGY 593, 594, 601 (2016); Moi Hoon Yap et. al., *Facial Behavioral Analysis: A Case Study in Deception Detection*, 4 BRIT. J. APPLIED SCI. & TECH. 1485, 1486, 1489, 1492, 1494 (Feb. 5, 2014); Sherry H. Stewart et. al., *Anxiety Sensitivity and Negative Interpretation Biases: Their Shared and Unique Associations with Anxiety Symptoms*, 34 J. PSYCHOPATHOLAGY & BEHAV. ASSESSMENT 332, 334–35 (Apr. 14, 2012).

[87] *See, e.g.*, Sabrine Windmann et. al., *Cognitive and Neural Mechanisms of Decision Biases in Recognition Memory*, 12 CEREBRAL CORTEX 808, 810–11 (Aug. 2002); *see, e.g.*, Cook et. al., *supra* note 15.

[88] *See* Hannah Devlin, *Discrimination by Algorithm: Scientists Devise Test to Detect AI Bias*, THE GUARDIAN (Dec. 19, 2016), https://www.theguardian.com/technology/2016/dec/19/discrimination-by-algorithm-scientists-devise-test-to-detect-ai-bias, https://perma.cc/EDT9-3AML; Jesse Emspak, *How a Machine Learns Prejudice,* SCI. AM. (Dec. 29, 2016), https://www.scientificamerican.com/article/how-a-machine-learns-prejudice/, https://perma.cc/PGD9-A5PB; Windmann et al., *supra* note 87; Cook et al., *supra* note 15.

ERP, the human subjectivity comes into play at an earlier moment in time, when a human being writes the algorithms that dictate the computer's interpretation of results – for example, when a programmer determines whether an observed physiological change is a "real" change or an artifact (sensitivity).[89] The programming process is itself, therefore, is subjective and subject to human bias and error.[90] Perhaps more concerning, subjectivity in algorithm programming is harder to detect and less likely to be discovered than human error in test administration, particularly in light of the proprietary nature of forensic software.[91]

---

[89] *See* Windmann et. al., *supra* note 87, at 815; *see* Devlin, *supra* note 88.

[90] *See* Guo-Zhu Wen & De-Shuang Huang, *A Novel Spike Sorting Method Based on Semi-Supervised Learning*, ADVANCED INTELLIGENT COMPUTING THEORIES AND APPLICATIONS 605, 606 (De-Shuang Huang et al., eds., 2008); Matt Burgess, *Holding AI to Account: Will Algorithms Ever Be Free from Bias if They're Created by Humans?*, WIRED (Jan. 11, 2016), http://www.wired.co.uk/article/creating-transparent-ai-algorithms-machine-learning, https://perma.cc/F6JB-DE6S; *cf.* Elena Rusconi & Timothy Mitchener-Nissen, *Prospects of Functional Magnetic Resonance Imaging as Lie Detector*, 7 FRONTIERS IN HUMAN NEUROSCIENCE 22, 27 (2013) (noting the subjectivity inherent in fMRI-analysis algorithms); Lauren Kirchner, *When Discrimination Is Baked into Algorithms*, THE ATLANTIC (Sept. 6, 2015), https://www.theatlantic.com/business/archive/2015/09/discrimination-algorithms-disparate-impact/403969/, https://perma.cc/Z83D-Z36X; Windmann, et al., *supra* note 87.

[91] *See* Lydia Pallas Loren & Andy Johnson-Laird, *Computer Software-Related Litigation: Discovery and the Overly-Protective Order*, 6 FED. CTS. L. REV. 1, 32–33 (2012); Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 1976–78 (2017); Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245, 1250 (2016); Cory Altheide & Christa M. Miller, *Validating Proprietary Digital Forensic Tools: A Case for Open Source*, FORENSIC MAG. (Dec. 13, 2011), https://www.forensicmag.com/article/2011/12/validating-proprietary-digital-forensic-tools-case-open-source, https://perma.cc/C7UR-CUCV; Rebecca Wexler, *Convicted by Code*, SLATE (Oct. 6, 2015), http://www.slate.com/blogs/future_tense/2015/10/06/defendants_should_be_able_to_inspect_software_code_used_in_forensics.html, https://perma.cc/SX53-VJYY; Rebecca Wexler, *When a Computer Program Keeps You in Jail*, N.Y. TIMES (June 13, 2017), https://www.nytimes.com/2017/06/13/opinion/how-computers-are-harming-criminal-justice.html, https://perma.cc/FV8G-S93F; *see, e.g.*, Michael Harrington, *A Methodology for Digital Forensics*, 7 T.M. COOLEY J. OF PRAC. & CLINICAL L. 71, 72, 74–75 (2004);

