5-1-2008

Detecting Lies Using Demeanor, Bias, and Context

Max J. Minzner
University of New Mexico - School of Law

DETECTING LIES USING DEMEANOR, BIAS, AND CONTEXT

*Max Minzner**

INTRODUCTION

> A fundamental premise of our criminal trial system is that the jury is the lie detector.
>
> –*United States v. Scheffer*[1]

The Supreme Court's statement in *Scheffer* reflects the conventional wisdom of the American legal system—courts assume that jurors, by closely observing demeanor, can accurately determine whether a witness is lying.[2] The consensus in the legal and social science literature is almost the opposite. "It is considered axiomatic . . . that individuals are at best inaccurate at deception detection."[3] Among legal academics, demeanor is seen as essentially useless in detecting deception, and decisions about lie detection are right no more than half the time.[4]

The negative legal academic perspective is based largely on two law review articles, both written well over a decade ago.[5] While these articles accurately summarized the literature on deception detection at the time, more recent research provides reason to believe that the academic view is incomplete. Even though demeanor is often unhelpful, in certain situations jurors and law enforcement officers can distinguish true and false stories with substantial accuracy using other methods. More importantly, we can identify situations in which they

* Assistant Professor, Benjamin N. Cardozo School of Law, Yeshiva University; E-mail: minzner@yu.edu. I am grateful to Monica Hakimi, Margaret Lemos, Alexander Reinert, and Stewart Sterk for their comments and suggestions on earlier drafts.

1 523 U.S. 303, 313 (1997) (internal quotations omitted).

2 NLRB v. Dinion Coil Co., 201 F.2d 484, 487 (2d Cir. 1952) ("Demeanor evidence may sometimes mislead, but our courts regard it nevertheless as an excellent cue to the trustworthiness of testimony.").

3 A.P. Hubbell, et al., *The Relative Effects of Timing of Suspicion and Outcome Involvement on Biased Message Processing*, 68 COMM. MONOGRAPHS 115, 115 (2001).

4 *See infra* Part I.

5 *See infra* text accompanying notes 24-39.

are likely to make the right decision about whether a witness is lying and, in turn, those situations in which their judgment is likely to be no better (and perhaps substantially worse) than a coin flip.

Part I provides a brief discussion of the divide between judges and academics in their opinion of the accuracy of legal credibility decisions and the role of demeanor evidence in making those determinations. Judges have generally assumed juries make accurate credibility decisions and believe demeanor is the mechanism for deciding whether a witness is telling the truth. Starting in the early 1990s, though, legal academics broke from this consensus view based on a series of social science studies demonstrating that test subjects in laboratory experiments correctly determined when a person was lying only slightly more than half the time.

Part II reviews more recent social science findings that demonstrate that the general consensus among legal academics is at least incomplete. Current studies show that the reality is more complicated than the legal literature reflects. In certain conditions, juries and law enforcement can differentiate between truth and lies. First, biases about witness credibility play a large role in determining whether deception will be caught. If a jury expects to hear the truth from a particular witness and that witness generally tells the truth, the jury will usually get the decision right. If, in turn, a witness who is expected to lie generally tells the truth, decisions about deception will usually be wrong. Second, context can dramatically improve lie detection accuracy. If the jury possesses information about the events that remains undisclosed to the witness, the witness is likely to be accurately classified as a liar or a truth teller. However, if private information is unavailable, observers are unlikely to make the right decision.

Part III applies these results in the criminal context. In particular, I examine decisions about deception by juries and by law enforcement regarding two types of speakers. First, I consider the attempt to classify correctly statements by cooperating witnesses testifying for the government or being interviewed as part of the cooperation process. Second, I look at situations where the defendant or a suspect is the witness, either when testifying at trial or during police interrogation.

Part IV concludes by placing these results in the context of the use of social science by legal academia. While the initial reviews of the value of demeanor evidence reflected a nuanced view of the social science literature, at least in the demeanor context, legal academics have tended to rely on the sound bite version of social science studies and failed to return to the original sources. To the extent that academics hope to persuade courts of the value of social science, more nuance is better.

I. LIE DETECTION THROUGH DEMEANOR

Courts often claim that demeanor is the way to distinguish the truth from lies. Several areas of law rest on a strong belief in the connection between demeanor and credibility. First and foremost, the Supreme Court has relied on the jury's ability to judge credibility from demeanor as a core policy foundation underlying the Confrontation Clause requirement of live witness testimony. Since *Mattox v. United States*, the Court has seen the benefit of the Clause as giving the accused "an opportunity . . . of compelling [the witness] to stand face to face with the jury in order that they may look at him, and judge by his demeanor upon the stand and the manner in which he gives his testimony whether he is worthy of belief."[6] The Court has continued to rely on the value of demeanor in setting the scope of the Confrontation Clause's requirements.[7] Along with the oath and cross-examination, the requirement that the jury be in a position to observe witness demeanor in order to judge credibility is the means of ensuring "that evidence admitted against an accused is reliable and subject to the rigorous adversarial testing that is the norm of Anglo-American criminal proceedings."[8] Indeed, demeanor alone can decide a criminal case. Juries may reject a defendant's testimony solely based on demeanor. Once the jury concludes a defendant has committed perjury, they are entitled to take that as affirmative evidence of guilt.[9]

Second, the role of demeanor in assessing witness credibility provides one of the standard (and oldest) justifications for appellate deference to lower court fact finding.[10] Whether a question is an issue of law or an issue of fact often turns on whether a credibility determination needs to be made based on demeanor.[11] The Court has identified the connection between demeanor and credibility as a source

---

6 Mattox v. United States, 156 U.S. 237, 242-43 (1895); Donelly v. California, 228 U.S. 243 (1913).

7 Confrontation "permits the jury that is to decide the defendant's fate to observe the demeanor of the witness in making his statement, thus aiding the jury in assessing his credibility." California v. Green, 399 U.S. 149, 158 (1970).

8 Maryland v. Craig, 497 U.S. 836, 846 (1990).

9 Wright v. West, 505 U.S. 277, 296 (1992).

10 *See* Rice v. Collins, 546 U.S. 333, 342-44 (2006) (Breyer, J., concurring) (pointing out the role of demeanor in credibility determination as justification for appellate deference); Lilly v. Virginia, 527 U.S. 116, 137 (2001) (describing demeanor as a "factor uniquely suited to the province of trial courts"). The Supreme Court relied on the role of demeanor in making credibility decisions as a reason for deference to lower courts in admiralty cases almost as far back as the Civil War. *See* The Quickstep, 76 U.S. 665, 669 (1869).

11 Thompson v. Keohane, 516 U.S. 99, 111, 114 (1995); Miller v. Fenton, 474 U.S. 104, 113-14 (1985).

of appellate deference to trial court decisions to disqualify jurors,[12] to state court decisions in the habeas context, [13] to district court evaluations of the credibility of the reasons proffered by the government during *Batson* challenges,[14] and for the very existence of the clearly erroneous standard under Federal Rule of Civil Procedure 52(a).[15]

Law enforcement training manuals have generally taken the same view—accurate lie detection depends on demeanor evidence. Police officers often must evaluate whether witnesses and suspects are telling the truth during interviews. The most influential current training method for law enforcement is the Reid technique, outlined in Reid and Inbau's book *Criminal Interrogation and Confessions*.[16] The authors claim to have trained over 150,000 law enforcement officers in the past 35 years. Inbau and Reid explicitly encourage law enforcement officers to evaluate demeanor. For example, they assert that "a suspect who does not make direct eye contact is probably withholding information."[17] Similar texts express comparable views.[18] While these texts are based on the experience of law enforcement officers working in the field, they do not rely on any formal empirical findings for their belief in the value of demeanor.