D. Ecological Validity

[24] The studies that purport to validate these new lie-detection techniques involve artificial, mock-crime scenarios in small-scale experimental research studies.[92] Under laboratory study conditions, ODT subjects are given a questionnaire about their assigned guilt or innocence of mock crimes (academic fraud and falsified drivers' licenses) while the eye tracker records their eye measurements and runs an algorithm to determine if they are being truthful or deceptive while answering the controlled test questions.[93]

[25] To achieve external validity, which is critical to evidentiary reliability, experimental results need to be replicable under real-world conditions.[94] Translating successful laboratory protocols into the criminal-justice system, however, is notoriously difficult.[95]

Lauren Kirchner, *Negligent DNA Testing Has Affected Thousands of New York Criminal Cases*, PAC. STANDARD (Sept. 7, 2017), https://psmag.com/news/negligent-forensic-dna-testing-has-affected-thousands-of-ny-criminal-cases, https://perma.cc/2CVS-HKMT; Lauren Kirchner*, Sentenced by an Algorithm: Where Traditional DNA Testing Fails*, *New Technology Takes Over*, PAC. STANDARD (Nov. 9, 2016), https://psmag.com/news/sentenced-by-an-algorithm-where-traditional-dna-testing-fails-new-technology-takes-over, https://perma.cc/7KWS-TVAV.

[92] *See* Daniel D. Langleben & Jane Campbell Moriarty, *Using Brain Imaging for Lie Detection: Where Science, Law and Research Policy Collide*, 19 PSYCHOL. PUB. POL'Y L. 222, 222–234 (2013).

[93] *See* Tariff Classification Ruling, *supra* note 17, at *1.

[94] *See* Langleben & Moriarty, *supra* note 92, at 14.

[95] *See* Rusconi & Mitchener-Nissen, *supra* note 90 ("The propriety of equating simulated scientific testing with real life scenarios for the purpose of evidence is highly questionable.").

[26] Test subjects are recruited from the university community, comprise mostly college students, and those assigned to lie during the tests were following the researchers' instructions in doing so.[96] One would expect these test subjects to have a different mix of socio-economic, educational, racial, and cultural diversity than criminal-justice-involved individuals. The developers of the ODT technique concede that it is ineffective with "poor readers,"[97] a group that is over-represented in the criminal-justice system. However, a more serious application problem arises from the fact that the test subjects in this artificially constructed context lack the significant situational consequences that motivate actual legal-system participants to deceive (or *beat* the test), which has been shown, in other contexts, to frustrate the generalizability of classification results.[98] Test subjects have a different risk/benefit ratio in the decision whether to lie than individuals involved in a real-world criminal investigation or prosecution.[99] Test subjects lie because they are instructed to do so, and the studies are designed with the assumption that the subjects will follow these directions, largely because they have no reason not to in an experimental setting.[100]

[27] The "administrators" of the tests are likely to be different in the real world as well, which gives rise to significant differences in both questioner

---

[96] *See* Kircher & Raskin, *supra* note 9, at 161; Travis L. Seymour, et al*., Combining Blink, Pupil, and Response Time Measures in a Concealed Knowledge Test,* FRONTIERS IN PSYCHOL., Feb. 4, 2013, at 3.

[97] Kircher & Raskin, *supra* note 9, at 169.

[98] *See* Maureen O'Sullivan et al*., Police Lie Detection Accuracy: the Effect of Lie Scenario*, 33 L. & HUM. BEHAV. 530, 531 (2009) ("[H]igh stakes deception scenarios – in which liars face significant consequences for getting caught in their lies, and significant benefits for getting away with them – are important for experimental realism and provide the kinds of relevant deception clues necessary for accurate lie detection.") (citations omitted).

[99] *See id.* at 532.

[100] *See* S.A. Spence, *et al., Behavioural and Functional Anatomical Correlates of Deception in Humans,* 12 NEUROREPORT 2849 (2001).

expectancy and subject effects in comparison to the research studies purporting to validate these techniques as lie detectors.[101] Police personnel typically administer polygraph examinations.[102] One would expect the same at least for the ODT test. Police in the United States generally use Reid-type interrogation techniques that include confrontational tactics like evidence ploys and repeated insistence on the subject's guilt as tools to increase the likelihood of obtaining confessions.[103] This type of interrogation environment has very different stakes for its participants than the laboratory environment in which a researcher has endorsed the lie and, in the process, conveyed to the subject no stake in whether the subject is telling the truth or lying other than scientific interest in the results of the test.[104] One would expect the real-world expectancy effects of actual police interrogation to significantly alter the underlying mental processes that connect the observed physiological response with its hypothesized correlated deception.[105]

---

[101] *See* O'Sullivan et al., *supra* note 98.

[102] *Id*.

[103] *See* Allison D. Redlich et al*., The Police Interrogation of Children and Adolescents*, INTERROGATIONS, CONFESSIONS AND ENTRAPMENT 107, 109 (2004), http://www.albany.edu/scj/documents/Chapter05Lassiter.pdf, https://perma.cc/SSX2-UTHP (explaining how current law permits the police to use "trickery and deception," including telling suspects that they possess inculpatory evidence that they do not, to obtain confessions); Miriam S. Gohara, *A Lie for a Lie: False Confessions and the Case for Reconsidering the Legality of Deceptive Interrogation Techniques*, 33 FORDHAM URBAN L.J. 100, 117–18 (2006).

[104] *See* John R.P. French, Jr. & Bertram Raven, *The Bases of Social Power*, STUDIES IN SOCIAL POWER, 259, 267 (1959), http://www.communicationcache.com/uploads/1/0/8/8/10887248/the_bases_of_social_power_-_chapter_20_-_1959.pdf, https://perma.cc/UR2W-5A56 (explaining how the social power of coercion can foster the fear that disobedience will be punished).

[105] *See* Kamila E. Sip et al*., The Production and Detection of Deception in an Interactive Game,* 48 NEUROPSYCHOLOGIA 3619 (2010), http://www.sciencedirect.com/science/article/pii/S0028393210003672?via%3Dihub, https://perma.cc/ARK7-JFP5 (noting that that deception is a complex act that cannot be exclusively associated with telling a falsehood and that it is facilitated by hierarchical

E. Countermeasures

[28] Older, more established lie-detection methods have been proven to be susceptible to countermeasures.[106] At present, however, the vulnerability of the ODT and ERP tests to countermeasures is unknown.[107]

[29] Because the underlying causal mechanism linking the observable physiological phenomenon with a deceptive mental state is unknown, countermeasures that seek to de-link the first step (lying) from the final step (pupil dilation or brain activity) would target the unknown intermediate variable.[108] If the "cognitive effort" theory of neuroscience is correct, then

decision-making and risk evaluation); *see also* NRC Report, *supra* note 2 (documenting how variations in the setting, test administrator, and question format affects the accuracy of the polygraph); *see generally* Janice Nadler, *No Need to Shout: Bus Sweeps and the Psychology of Coercion*, 2002 SUP. CT. REV. 153, 168–97 (2002) (explaining the actor-observer bias, pursuant to which, observers like police officers tend to considerably overestimate the voluntariness of others' actions and underestimate the effect of politeness rules for expressing and understanding commands stated as "requests," the coercive effect of narrowing personal space, deference to status, and reduced deliberation under time pressure); Andrew E. Taslitz, *Racial Profiling, Terrorism, and Time*, 109 PENN ST. L. REV. 1181, 1181–96 (2005) (discussing the actor-observer bias); Kamila E. Sip et al*., What if I Get Busted? Deception, Choice, and Decision-Making in Social Interaction*, 6 FRONTIERS IN NEUROSCI. 58 (2012) (finding that an individual's decision to lie is more affected by the potential risk of social confrontation than by the claim itself).

[106] *See* G. Ganis et al*., Lying in the Scanner: Covert Countermeasures Disrupt Detection by Functional Magnetic Resonance Imaging*, 55 NEUROIMAGE 312 (2011) (demonstrating the effectiveness of task-tailored countermeasures in "fooling" an fMRI); Matt Zapotosky, *Indiana Man Accused of Teaching People to Beat Lie Detector Tests Faces Prison Time*, WASH. POST (Aug. 31, 2013), https://www.washingtonpost.com/local/indiana-man-accused-of-teaching-people-to-beat-lie-detector-tests-faces-prison-time/2013/08/31/a7cbe74a-08ea-11e3-9941-6711ed662e71_story.html?utm_term=.77bfaaea3034, https://perma.cc/6WAE-3FWE.