Legal commentators traditionally adopted the same approach to demeanor as courts and law enforcement, concluding that the demeanor of a witness was "the solution of the always difficult problem of determining the truthfulness of his testimony."[19] Beginning in the

---

12 Judges must "reach conclusions as to impartiality and credibility by relying on their own evaluations of demeanor evidence and of responses to questions." Rosales Lopez v. United States, 451 U.S. 182, 188 (1981); *see also* Patton v. Yount, 467 U.S. 1025, 1038 (1984).

13 Marshall v. Lonberger, 459 U.S. 422, 434 (1983). Demeanor is also one of the reasons why evidentiary hearings are required in some habeas cases. "Where an unresolved factual dispute exists, demeanor evidence is a significant factor in adjudging credibility." Townsend v. Sain, 372 U.S. 293, 322 (1963).

14 Hernandez v. New York, 500 U.S. 352, 365 (1991). In the multistage *Batson* process for objecting to the government's use of race in the striking of jurors, a defendant must first make a preemptory showing that the government is engaging in race-based strikes. Next, the prosecutor must provide a race-neutral justification for striking the juror. Finally, the court must evaluate the justifications and determine whether the government engaged in purposeful discrimination. At the second stage, the district court must evaluate the credibility of the prosecution's claimed race neutral justification. The Supreme Court requires substantial appellate deference for that credibility determination because of the importance of demeanor in that decision. *See, e.g.*, Miller-el v. Cockrell, 537 U.S. 322, 328, 339-41 (2003).

15 *See* FED. R. CIV. P. 52(a); Anderson v. City of Bessemer, 470 U.S. 564, 575 (1985).

16 FRED E. INBAU ET AL., CRIMINAL INTERROGATION AND CONFESSIONS (Jones and Bartlett 2004) (1962).

17 *Id.* at 151.

18 DAVID E. ZULAWSKI & DOUGLAS E. WICKLANDER, D.E., PRACTICAL ASPECTS OF INTERVIEW AND INTERROGATION (CRC Press 1993) (1992).

19 Henry S. Sahm, *Demeanor Evidence. Elusive and Imponderables*, 47 A.B.A. J. 580, 580 (1961). "[D]emeanor evidence 'may satisfy the tribunal, not only that the witness' testimony is not true, but that the truth is the opposite of his story.'" Fleming James, Jr., *Sufficiency of the Evidence and Jury-Control Devices Available Before Verdict*, 47 VA. L. REV. 218, 223-24 (1961)

1960s, though, social scientists started laboratory experiments on the value of demeanor evidence.[20] These studies found that while the subjects had strong beliefs about cues of deception, these beliefs bore little relationship to reality.[21] For instance, while subjects believed that witnesses who were lying tended to have shifty eyes, witnesses who were lying did not avert their gaze more frequently than witnesses telling the truth. More importantly, when presented with an equal number of true and false statements, lie detection accuracy hovered around the level of chance. Evaluations of deception detection were accurate between 45% and 60% of the time.[22] Studies of law enforcement officers show similar results. While some studies show that law enforcement officers outperform civilians, the results are mixed at best.[23]

These social science results did not immediately penetrate the legal literature.[24] In the early 1990s, though, law review authors began to present arguments doubting the value of demeanor in making credibility determinations.[25] In particular, Olin Wellborn[26] and Jeremy

---

(quoting Dyer v. McDougall, 201 F.2d 265, 269 (2d Cir. 1952)). Demeanor was viewed not just as a mechanism to determine whether a witness was truthful, but also whether he has the capacity to observe what he claimed to have seen. "When an eyewitness testifies at trial, the defendant's most valuable weapon is the opportunity to probe the witness' subjective abilities so that the jury may consider his responses and demeanor in assessing the credibility of his testimony." James W. Jennings, *Preserving the Right to Confrontation—A New Approach to Hearsay Evidence in Criminal Trials*, 113 U. PA. L. REV. 741, 751 (1965).

20 While serious research did not begin until the 1960s, the first studies occurred in the 1920s and found that jurors hearing live testimony performed worse on deception judgments than those who reviewed transcripts. *See* William M. Marston, *Studies in Testimony*, J. AM. INST. CRIM. L. & CRIMINOLOGY 5, 22 (1924).

21 For summaries of these results, see Joseph W. Rand, *The Demeanor Gap: Race, Lie Detection, and the Jury*, 33 CONN. L. REV. 1, 7-14 (2000) and 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, 1190-94 (1993).

22 *See* Miron Zuckerman et al., *Verbal and Nonverbal Communication of Deception*, 14 ADVANCES IN EXPERIMENTAL SOC. PSYCHOL. 1, 26 (1981).

23 *See* Samantha Mann et al., *Detecting True Lies: Police Officers' Ability to Detect Suspects' Lies*, 89 J. APPLIED PSYCHOL. 137, 137 (2004); Christian A. Meissner & Saul M. Kassin, *"He's Guilty!": Investigator Bias in Judgments of Truth and Deception*, 26 LAW & HUM. BEHAV. 469, 470 (2002).

24 As late as 1985, articles were emphasizing the importance of demeanor evidence in determining witness credibility. *See* Edward J. Imwinkelried, *Demeanor Impeachment: Law and Tactics*, 9 AM. J. TRIAL ADVOC. 183, 186 (1985).

25 James P. Timony, *Demeanor Credibility*, 49 CATH. U. L. REV. 903, 906 (2000). Some law professors had previously noted the weak empirical foundations for the strong judicial statements about demeanor, but none had much impact. *See* Edward H. Cooper, *Directions for Directed Verdict: A Compass for Federal Courts*, 55 MINN. L. REV. 903, 934 (1970) ("[C]ompared to persons making judgments based on recordings or transcripts of the same testimony, persons making judgments on the basis of observation of the 'witnesses' form significantly *less* accurate judgments."); Charles T. McCormick, *Deception-Tests and the Law of Evidence*, 6 TENN. L. REV. 108, 127 & n.50 (1928) (noting that jurors relying on demeanor were only correct in their determinations 48% of the time). This later article is the earliest statement in the law review literature of which I am aware disputing conventional beliefs about demeanor.

Blumenthal[27] prepared detailed surveys of the then-current psychological literature on deception detection. Wellborn's 1991 article was the first to introduce psychological studies of deception detection to a legal audience in a systematic way.

Relying on studies from the 1960s through the early 1980s, Professor Wellborn pointed out that test subjects were *less* accurate in detecting deception when they observed the speaker on videotape rather than when they reviewed transcripts or listened to an audio recording.[28] In addition, experimental subjects did not make significantly better deception judgments when observing cues from the speakers' face or body. Wellborn recognized the potentially significant differences between laboratory experiments and the courtroom environment, particularly the presence of context, cross-examination, deliberation, and preparation, but he dismissed these as mechanisms that might improve decisions about deception and argued that they are in fact likely to reduce the validity of deception determinations.[29] Wellborn concluded that "the experimental evidence indicates that ordinary observers do not benefit from the opportunity to observe nonverbal behavior in judging whether someone is lying" and cited the now-conventional wisdom that "most people cannot do much better than chance in discerning lies in laboratory conditions."[30] Even more troubling, mock jurors believe that they are far better lie detectors than they actually are.[31]

Blumenthal concurs in Wellborn's analysis. Blumenthal emphasized that the public held strong views about nonverbal behavior that indicated deception.[32] However, the cues "popularly believed to manifest deception are qualitatively and quantitatively different from those which are actually observed during deception."[33] People generally believed that a reduction in smiling, an increase in furtive glances, fidgeting, and gaze avoidance indicated that the witness was lying. In fact, none of these beliefs had support in the social science literature.[34] Blumenthal concluded that "it is unforgivable that the legal system deliberately ignores demonstrated, relevant findings about demeanor evidence and willfully adheres to an ineffectual traditional approach."[35]

---

26 Olin Guy Wellborn III, *Demeanor*, 76 CORNELL L. REV. 1075 (1991).

27 Blumenthal, *supra* note 21.

28 Wellborn, *supra* note 26, at 1086-87.

29 *Id.* at 1079.