[107] *See* Ganis, *supra* note 111, at 312.

[108] *See* Terri Patterson, The Effect of Cognitive Load on Deception, 3 (Oct. 2, 2009) (unpublished Ph.D. dissertation, Florida International University),

an individual could defeat these tests, either by increasing their cognitive effort in the initial (truthful) calibration questions (perhaps by performing mathematical calculations or attempting to recall a long, memorized list of items during baseline questioning), or decreasing their cognitive effort while lying during test administration (perhaps by memorizing and rehearsing a lie repeatedly before an ODT or ERP interview).[109] If individuals are capable of manipulating either their baseline physiology or their physiological response during deceptive behavior, then these tests lack reliability in application. Unfortunately, there have been few attempts to study the effectiveness of countermeasures in defeating the ODT and ERP tests, despite the fact that countermeasures are one of the reasons courts are resistant to admit the results of polygraph and fMRI tests.[110] The little literature that exists on the subject suggests that countermeasures are effective in defeating the ODT test.[111]

PART III. GENERALIZATION

[30] These validity concerns in turn give rise to a related set of concerns regarding the reliability of generalizing from these group-level studies to the brain function of any given individual whose mental state is relevant to a criminal investigation or prosecution.[112] In developing these techniques, researchers study large groups of subjects to determine "normal" ranges of response to various sensory and cognitive stimuli, most pertinently for the

http://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=1174&context=etd, https://perma.cc/8DYW-67UP.

[109] *See id.* at 31.

[110] *The Lie Behind Lie Detectors*, WIRED, (Mar. 15, 2006), https://www.wired.com/2006/03/the-lie-behind-lie-detectors/, https://perma.cc/9XNB-6PVN.

[111] *See* Seymour, *supra* note 96, at 14.

[112] *See* Matthias Gamer & Wolfgang Ambach, *Deception Research Today*, 5 FRONTIERS IN PSYCHOL. 1, 1–2 (2014).

criminal justice system: truth, concealment, and affirmative misrepresentation. To generate the readings for "truth" and "lie," they average these research data sets to generate group-level results.[113] For example, the theory of ODT is based on the finding that the average pupil size of an individual who is lying is larger than that of that same individual that is telling the truth because human pupils are generally smaller during deception than during truth telling.[114]

A. Extrapolating from the Aggregate

[31] Validation studies of ODT and ERP tests are based on statistical averages across large groups of test subjects.[115]The theory behind the validation studies is that the phenomenon that correlates with lying, like pupil size and changes in brain activity, across a large number of test subjects, captures universal similarities with "lying," rather than singular physiological reflexes associated with individual instances of lying during the experiments.[116] The group-to-group nature of these studies gives rise to a classic extrapolation problem that exists between laboratory and field research and making predictions about deception in real-world application in the criminal-justice system.[117] In order to validate the hypothesis that a

[113] *See* LUCK, *supra* note 22, at 24 ("In most ERP experiments, an averaged ERP waveform is constructed at each electrode site for each subject in each condition." The amplitude or latency of a component of interest is then measured in each one of these waveforms, and these measured waveforms are then entered into a statistical analysis just like any other variable.").

[114] *See* Seymour, *supra* note 96, at 2.

[115] *See* Pooja Patnaik et al., *Generalizability of an Ocular-Motor Test for Deception to a Mexican Population*, 6 INT'L J. APPLIED PSYCHOL. 1, 7 (2016); *See* Langleben, *supra* note 3, at 1–2, 17.

[116] *See* LUCK, *supra* note 22.

[117] *See* John A. List & Steven D. Levitt, *What Do Laboratory Experiments Tell Us About the Real World?*, 1, 16 (2005), http://pricetheory.uchicago.edu/levitt/Papers/LevittList2005.pdf, https://perma.cc/5R72-CA75.

given individual was lying or concealing information based on pupil reaction, it would be necessary to determine whether that *particular* individual's pupil reaction indicated deception, which is a different question than the question of whether pupil size generally indicates deception. The group-level results, therefore, even if valid, have no predictive value at the individual level, which would require a differential approach tailored to the particular individual whose veracity was at issue.