30 *Id.* at 1088.

31 *See* George Fisher, *The Jury's Rise as Lie Detector*, 107 YALE L.J. 575, 707 (1997); Wellborn, *supra* note 26, at 1088.

32 *See* Blumenthal, *supra* note 21, at 1194.

33 *Id.*

34 *Id.* at 1194.

35 *Id.* at 1204.

Neither article gained traction among courts or legislatures. Only three decisions have cited either Blumenthal or Wellborn for the proposition that demeanor evidence is overrated; all three do so merely in passing.[36] The Federal Rules have preserved the traditional view on demeanor. Rule 43 continues to require that testimony be taken live in open court. In limiting the use of video testimony, the advisory committee notes stress that "[t]he importance of presenting live testimony in court cannot be forgotten. . . . The opportunity to judge the demeanor of a witness face to face is accorded great value in our tradition."[37] Of course, the importance of demeanor remains a justification for excluding hearsay evidence under the Rules of Evidence,[38] and the use of demeanor evidence remains a standard part of jury instructions.[39]

In contrast to the reaction among courts, both the Blumenthal and Wellborn articles have been widely adopted in the law review literature. Legal commentators have generally accepted the view that "psychological studies strongly indicate that observers do no better than pure chance in evaluating live witnesses."[40] Indeed, Blumenthal and

---

36 Edmunds v. Deppisch, 313 F.3d 997, 1000 (7th Cir. 2002); Morales v. Artuz, 281 F.3d 55, 61 (2d Cir. 2001); Koskela v. Willamette Indus, 978 P.2d 1018, 1028 & n.14 (Or. Ct. App. 1999), *rev'd on other grounds*, 15 P.3d 548 (Or. 2001). I am not aware of any court citing to any of the original social science literature on demeanor discussed in this Article or in the Blumenthal and Wellborn pieces.

37 FED. R. CIV. P. 43 advisory commitee notes, 1996 Amendment.

38 FED. R. EVID. advisory commitee notes, Intro. to Art. VIII; Williamson v. United States, 512 U.S. 594, 598 (1994).

39 *See, e.g.*, JUDICIAL COMM. ON MODEL JURY INSTRUCTIONS FOR THE EIGHTH CIRCUIT, MANUAL OF MODEL CRIMINAL JURY INSTRUCTIONS FOR THE DISTRICT COURTS OF THE EIGHTH CIRCUIT § 3.04(2007); PATTERN CRIMINAL JURY INSTRUCTINOS FOR THE DISTRICT COURTS OF THE FIRST CIRCUIT ¶ 3.06 (2007); ELEVENTH CIRCUIT PATTERN JURY INSTRUCTIONS (CRIMINAL CASES) ¶ 2.2 (2003).

40 David Crump, *The Case for Selective Abolition of the Rules of Evidence*, 35 HOFSTRA L. REV. 585, 610 (2006) (citing Blumenthal and Wellborn). Recent articles taking the view that demeanor is generally unhelpful include Leo Kittay, Note, *Admissibility of FMRI Lie Detection*, 72 BROOK. L. REV. 1351, 1388 (2007) ("[S]tudies have shown that people are not good at determining anothers' truthfulness."); John Leubsdorf, *Presuppositions of Evidence Law*, 91 IOWA L. REV. 1209. 1252 (2006) ("Empirical research shows that jurors are not able to tell by observing witnesses which of them is telling the truth."); Deana Kim El-Mallawany, Comment, Johnson v. California *and the Initial Assessment of Batson Claims*, 74 FORDHAM L. REV. 3333, 3354 (2006) ("Studies have shown, however, that judging credibility based on demeanor is susceptible to great inaccuracy." (citing Wellborn)); Michelle J. Anderson, *The Legacy of the Prompt Complaint Requirement, Corroboration Requirement, and Cautionary Instructions on Campus Sexual Assault*, 84 B.U. L. REV. 945, 948 n.12 (2004) ("[U]ntrained individuals cannot do much better than chance in discerning lies under experimental conditions."); Chad M. Oldfather, *Appellate Courts Historical Facts, and the Civil Criminal Distinction*, 57 VAND. L. REV. 437, 458 (2004) ("The experimental evidence strongly suggests that the ability to observe demeanor is of no value in assessing witness credibility."); Charles H. Koch, 11 IND. J. GLOBAL LEGAL STUD. 139, 156 & n.64 (2004) ("That the one who hears the witness is best able to evaluate the testimony is a fundamental assumption of the common law. However, this assumption may not be supportable in fact."); John G. Douglass, *Virtual Cross Examination: The Art of Impeaching Hearsay*, 34 J. MAR. L. & COM. 149, (2003) ("In fact, there is no empirical

Wellborn have largely displaced the original social science research among legal commentators. Aldert Vrij and Bella DePaulo, two of the most prolific deception detection researchers, have fewer citations combined than either Blumenthal or Wellborn has individually.[41]

We now are in a situation where there are two competing consensus views about demeanor and credibility. Legal critics deride demeanor evidence and conclude that lie detection is essentially impossible; courts depend on it. As the next section shows, neither view is right. Wellborn and Blumenthal's studies were accurate and thoughtful summaries of the psychological literature regarding demeanor evidence at the time they were written. Today's findings, though, are far more complex than the then-current research and give us a much greater ability to identify those situations in which we need to worry about mistaken judgments about credibility and those in which we do not.

II. DECEPTION DETECTION RESEARCH: DEMEANOR, BIAS AND CONTEXT

Research on deception detection has generally presented only limited variations on a theme. These studies usually involve displaying either live or video testimony to college students and asking them whether the witness was lying. Part II.A describes the current state of the research on deception detection and demeanor. In general, the research supports the conclusion that demeanor plays little role in improving accuracy. Part II.B, though, demonstrates that this result does not support the further conclusion that deception detection accuracy is no better than flipping a coin. Indeed, results suggest that in

---

evidence that jurors' assessments of credibility based on witness demeanor are likely to be any more reliable than random guesses."); Edward J. Imwinkelried, *Trial Judges—Gatekeepers or Usurpers*, 84 MARQ. L. REV. 1, 32 (2000) ("However, the most recent psychological research indicates that a witness's demeanor on the stand can be so idiosyncratic that it is often an unreliable indicator of the witness's truthfulness."); Robert J. Condlin, *What's Really Going On? A Study of Lawyer and Scientist Inter-Disciplinary Discourse*, 25 RUTGERS COMPUTER & TECH. L.J. 181, 249-50 (1999) ("[T]here are no reliable standards for deciding when an eyewitness is telling the truth."); Donald Dripps, *Miscarriages of Justice and The Constitution*, 2 BUFF. CRIM. L. REV. 635, 670-71 (1999) ("Empirical evidence overwhelming indicates that ordinary people cannot detect deception by observing demeanor."); Chris W. Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227, 1245 (2001) ("The experimental evidence on lay assessment of demeanor casts serious doubt on the ability to human subjects to assess witness credibility"); and James F. Flanagan, *Redefining the Role of the State Administrative Law Judge*, 54 ADMIN. L. REV. 1355, 1397 & n.179 (2002) ("In general, most people have difficulty detecting lies. Thus, relying on non-verbal behavior produces results roughly equivalent to chance.").

[41] A Westlaw search of the JLR database performed July 12, 2007 for "demeanor /50 (vrij depaulo)" produced 14 hits, while similar searches for "demeanor /50 wellborn" and "demeanor /50 blumenthal" produced 89 hits and 36 hits, respectively.

certain circumstances, we should expect accuracy to be substantially higher or lower than the 50% mark. First, when the evaluators of deception are biased in the correct way, accuracy is significantly improved. Second, when evaluators have secret information about the events described, accuracy is also much higher.

A. *Demeanor and Deception Detection*

The traditional approach to demeanor in the legal system depends on two assumptions: first, that liars and truth-tellers reliably exhibit different cues; and second, that observers can detect these cues. Two recent meta-analyses[42] have summarized the current findings relating to the existence of deception cues. First, a 2003 study by Bella DePaulo and co-authors reviewed 116 studies examining deception detection.[43] They found that, in general, few reliable cues to deception exist and in particular, the cues widely believed by the public to signify deception generally do not. As mentioned above, gaze aversion, speech disturbances (use of "ah" and "um"), longer and more pauses, eye blinking, and fidgeting are all usually considered signs of deception. In fact, liars do not avert their gaze any more frequently than honest witnesses,[44] and the same is true for these other behaviors widely associated with deception; DePaulo finds that all of these effect sizes are "small."[45] Sporer's 2007 study of a more limited set of cues came to similar conclusions. Looking at 12 standard behaviors identified with deception, including, among others, gaze aversion, smiling, and fidgeting, the Sporer study found little evidence that these nonverbal cues bore any relationship to deceit.[46]

---

[42] Siegfried L. Sporer et al., *Moderators of Nonverbal Indicators of Deception: A Meta-analytic Synthesis*, 13 PSYCHOL. PUB. POL'Y & L. 1 (2007); Bella M. DePaulo et al., *Cues to Deception*, 129 PSYCHOL. BULL. 74 (2003). A meta-analysis is a statistical technique that takes other social science studies as the observational units and tries to measure an effect size across all of the studies. DePaulo et al. use Cohen's d to measure effect size while Sporer et al. use the functionally equivalent Pearson's correlation measure r. *Id.* at 89; Sporer, *supra*, at 2 & n.2; JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 20-24 (1988). This choice of measures of effect size, as well as the general theory underlying meta-analyses, is subject to quite reasonable criticism. *See* Chris W. Sanchirico, *Evidence, Procedure, and the Upside of Cognitive Error*, 57 STAN. L. REV. 291, 311 & n. 93 (2004); RICHARD BERK, REGRESSION ANALYSIS: A CONSTRUCTIVE CRITIQUE, 196-200 (2004). Despite these critiques, meta-analysis and the use of d or r remain relatively standard among applied social scientists.

[43] DePaulo et al., *supra* note 42.

[44] *Id.* at 93 (noting that the effect size of the estimates of gaze aversion and eye contact were not significantly different than zero).

[45] *Id.* at 95. DePaulo, consistent with Cohen, selects a threshold of $-.02<d<0.2$ as defining a "small" effect size. One of the traditional indicators of deception, a rise in vocal pitch, exceeded this threshold in the DePaulo study. *Id.*

[46] Sporer did find that liars decreased their hand and finger movements and that this effect size was greater than "small." Sporer et al., *supra* note 42, at 17. Note that this result was not

Even if cues to deception exist, the relevant question for the legal system is whether observers can detect these cues. Extensive studies have considered the ability to distinguish between true and false statements in the experimental context. An early review of pre-1980 studies showed an average accuracy rate of 57%.[47] A similar 2000 review of studies from the previous twenty years found a similar result —an accuracy rate of 56.6%.[48] These results are the basis of the conclusion generally cited in the legal literature that lay observers are only slightly better than chance at detecting deception.

Most recently, Charles Bond and Bella DePaulo recently summarized the studies of deception detection in a large-scale meta-analysis.[49] They found a comparable summary statistic—study participants correctly classified truths and lies with a mean accuracy rate of only 53.46%.[50] This summary statistic, though, conceals some variation. Deception detection improves moderately when the receiver has access to a more varied medium—observers who see but do not hear the sender are less accurate than those who hear but do not see and both are less capable than observers who can both see and hear the sender.[51] Prepared lies are somewhat harder to detect than unprepared.[52] Overall, though, Bond and DePaulo conclude that "rates of lie detection vary within a narrow range . . . . within a few points of 50%."[53] In summary, these results are generally consistent with the dismal view among legal academics. Demeanor cues do not lead to accurate lie detection.

---

found in the DePaulo study—its results did not find a significant relationship between deception and hand movements. DePaulo et al., *supra* note 42, at 92.

47 R. E. Kraut, *Humans as Lie Detectors. Some Second Thoughts*, 30 J. COMM. 209 (1980).

48 ALDERT VRIJ, DETECTING LIES AND DECEIT 75 (2000).

49 Charles F. Bond Jr. & Bella M. DePaulo, *Accuracy of Deception Judgments*, 10 PERS. & SOC. PSYCHOL. REV. 214 (2006).

50 *Id.* at 222. This figure is the weighted mean; the unweighted mean was comparable at 53.98%. *Id.* at 219. While this effect appears small, it is strongly statistically significant. *Id.* at 222 (concluding that the weighted mean is significantly different from the null hypothesis that the mean is 50% at a level of $p < 0.0001$). Furthermore, it compares favorably to other effect sizes in social psychology. For instance, this is approximately the same effect size observed in support of the conclusion that men are, on average, more aggressive than women. Bond and DePaulo measure the effect size here using the same method as the DePaulo 2003 study mentioned above and is subject to the same criticism. *See* DePaulo et al., *supra* note 42. Bond and DePaulo find a d of 0.39, which is the equivalent of an r of 0.21. F.D. Richard and co-authors surveyed a number of meta-analytic studies in 2003 and provided r figures for comparison. *See* F.D. Richard et al., *One Hundred Years of Social Psychology Quantitatively Described*, 7 REV. GEN. PSYCHOL. 331 (2003). In that survey, they found that the average r for the conclusion that men were more aggressive than women was also 0.21.

51 Bond & DePaulo, *supra* note 49, at 226.

52 *Id.* at 227.

53 *Id.* at 231.

B. *New Results on Deception Detection*

While the social science results discussed in Section II.A support the view that examining demeanor alone does not lead to accurate credibility determinations, they do not demonstrate that real world credibility decisions are accurate only 50% of the time. More recent studies suggest that in certain situations, lie detection accuracy is likely to be substantially better than that figure suggests, while in others, it is likely to be substantially worse. First, the inaccuracy of demeanor cues only matters if jurors and police officers actually rely on them. While people purport to believe that demeanor cues indicate whether speakers are lying, recent work suggests that people do not actually use these cues to make decisions.[54] By contrast, while demeanor does not play a significant role in real-world deception detection, the context surrounding the speaker's statement does appear to matter. When observers have background information about the witness's statement, they use it and lie detection accuracy improves considerably. Finally, if context is unavailable, recent results suggest that bias plays an important role.[55] The observer's initial likelihood of believing the witness largely determines the ultimate conclusion about truthfulness.

Skeptics of legal lie detection assume that individuals rely heavily on demeanor cues. This assumption initially appears well-founded. In experimental studies, people claim to be able to identify demeanor cues to deception and express confidence in their ability to evaluate them.[56] However, this result is likely an artifact of the experimental design—experimenters may be effectively forcing people to rely on demeanor. In 2002, researchers surveyed almost 200 undergraduates about situations in which they had detected someone lying in their life and asked how the lie was uncovered.[57] The survey strongly indicates that demeanor plays a limited role in real-world lie detection. Respondents identified verbal and non-verbal behavior as a deception detection mechanism in only 11.4% of the reported lies and even then, the mechanism was generally used in combination with some other technique. Demeanor standing alone was only used to detect deception in 2.1% of the lies.[58]

If people do not rely on demeanor to detect deception, what does drive the decision to classify a witness as truthful? Study subjects reported relying heavily on context. They detected lies using

---

[54] *See infra* notes 56-58 and accompanying text.