B. Outliers

[32] Another classic problem with physiological lie detection has to do with outliers. Even if one could make meaningful generalizations about "normal" individuals based on controlled study in the aggregate, there will always be individual variations in the results of ODT or ERP tests that affect their accuracy with regard to any given test subject.[118] To use ODT as an example, even though there may be an average human pupil size "at rest" and an average amount of dilation that occurs during lying, the amount of variation between the lying pupil and the truth-telling pupil itself would be expected to vary among individuals, along with baseline pupil size.[119] There are also certainly individuals whose pupils react more, less, or differently than the "average" test subject. What this means for the ODT's classification accuracy (the rate at which it can accurately sort the guilty from the innocent) is that an individual with a larger than average pupil and a smaller than average size variation when lying may present as a false

[118] *See* Gary Bargary et al., *Individual Differences in Human Eye Movement: An Oculomotor Signature?*, VISION RESEARCH (2017), http://www.sciencedirect.com/science/article/pii/S0042698917300391, https://perma.cc/D4ZT-VX2V.

[119] *See* Anne E. Cook et al., *Lyin' Eyes: Ocular-Motor Measures of Reading Reveal Deception*, 18 J. OF EXPERIMENTAL PSYCHOLOGY: APPLIED 301, 302 (2012), http://psycnet.apa.org/fulltext/2012-10769-001.pdf, https://perma.cc/L8M4-UDWH.

positive (*i.e.* the ODT is more likely to classify the individual as lying when s/he is telling the truth).[120]

[33] The unproven assumptions that all human brains process lies in the same way, and that all human eyes respond to those lies in the same way, have not only not been validated, but probably could not be. Because these tests all deal in aggregate averages, they are not going to be accurate for people whose pupils or brains behave in ways that deviate significantly from the mean of the tested population.[121]

[34] Conversely, if one or more of the test subjects in a validation experiment is an outlier in terms of pupil reaction or brain activity, then the entire aggregate range will be skewed.[122] This can lead to what statisticians classify as a "Type I error:" "concluding that a difference is real when it was actually a result of a random variation."[123] There is a distinct possibility that the test subjects in the experiments designed to validate these new techniques, collectively, are outliers, significantly narrowing the field to which one could ever hope to generalize about the test findings: they are presumably novice liars, which may not be true of the criminal-justice population.

---

[120] Kyoung Whan Choe et al., *Pupil Size Dynamics During Fixation Impact the Accuracy and Precision of Video-Based Gaze Estimation*, VISION RESEARCH (2015), https://ac.els-cdn.com/S0042698915000024/1-s2.0-S0042698915000024-main.pdf?_tid=71df1a00-bb2c-11e7-98d9-00000aab0f26&acdnat=15091187030254d64b887e7daa46aacbdf2b20d5af, https://perma.cc/85UV-K2HG.

[121] *But cf.* Cook et al.*, supra* note 124 (describing the methods of ODT Testing); BRAINWAVE SCIENCE, http://www.brainwavescience.com, https://perma.cc/VZE8-H7ZM (last visited Aug. 19, 2017) (Describing ERP tests).

[122] *Cf. The Effect of Outliers*, STATISTICS LECTURES (Oct. 25, 2017), http://www.statisticslectures.com/topics/outliereffects, https://perma.cc/F3ZF-PF5Z.

[123] *See* LUCK, *supra* note 22, at 24.

[35] In addition to subjects (either test subjects or real-world applicants of the techniques) who are outliers in terms of the relationship between their observed responses to these tests and their underlying mental state, there are also outliers with regard to the dependent variable (the unobserved mental correlate) for whom these tests, even if accurate for many, would generate inaccurate results.[124] Many mental disorders, such as developmental disabilities, autism, psychopathy, personality disorders, and delusional mental illnesses like schizophrenia, brain lesions, affect the afflicted individual's ability to lie successfully,[125] and individuals who suffer from these disorders tend to be over-represented in the criminal-justice population.[126] The sensitivity and specificity concerns discussed above are heightened for these outlier populations.[127] The lack of the presence of individuals suffering from these disorders in the pool of test subjects gives rise to concern about the applicability of the test results given their over-representation in the criminal-justice system, with previous studies showing them to be less amenable to other purported lie-detection methods.[128] Juveniles are another likely outlier population. Cognitive neuroscience already suggests that juvenile brains, which are not yet finally developed,

[124] *Cf.* Statistics Lectures, *supra* note 127.