[55] *See infra* notes 59-68 and accompanying text.

[56] *See* Bella DePaulo et al., *The Accuracy-Confidence Correlation in the Detection of Deception*, 1 PERS. & SOC. PSYCHOL. REV. 346, 346 (1997).

[57] Hee Sun Park et al., *How People Really Detect Lies*, 69 COMM. MONOGRAPHS 144 (2002).

[58] *Id.* at 150-51.

information received from third parties, physical evidence, and solicited confessions.[59] The timing of lie detection also supports the relative unimportance of demeanor. Demeanor cues should lead to immediate lie detection, but subjects reported detecting lies long after they took place. Over 80% of the lies were discovered more than an hour after they were told and over 60% were uncovered over a day later.[60]

A recent study of juror questioning further suggests that observers rely heavily on context when making credibility decisions. When context is not available to jurors, they seek it out. Shari Seidman Diamond and her co-authors recently examined the questions that jurors posed to witnesses in 50 Arizona civil trials, where court rules provide for juror questioning.[61] A plurality of the questions asked (42%) fell into the category they label "cross-checking," questions asked to generate evidence from disinterested witnesses to compare to the testimony of witnesses whose credibility is more dubious.[62] Jurors not only sought out this information, they were displeased when they could not get it. The study had access to videotapes of jury deliberations and when reviewing the video, they found that "jurors often bemoan the absence of a witness to the events who is not associated with one of the parties, and would thus presumably offer a more trustworthy, or at least independent account of the events that transpired."[63]

Observers not only use context, they also use it effectively. Recent studies show that detailed questioning by an interrogator with an informational advantage substantially increases lie detection accuracy. In two studies, Maria Hartwig and collaborators tested the effect of questioning when the questioner knows information that the subject does not know and discloses it during the interview.[64] In both studies, the experimenter had the subject engage in a mock crime and provided the interrogator evidence about the crime that was unknown to the subject. In the first study, the experimenters videotaped interrogations where the questioner engaged in either early or late disclosure of the information. Either the witness learned the secret information at the

---

[59] Considering those lies that were detected by a single method, the respective figures for third party information, physical evidence, and solicited confessions were used in 32%, 18%, and 3.6% of the cases. For lies detected by a combination of methods, the respective figures are 52.1%, 30.9%, and 18.6% of cases. *Id.*

[60] *Id.* at 152. As Park et al. note, these results are likely biased in favor of shorter response times because respondents were directed to think of a recent lie. *Id.*

[61] *See* Shari Seidman Diamond et al., *Juror Questions During Trial: A Window Into Juror Thinking*, 59 VAND. L. REV. 1927 (2006).

[62] *Id.* at 1956-57.

[63] *Id.* at 1960.

[64] Maria Hartwig et al., *Strategic Use of Evidence During Police Interviews: When Training to Detect Deception Works*, 30 LAW & HUM. BEHAV. 603 (2006) [hereinafter Hartwig I]; Maria Hartwig et al., *Detecting Deception via Strategic Disclosure of Evidence*, 29 LAW & HUM. BEHAV. 469 (2005) [hereinafter Hartwig II].

start of the interview before questioning began or alternatively, the interviewer waited to disclose the information until after the subject had already told a story. When the information was disclosed late in the interview, accuracy decisions increased substantially. Judges who observed late disclosure of evidence were correct in 61.7% of cases compared to merely 42.8% of cases in which the judges observed early disclosure of the evidence.[65] Not surprisingly, when presented evidence in advance, liars incorporated it into the lies they told, making deception detection more difficult.

Hartwig's more recent study produced even stronger effects. In the second study, Hartwig trained police investigators to disclose evidence late and ask questions relating to the concealed evidence without revealing it. Experimenters then compared the deception detection ability of trained questioners to an untrained control group. Interrogators trained to disclose evidence late correctly classified 85% of true statements and 85.7% of false statements, compared to untrained interrogators, who classified correctly only 57.1% of true statements and 55% of false statements.[66] The study not only observed large effects, but also identified a mechanism. Trained questioners successfully induced inconsistencies in false stories.[67] The DePaulo meta-analysis results support the findings of both of these studies. While liars do not give off demeanor cues, they do tell stories that are less logical, less consistent, and contain fewer details than those of truth-tellers.[68] These traits should be far more apparent if the questioner has an informational advantage and can use it to craft good questions.

When context is unavailable, what determines accuracy rates? Researchers have consistently seen a veracity effect in deception studies—people classify truthful messages more accurately than false ones. The DePaulo study found that people correctly classify 61.34% of truthful messages as truthful but only 47.55% of false messages as false.[69] These results are robust and have been demonstrated widely

---

[65] Hartwig I, *supra* note 64, at 477. Notably, this study shows a reversal of the truth bias—late disclosure of evidence led to accurate judgments in 67.6% of false statements but merely 53.8% of true statements. *Id.*

[66] Hartwig II, *supra* note 64, at 613.

[67] *Id.* at 617.

[68] In addition to physical demeanor cues, DePaulo and colleagues expanded their study to include cues relating to the external verifiability of stories told. These experimenters found that liars provide fewer details than truth tellers, are less likely to embed their stories in context, and are more likely to tell implausible stories, stories that are illogical, and stories with internal discrepancies. DePaulo et al., *supra* note 42, at 92, 94-97. These findings support the Hartwig results.

[69] *Id.* at 223. These figures reflect the unweighted means. They are comparable to earlier results recognizing the differential evaluations of true and false statements. *See* VRIJ, *supra* note 48, at 69 (finding that observers detect true statements at a 67% accuracy rate but false statements only 44% of the time).

across the literature.[70] The source of the veracity effect is a "truth bias." Experimental subjects are substantially more likely to decide a statement is truthful than deceitful.[71] As a result, they are far more accurate at identifying truths than lies.[72] The source of the truth bias is unclear, although one commonly proposed explanation is that the bias results from a version of the availability heuristic. Because most statements encountered in ordinary life are true, people tend to assume that statements are more likely than not to be accurate.[73]

Regardless of the source of the truth bias, it has important implications for deception detection. Levine and his co-authors have recently demonstrated that whether true statements are believed and false statements disbelieved is largely a function of two factors: the strength of the truth bias and the base rate of message truthfulness.[74] When receivers are more likely to evaluate messages as true than false, these receivers will generally classify messages correctly when most messages are truthful but will usually be wrong when most messages are false. A symmetric result holds for a lie bias. This result should not come as a surprise—people who tend to believe most statements will usually be accurate if most messages are truthful. This tendency suggests that the common finding that deception detection accuracy rates hover around 50% is an artifact of the research design, holding only if half of the messages are true and half are false.

These results do not mean that, absent context, people are better at judging deception than random chance would predict. Indeed, it

---

[70] *See* Samantha Mann et al., *Detecting True Lies: Police Officers' Ability to Detect Suspects' Lies*, 89 J. APPLIED PSYCHOL. 137, 137 (2004) ("[J]udges are more likely to consider that messages are truthful than deceptive and, as a result, truthful messages are identified with relatively high accuracy (67%) and deceptive messages with relatively low accuracy (44%).").

[71] For instance, the DePaulo study found that 55.23% of statements are coded as true even though the standard practice in deception detection experiments is to include an equal mix of true and false statements. This figure also reflects the weighted mean. Again, the unweighted mean was comparable at 56.86%. *Id.* at 223. *See generally* Kevin John Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 286 (2006).

[72] Levine et al. demonstrate this result in a very straightforward manner. They show highly positive correlations (ranging from 0.5 to 0.66) between truth bias and truth accuracy and highly negative correlations between truth bias and lie accuracy (ranging from -.075 and -0.82). By contrast, there is limited correlation between truth and lie accuracy. Timothy Levine et al., *Accuracy in Detecting Truths and Lies: Documenting the "Veracity Effect,"* 66 COMM. MONOGRAPHS 125, 141 (1999).

[73] VRIJ, *supra* note 48, at 69; Maureen O'Sullivan et al., *The Effect of Comparison on Detecting Deceit*, 12 J. NONVERBAL BEHAV. 203, 205 (1988). Vrij suggests that general rules of social politeness or stereotypes about deceptive behavior might also cause a truth bias. VRIJ, *supra*, at 69.

[74] *See* Timothy Levine et al., *Deception Detection Accuracy is a Predictable Linear Function of Message Veracity Base Rate: A Formal Test of Park and Levine's Probability Model*, 73 COMM. MONOGRAPHS 243 (2006); Hee Sun Park & Timothy Levine, *A Probability Model of Accuracy in Deception Detection Experiments*, 68 COMM. MONOGRAPHS 201 (2001); Levine, *supra* note 72. The description in the text elides some of the complexities of the analysis which are outlined further in the Appendix.

supports exactly that conclusion. If the truth bias and the base rate determine accuracy rates, we can still think of the classification of statements as true or false as the outcome of a coin flip, but not a *fair* coin flip. Instead, at least in the experimental context, the coin is weighted to come up true more than 50% of the time.[75] As the next section shows, different observers have different biases that may lead to more (or less accurate) classification of liars and truth-tellers. Jurors and law enforcement may have coins that are biased in different directions.

III. BIAS, CONTEXT, COOPERATORS AND DEFENDANTS

What do these results about demeanor, bias, and context indicate about lie detection in the legal system? The results about context are relatively straightforward. When a legal decision-maker possesses private information and uses it to probe the witness's story, lie detection accuracy will improve. As a result, we should expect inaccurate lie detection in single-witness cases, where all of the evidence about the event comes from one source. By contrast, we can be more optimistic when more information is available. If the police and juries have multiple independent sources of information about an event, they are far more likely to accurately decide which witnesses to believe. The results with respect to bias are more complicated. Rather than examining generic lie detection ability, we need to investigate whether the bias fits well with the base rate. Do jurors and law enforcement tend to be truth-biased when witnesses are likely to be truthful, and, alternatively, do they tend to be lie-biased in situations where witnesses are likely to be lying? Note that, from a legal standpoint, we are not primarily interested in whether particular decision makers are truth-biased or lie-biased in the abstract. Instead, we are concerned with the fit between bias and reality. Do the legal structures cause juries and law enforcement to be skeptical of the lying witnesses and credulous of the truthful witnesses? If so, we should be optimistic about lie detection decisions. In turn, if there is a lack of fit, i.e., jurors and police officers are lie-biased toward generally truthful witnesses and vice versa, we should be very pessimistic—under those circumstances, lie detection decisions are apt to be right far less than half the time.

In general, then, lie detection accuracy should be heterogeneous—lie detection decisions should be generally correct when they are made

---

[75] *See* Levine et al., *supra* note 74, at 256. Levine, et al. effectively model deception accuracy rates knowing simply the truth bias and the base rate, suggesting that we can view each truth/lie evaluation as simply a single draw from a Bernoulli distribution with parameter θ where θ represents the truth bias.

by individuals with substantially more information than the witness and when their predispositions about whether the witness will be truthful or deceitful match reality. However, when the witness is aware of what the receiver knows about the incident, or when the receiver's bias does not match the base rate, lie detection accuracy is likely to fall well below the 50% mark. In this Part, I discuss the implications of these results in two important criminal law contexts: statements by cooperating witnesses and by defendants.

A. *Detecting Lies by Cooperating Witnesses*

Cooperating witnesses have become an essential feature of the federal criminal system. If the government believes that a defendant is providing truthful information and believes that the defendant may be of assistance in future cases, the government will offer the defendant the opportunity to plead guilty to some or all of the crimes he committed pursuant to a written cooperation agreement. At the time of the cooperator's sentencing, the government submits a letter to the court pursuant to Section 5K1.1 of the sentencing guidelines,[76] outlining the cooperator's substantial assistance provided to the government, including information on whether the cooperator participated in undercover investigations, testified at trial, or aided in other ways.

Detecting cooperator lies begins with the government. Cooperators in the federal system engage in extensive proffer sessions with the government prior to being signed up to cooperation agreements. A primary, if not the sole, decision to be made in the proffer session is whether the potential cooperator is telling the truth. This environment is one where the value of context should lead us to expect relatively good detection accuracy, particularly in cases where the government already knows a great deal about the information provided by the cooperator. Indeed, federal prosecutors express substantial confidence in the cooperation process.[77] That confidence is likely justified when the cooperator does not learn what information the government already knows. However, lie detection probably varies considerably from case to case. Ellen Yaroshefsky's 1999 survey of former Southern District of New York Assistant United States Attorneys asked about the information provided to cooperators, and at least three of the AUSAs surveyed indicated that prosecutors at least occasionally provided information to cooperators either intentionally or

---

[76] *See* U.S. SENTENCING GUIDELINES MANUAL § 5K1.1 (2007).

[77] *See* Ellen Yaroshefsky, *Cooperation with Federal Prosecutors: Experiences of Truth Telling and Embellishment*, 68 FORDHAM L. REV. 917, 932 (1999) (finding that three-quarters of prosecutors surveyed believe that the government obtain truthful information from cooperators).

accidentally.[78]

Yaroshefsky's study also provides information about bias, specifically the type of deception that law enforcement is trying to detect. By and large, prosecutors believe that cooperators frequently minimize their own criminal conduct[79] but rarely falsely implicate others.[80] The government expects cooperators to tell stories that both reduce their role in the offense from participant to witness and that subtract from the number of criminal transactions involved. If the government's presumption is correct, lie detection is likely to be very accurate in proffer sessions. On the other hand, if cooperators often falsely implicate defendants, lie detection is likely to fail. We do not know enough about base rates in this context to draw strong conclusions about whether this bias is correct or incorrect.

Once the government has decided that a cooperator is truthful, the burden of lie detection shifts to the jury. Jurors have strong, well-recognized biases about witness testimony,[81] but we do not know their biases about *cooperator* testimony. In the federal system, cooperation agreements require cooperating witnesses to either tell the truth at trial or lose the benefits of the cooperation. Prosecutors are generally confident that the truth-telling provisions work, while commentators are often skeptical about the effectiveness of these incentives.[82] However, we have little actual data about the rates at which cooperators lie.[83]

---

[78] *Id.* at 960. With respect to paid confidential informants, DOJ Guidelines strongly encourage not disclosing information to them. *See* OFFICE OF THE ATTORNEY GEN., THE ATTORNEY GENERAL'S GUIDELINES REGARDING THE USE OF CONFIDENTIAL INFORMANTS (Oct. 2003) (redacted version), *available at* http://www.fas.org/irp/agency/doj/fbi/dojguidelines.pdf. No comparable guidelines exist with respect to cooperating witnesses, but my experience as an AUSA in the Eastern District of New York reflected similar training. AUSAs were strongly discouraged from disclosing information learned from other sources to cooperating witnesses.

[79] Yaroshefsky, *supra* note 77, at 957 & n.189.

[80] *Id.* at 933.