[125] *Cf.* Temple Grandin & Sean Barron, THE UNWRITTEN RULES OF SOCIAL RELATIONSHIPS (2005) (describing how individuals suffering from autism understand and experience the social world differently than others).

[126] *See* Matt Vogel et al., *Mental Illness and the Criminal Justice System*, 8 Sociology Compass 627 (2014).

[127] *See* Daniel Langleben et al., *True Lies: Delusions and Lie-Detection Technology*, 34 J. PSYCHIATRY L. 351, 363 (2006) (finding that fMRI "lie-detection" technology is unreliable when used on individuals suffering from delusional disorders).

[128] *See id.* at 363; *see also*, Richard A. Friedman, *Behavior; Truth About Lies: They Tell a Lot About a Liar,* N.Y. TIMES (Aug. 5, 2003) http://www.nytimes.com/2003/08/05/health/behavior-truth-about-lies-they-tell-a-lot-about-a-liar.html, https://perma.cc/AA3E-VYVK; *see generally* Adrian Raine et al., *Corpus Callosum Abnormalities in Psychopathic Antisocial Individuals*, 60 ARCH. GEN. PSYCHIATRY 1134, 1139–40 (2003) (finding that individuals suffering from psychopathy have structural differences in their brain).

behave differently than adult brains.[129] Similar concerns arise regarding individuals who are under the influence of controlled substances, which have significant effects on pupil size and movement,[130] who are also over-represented in the criminal-justice system compared to the general population.[131]

[36] Neuroscience studies suggest that "normal" neural circuits are a necessary prerequisite for the ability to lie.[132] This, in turn, suggests that an individual with neural-circuit damage of some kind would have different results in these lie-detection tests than "normal" individuals.[133] At the other end of the spectrum, good liars, who have been shown to be able to mask the emotions that trigger negative polygraph results, may also be able to mask the emotions that may be the intermediate cause of pupil change.[134]

[37] Currently, no data exists about the number of individuals who cannot successfully take an ODT or ERP test, or the possible reasons why the tests might not work on individuals who belong to those groups –

---

[129] *See* Staci A. Gruber & Deborah A. Yurgelun-Todd, *Neurobiology and the Law: A Role in Juvenile Justice?*, 3 OHIO ST. J. CRIM. L. 321, 329–30 (2006).

[130] *See* Jason Peragallo et al., *Ocular Manifestation of Drug and Alcohol Abuse*, 24 CURRENT OPINION IN OPHTHALMOLOGY 566, 567 (2013).

[131] *See Addiction and the Criminal Justice System*, NATIONAL INSTITUTES OF HEALTH (2010), https://report.nih.gov/NIHfactsheets/ViewFactSheet.aspx?csid=22, https://perma.cc/FZ22-SV24.

[132] *See, e.g.*, ANTONIO DAMASIO, DESCARTES' ERROR: EMOTION, REASON, AND THE HUMAN BRAIN 10 (1994).

[133] *See* Jerrod Brown et al., *Brain Injury and Confabulation: A Review for Caregivers and Professionals*, CONCORDIA UNIVERSITY, ST. PAUL (July 16, 2014), https://online.csp.edu/blog/criminal-justice-online/brain-injury-and-confabulation-a-review-for-caregivers-and-professionals, https://perma.cc/VE6V-CRS4.

[134] *See* Lenese Herbert, *Othello Error: Facial Profiling, Privacy, and the Suppression of Dissent,* 5 OHIO ST. J. CRIM. L. 79, 94–95 (2007).

individuals taking certain medications, from certain cultural backgrounds, or with certain mental disorders, for instance.[135]

PART IV. CONCLUSION

[38] There are concerns about the forensic use of these new technologies that are particularly acute in the criminal context. The gatekeeping that *Daubert* was supposed to accomplish notwithstanding,[136] the criminal justice system has a history of judicial failure to engage in this assigned gatekeeping functions[137] and premature application of theoretical research techniques in forensic contexts.[138] For example, this past year, when the

[135] *See The Truth About Lie Detectors (aka Polygraph Tests)*, AMERICAN PSYCHOLOGICAL ASSOCIATION, http://www.apa.org/research/action/polygraph.aspx, https://perma.cc/32GQ-6CNF (Aug. 5, 2004).