[81] One very commonly cited result is that jurors vastly overestimate the value of eyewitness evidence. For instance, one study found that mock jurors predict an accuracy rate of about 71% for eyewitness testimony, when the actual accuracy rate was only 13%. *See* Steven Penrod, & Brian Cutler, *Witness Confidence and Witness Accuracy*, 1 PSYCHOL. PUB. POL'Y & L. 817, 819 (1995); *cf.* Gary L. Wells et al., *Accuracy, Confidence and Juror Perceptions in Eyewitness Identification*, J. APPLIED PSYCHOL. 440, 447 (1979) (jurors believe 4 out of 5 mistaken identifications). For a general summary of cognitive psychology results with respect to direct evidence, see Kevin Jon Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 247-53 (2006). However, these results generally focus on the question of evaluating honest but inaccurate testimony rather than dishonest testimony.

[82] *Compare* Yaroshefsky, *supra* note 77, at 932 (finding that three-quarters of prosecutors surveyed believe that the government obtain truthful information from cooperators) *with* George C. Harris, *Testimony for Sale: The Law and Ethics of Snitches and Experts*, 28 PEPP. L. REV. 1, 51 (2000) (expressing skepticism about the value of the truth-telling provisions in cooperation agreements).

[83] What we do know is that false cooperator testimony is involved in a large number of convictions later identified as wrongful. *See* Samuel R. Gross et al., *Exonerations in the United States 1989 through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 543-44 (2005).

Given that we lack the base rate of cooperator honesty in the abstract, we should have significant concerns about whether jurors are biased in the correct direction. One can easily imagine that jurors might exhibit a lie bias and strongly discount cooperator testimony despite the truth-telling provisions. If cooperators are generally truthful, this lie bias would lead to highly inaccurate truthfulness evaluations. The alternative situation is equally possible; jurors may tend to believe cooperators even though they usually lie. The core point is that accurate deception detection is only likely to occur when base rates align with biases, i.e. when juror expectations about truth-telling match reality. We do not know much about either base rates or biases for cooperator testimony, but we have no reason to think that juror biases regarding cooperators' credibility match the reality of this credibility with any frequency.

There is also little reason to believe that context improves jury lie detection with respect to cooperator testimony. Jurors are unlikely to have concealed, private information about the stories that cooperators tell. Jurors receive the vast majority of their information about the case from the government. Cooperating witnesses only take the stand when the government (1) already believes the story that the cooperator is telling and (2) expects that story to be consistent with the other evidence in the case. Jurors, then, will only know private information that is potentially inconsistent with the cooperator testimony when that evidence is in the hands of the defendant. Defense investigation, of course, is widely recognized as limited compared to that of the government.[84]

These conclusions do not suggest that cooperators frequently lie. As discussed above, the government preparation of cooperators is often designed to avoid exactly that outcome, and we have reason to think that it works. However, these conclusions suggest that jurors do not bring much additional lie detection ability to the table. The deception detection skills that jurors have do not apply when cooperators testify.

B. *Detecting Lies by the Defendant*

We have reason to be somewhat more optimistic about jurors' ability to classify accurately the truthfulness of defendant testimony, although the results are still mixed. Jurors are far more likely to see the

---

[84] *See* Darryl K. Brown, *The Decline of Defense Counsel and the Rise of Accuracy in Criminal Adjudication*, 93 CAL. L. REV. 1585, 1602-03 (2005); Michael McConville & Chester L. Mirsky, *Criminal Defense of the Poor in New York City*, 15 N.Y.U. REV. L. & SOC. CHANGE 581, 762 (1986-87) (reporting study results that defense counsel only interviewed witnesses and visited crimes scenes in only 4% of nonhomicide cases).

defendant cross-examined by a questioner with private information since the government will often know information about the events of the crime that has not been disclosed to the defendant.[85] The value of context also suggests that we should be more confident that the jury will correctly classify the testimony of defendants in cases that are investigated more deeply. If the government has put more time and effort into investigating the crime, more private information will be available to the government and the jury at the time of trial. For instance, in reactive cases,[86] in which a crime occurs and the government investigates it after the fact, the government's information about the case is necessarily limited in scope. In proactive cases, in which the crime and investigation occur contemporaneously, the government possesses a wide range of available evidence-gathering tools, from wiretaps to more mundane physical surveillance. In these proactive cases, the private information known to the government will be much more extensive, and the jury's deception detection will be substantially improved.[87] We should expect that jurors should be more likely in these cases to classify correctly truthful defendants as honest and lying defendants as dishonest.

While the context analysis gives some reason for optimism, the bias question leads us in the other direction. We have limited data about whether real jurors have a truth bias or a lie bias for defendants.[88] Even studies analyzing the different but related question of the effect of defendant testimony on conviction rates have found effects that are unclear.[89] We also know little about base rates in this context. We do not know whether testifying defendants are generally truthful. Again, we have little reason for optimism that base rates match biases. There is

---

85 This is one of the classic justifications for the more limited nature of discovery in criminal cases than in civil cases. *See* H. Lee Sarokin & William E. Zuckerman, *Presumed Innocent? Restrictions on Criminal Discovery in Federal Court Belie This Presumption*, 43 RUTGERS L. REV. 1089, 1090-91 (1991); William J. Brennan, Jr., *The Criminal Prosecution. Sporting Event or Quest for Truth?*, 1963 WASH. U. L.Q. 279, 289.

86 *See* Kevin Washburn, *American Indians, Crime, and the Law*, 104 MICH. L. REV. 709, 718 (2006) (distinguishing between proactive and reactive cases); Tom Lininger, *Sects, Lies, and Videotape: The Surveillance and Infiltration of Religious Groups*, 89 IOWA L. REV. 1201, 1271-72 (2004) (same).

87 Some of this evidence will be disclosed to the defendant in discovery. For instance, a defendant will receive any wiretap recording including his statement. *See* FED. R. CRIM. P. 16(a)(1)(B) (requiring disclosure of defendants' recorded statements). Physical surveillance, recordings not involving the defendant, and a myriad of other types of evidence need not be disclosed in advance.

88 I am only aware of one study even remotely on point involving real jurors. *See* Mitchell J. Frank & Dawn Broschard, *The Silent Criminal Defendant and the Presumption of Innocence: In the Hands of Real Jurors, Is Either of Them Safe?*, 10 LEWIS & CLARK L. REV. 237, 268 (2006) (finding that 59.5% of jurors disagreed with the statement "I found the defendant credible").

89 "Three studies have examined defendant testimony at trials and the results are inconclusive." *See* Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research*, 7 PSYCHOL. PUB. POL'Y & L. 622, 680 (2000).

not much reason to expect jurors' preconceived notions about the likelihood of the defendant testifying truthfully to actually match reality.

By contrast, we know substantially more about the attempts of law enforcement to detect lies told by defendants. While law enforcement officials are not consistently better in deception detection than lay jurors, they are consistently different. In particular, studies show a weaker veracity effect for law enforcement. Training in detecting deception and law enforcement experience generally reduces truth bias.[90] Some studies show that certain law enforcement officials have a substantial lie bias.[91] Lie bias is likely to be substantially stronger in the real world as a result of the specifics of law enforcement training. The primary law enforcement training methodology, the Reid Technique,[92] instructs officers to decide whether an encounter is treated as an interrogation or an interview. While "interviews" are non-accusatory and designed for information gathering, "interrogations" are overtly accusatory and "should be conducted only when the investigator is reasonably certain of the suspect's guilt."[93] This division between interrogation and interview should construct a strong lie bias in the former context but truth bias in the later context.[94] The quality of deception detection is then going to be heavily dependent on the base rate. If the initial decision to treat the encounter as an interview or interrogation is good, individuals who are likely to lie are going to be expected to lie and accuracy will be high. On the other hand, if the initial decision is poor, the base rate and bias will be mismatched and decisions about truthfulness will frequently err. This result also suggests that lie detection ability is going to be heterogeneous across law enforcement. Officers who make good initial decisions to sort witnesses into the interview and interrogation categories will make good lie-detection decisions, while those whose initial decisions are poor will succeed in lie detection at much lower rates.