[136] *See* Green, *supra* note 28, at 597.

[137] *See* Jane C. Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES' J., 2005 16, 28 (2005) ("The single most important observation about judicial gatekeeping of forensic science is that most judges under most circumstances admit most forensic science. There is almost no expert testimony so threadbare that it will not be admitted if it comes to a criminal proceeding under the banner of forensic science."); *See also* Jessica G. Cino, *An Uncivil Action: Criminalizing Daubert in Procedure and Practice to Avoid Wrongful Convictions*, 119 W. VA. L. REV. 651 (2016) (discussing how a lack of gatekeeping has led to problems in many cases such as wrongful convictions, etc.); *See e.g.*, Jessica M. Sombat, *Latent Justice: Daubert's Impact on the Evaluation of Fingerprint Identification Testimony*, 70 FORDHAM L. REV. 2819, 2842 (2002) (discussing how some judges, like the one in *Llera Plaza I*, exercised gatekeeping discretion which was not precedent).

[138] Several forensic techniques, which were widely used in criminal adjudication over lengthy periods of time, have subsequently been debunked, often after a spate of documented wrongful convictions. The bite-mark comparisons once regularly performed by forensic odontologists, forensic pattern analysis of fire scenes once regularly performed by arson investigators, and microscopic hair-comparison and comparative bullet-lead analysis ("CBLA") techniques that were spearheaded by the FBI for decades, have all subsequently been shown to have lacked the validity and accuracy that their proponents once claimed that they had. *See* People v. Acri, 662 N.E.2d 115, 117 (Ill. App. Ct. 1996) (holding that dog alerts to accelerants at a fire scene that were unconfirmed by laboratory analysis did not meet the *Frye* test for admission of scientific

Oregon Legislature had concerns about the validity and fairness of the new ODT tests, it introduced a bill to reign in their use, but only in the context of employment screening.[139] While banning them from employment use, the proposed bill specifically authorized their use by police and other

evidence); Ragland v. Commonwealth, 191 S.W.3d 569, 580 (Ky. 2006) (holding that a trial court's admission of the FBI's debunked CBLA evidence despite its reliability problems was clearly erroneous); Clemons v. State, 896 A.2d 1059, 1070 (Md. 2006) (holding that CBLA was not admissible under *Frye* because its processes and underlying assumptions were not generally accepted as valid and reliable); State v. Behn, 868 A.2d 329, 331 (N.J. Super. Ct. App. Div. 2005) (holding that the debunked CBLA technique was "based on erroneous scientific foundations"); JOHN J. LENTINI, SCIENTIFIC PROTOCOLS FOR FIRE INVESTIGATION 482–501 (2d. ed. 2013) (describing the debunking of traditional fire-pattern analysis); Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 83–85, 92 (2008) (suggesting forensic-science techniques, including hair comparison and bite-mark analysis, were insufficient in wrongful convictions and even false in some studied cases); Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 14–15 (2009) (discussing how problematic using hair comparison and other forensic science methods were in finding a correct conviction); Edward J. Imwinkelried & William A. Tobin, *Forensics Symposium: The Use and Misuse of Forensic Evidence; Comparative Bullet Lead Analysis (CBLA) Evidence: Valid Inference or Ipse Dixit?,* 28 OKLA. CITY U. L. REV. 43, 44–46 (2003) (suggesting there is valid evidence to contest the sufficiency of the previously well-established comparative bullet lead analysis used in crime scene investigations); Thomas R. May, *Fire Pattern Analysis, Junk Science, Old Wives Tales, and Ipse Dixit: Emerging Forensic 3D Imaging Technologies to the Rescue?*, 16 RICH. J.L. & TECH. 13 (2010), http://jolt.richmond.edu/v16i4/article13.pdf, https://perma.cc/93MF-9ASK (suggesting that previous fire investigations suffered many evidential problems, but new forensic 3D imaging technologies will serve as an improvement); Caitlin M. Plummer & Imran J. Syed, *"Shifted Science" Revisited: Percolation Delays and the Persistence of Wrongful Convictions Based on Outdated Science*, 64 CLEV. ST. L. REV. 483, 491-96 (2016) (describing various problems in fire-pattern analysis based on learning of the flashover problem as well as the cumbersome process of adopting NFPA 921); Eric Lichtblau, *F.B.I. Abandons Disputed Test for Bullets from Crime Scenes*, N.Y. TIMES (Sept. 2, 2005) at A12 (discussing how the F.B.I. has adopted a new method of testing bullet lead and reviews of prior cases are being offered).

[139] *See* H.B. 2545, 79th Legis. Assem., Reg. Session (Or. 2017).

criminal investigators in investigating crimes and supervising probationers and parolees.[140]

[39] If forensic science in American courts were to have a motto, "under validated and over sold" would be a serious contender.[141]This rush to admit questionably valid scientific evidence in the criminal context can be explained by several factors unique to criminal adjudication, all of which are extensively covered in the criminal-law and political-science literature. One is the lack of a culture of science and internal validation at police-run crime labs.[142] Another is what William Stuntz termed the "pathological politics" of criminal law making.[143] A third is the oft-bemoaned lack of competency at and resources for prosecutorial and public-defense agencies, which prevents more meaningful adversarial vetting of these techniques.[144]

---

[140] *See id.*

[141] *See* Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892 (2005), http://science.sciencemag.org/content/309/5736/892/tab-pdf, https://perma.cc/9YL5-6TJ5.

[142] *See* Harry T. Edwards, *Solving the Problems That Plague the Forensic Science Community*, 50 JURIMETRICS J. 5, 14 (2009); Geoffrey S. Mearns, *The NAS Report: in Pursuit of Justice*, 38 FORDHAM URB. L. J. 429, 430 (2010); Jennifer L. Mnookin et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725, 775–76 (2011); Saks & Koehler, *supra* note 146 at 893; *see generally* Karien Knorr-Cetina, *Epistemic Cultures: How the Sciences Make Knowledge*, 26 SCI., TECH. & HUMAN VALUES 390 (1999), https://www.jstor.org/stable/pdf/690270.pdf, https://perma.cc/6UG3-YQN3.

[143] *See generally* William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505 (2001) (arguing that criminal law's breadth and severity flow not from electoral politics but from institutional politics).

[144] *See* Jennifer E. Laurin, *Remapping the Path Forward: Toward a Systemic View of Forensic Science Reform and Oversight*, 91 TEXAS L. REV. 1051, 1094–95 (2013) (explaining the relationship between under-resourced prosecutors' offices and their failure to demand high-quality forensic-science evidence from police investigators); Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH S107, S110 (2005); *see also* Daniel Givelber, *Lost Innocence: Speculation and Data About the Acquitted*, 42 AM. CRIM. L. REV. 1167,

[40] The criminal justice system is supposed to be more concerned with avoiding false positives (false alarms) than false negatives (misses).[145] This distinguishes from other application settings for these technologies (*e.g.* employment screening, family law cases). If anything, the investigation and prosecution of crimes should be the last context in which a new theoretical technology is put to use.

1180–81 (2004) (describing how prosecutors and defense attorneys rely only on readily-available forensic-science during their pretrial decision making); *see, e.g.*, Elmore v. Ozmint, 661 F.3d 783, 872 (4th Cir. 2011) (holding that defense counsel's failure to investigate the State's forensic evidence constituted ineffective assistance of counsel); *see* Driscoll v. Delo, 71 F.3d 701, 709 (8th Cir. 1995) (holding that defense counsel rendered ineffective assistance for failing to challenge the State's serology evidence); *see* Letemps v. Secretary of Fla. Dept. of Corrections, 114 F. Supp. 3d 1216, 1231 (M.D. Fla. 2015) (holding that defense counsel's failure to investigate the testing procedures used by the State's forensic serologist on a semen stain from the victim's robe constituted ineffective assistance of counsel); *see* Commonwealth v. Epps, 53 N.E.3d 1247, 1268 (Mass. 2016) (holding that Epps was deprived of his constitutional right to present a defense by counsel's failure to call a forensic medical expert to rebut the Commonwealth's evidence of Shaken Baby Syndrome ("SBS")); *see* People v. Ackley, 870 N.W.2d 858, 867 (Mich. 2015) (holding that defense counsel's failure to consult a forensic pathologist to counter the State's SBS evidence constituted ineffective assistance of counsel).

[145] *See* In Re Winship, 397 U.S. 358, 372 (1970) (Harlan, J., concurring) (". . . . it is far worse to convict an innocent man than let a guilty man go free.").