In law enforcement interrogations, context would initially seem to

---

90 *See* Timothy Levine et al., *Testing the Effects of Nonverbal Behavior Training on Accuracy in Deception Detection with the Inclusion of a Bogus Training Control Group*, 69 W.J. COMM. 203, 213-15 (2005) (effect of training on truth bias); Meissner & Kassin, *supra* note 23, at 478 (police investigators substantially more likely to identify statements as deceitful than students).

91 *See* Mark G. Frank & Thomas Hugh Feeley, *To Catch a Liar: Challenges for Research in Lie Detection Training*, 31 J. APPLIED COMM. RES. 58, 65 (2003).

92 The Reid technique is named after John Reid, one of the authors of *Criminal Interrogation and Confessions*, the standard law enforcement interrogation training manual. *See* INBAU, *supra* note 16. Reid claims to have trained over 500,000 law enforcement officials in interrogation techniques since the early 1970s. *See* John E. Reid & Assoc., Inc., Interviewing & Interrogating, http://www.reid.com/training_programs/interview_overview.html (lasted visited Feb. 18, 2007).

93 *See* INBAU, supra note 16, at 7-8.

94 More accurately, it should construct a lie bias for exculpatory statements and a truth bias for inculpatory statements in the interrogation context.

allow good lie detection. Law enforcement officials will often have the capacity to generate a great deal of private information and not disclose it to the defendant. Law enforcement training on this point suggests some internalization of these norms. Inbau and Reid encourage a strong mastery of underlying facts.[95] The *FBI Law Enforcement Bulletin*, the leading law enforcement periodical,[96] frequently publishes articles on interrogations that strongly encourage extensive preparation prior to an interrogation.[97] However, the advice to law enforcement on *concealment* of evidence is less heartening. While Reid and Inbau caution against "revealing to the suspect at the outset of the interrogation all the specific evidence that implicates him,"[98] other sources take a different view, including some which encourage disclosing evidence to suspects in order to evoke a confession.[99]

In practice, though, there is further reason for pessimism about the accuracy of lie detection during interrogation. First, we lack information about the reality of law enforcement training. The work that has been done on this subject suggests that law enforcement officers receive limited training on deception detection.[100] Second, studies on the practices of law enforcement interrogation show that evidence is frequently disclosed to defendants. Richard Leo's seminal 1996 study of real-world interrogations found that in 85% of interrogations, officers began by confronting the suspect with at least some of the evidence against him.[101] Officers who surrender their informational advantage are likely to impair their accuracy

---

95 *Id.* at 12.

96 The LEB is published by Law Enforcement Communication Unit of the FBI Academy and according to its website, is "the most widely read law enforcement publication in the world . . . with an estimated readership of 200,000" readers in the law enforcement community. *See* The FBI Academy: Law Enforcement Communication Unit, http://www.fbi.gov/hq/td/academy/lecu/lecu.htm (last visited July 2, 2007).

97 *See, e.g.*, Andre B. Simons & Brian Parsi Boetig, *The Structured Investigative Interview*, FBI L. ENFORCEMENT BULL., June 2007, at 9, 10; Randy Bowling & Dave Resch, *Child Pornography Cases*, FBI L. ENFORCEMENT BULL., Mar. 2005, at 1, 2; Scott O'Neal, *Interviewing Self-confident Con Artists*, FBI L. ENFORCEMENT BULL., Mar. 2002, at 16, 19.

98 INBAU, *supra* note 16, at 235.

99 *Compare* Bowling & Resch, *supra* note 97, at 6 (suggesting that officers "show case facts and refer to real or implied evidence to convince the suspect of the futility of denial"); David Vessel, *Conducting Successful Interrogations*, FBI L. ENFORCEMENT BULL., Oct. 1998, at 1, 4 (encouraging confronting subjects "with the facts and issues surrounding the incidents") *with* Michael R. Napier & Susan H. Adams, *Criminal Confessions: Overcoming the Challenges*, FBI L. ENFORCEMENT BULL., Nov. 2002, at 9, 12 (instructing interrogators on how to ask questions without disclosing information).

100 A recent survey of Texas law enforcement found that on average, officers participated in one lecture course on interviewing and in most cases, the training only involved interrogations of suspects rather than interviews of witnesses or victims. *See* Lori H. Colwell et al., *The Training of Law Enforcement Officers in Deception Detection: A Survey of Current Practices and Suggestions for Improving Accuracy*, 9 POLICE Q. 275, 282 (2006).

101 *See* Richard A. Leo, *Inside the Interrogation Room*, 86 J. CRIM. L. & CRIMINOLOGY 266, 277 (1996).

substantially.

CONCLUSION

There remains a great deal that we do not know about lie detection that should matter to the legal system, but we know more than the legal literature usually reflects. While the newest results on lie detection support the now-traditional view in legal academia that demeanor is not a valuable tool in making credibility decisions, they undermine the further conclusion that accurately detecting lies is impossible and, as a result, we should view credibility decisions by juries as no better than a coin flip. When bias and context are incorporated, we should expect that deception detection accuracy is highly heterogeneous and varies substantially based on the situation. We just do not yet know enough about bias, base rates, and the value of context to say whether the social science evidence supports the currently skeptical view on legal lie detection.

We also do not know whether some common policy recommendations survive this new evidence. For instance, commentators have criticized the fact that jurors are instructed to rely on demeanor. While there is reason to be concerned about these instructions, there is significant evidence that jurors are not heavily swayed by jury instructions. Studies of jurors show relatively limited comprehension and retention of jury instructions.[102] Similarly, we know little about the relative accuracy of judges and juries. For instance, we do not know whether the group structure of a jury makes them better lie detectors than judges. Studies are mixed on whether groups are more successful at detecting lies than individuals.[103] We also do not know whether the fear of reversal creates incentives for judges to

---

[102] *See* Shari Seidman Diamond & Mary R. Rose, *Real Juries*, 1 ANN. REV. L. & SOC. SCI. 255, 272 (2005); Alan Reifman et al., *Real Jurors' Understanding of the Law in Real Cases*, 16 LAW & HUM. BEHAV. 539, 550 (1992). This is a nice example of a situation in which a cognitive error may help the legal system avoid a problem that it would otherwise face. *See* Chris William Sanchirico, *Evidence, Procedure, and the Upside of Cognitive Error*, 57 STAN. L. REV. 291, 364 (2004).

[103] Two studies have analyzed the relative abilities of individuals and groups in deception detection with slightly different results. *Compare* Mark Frank et al., *Individual and Small Group Accuracy in Judging Truthful and Deception Communication*, 13 GROUP DECISION & NEGOTIATION 45 (2004) (finding that groups and individuals judged truthful statements correctly at similar rates but that groups were more accurate in judging false statements), *with* E. Park et al., *Group and Individual Accuracy in Deception Detection*, 19 COMM. RES. REP. 99 (2002) (finding no difference between group and individual accuracy for either true or false statements). These mixed results are comparable to the findings in the broader literature on the role cognitive bias in the group versus individual contexts. *See* N. Kerr et al., 103 PSYCHOL. REV. 687, 713 (1996) (concluding that there is no simple empirical answer to whether groups or individuals are likely to make biased judgments).

make better decisions than juries who are likely immune to this concern. We do not know what effect motivation has on accuracy.[104] As a result, we do not know whether this evidence supports or undermines the notion of appellate deference. Perhaps the best lesson to be drawn from these new results is the value of caution in the face of new social scientific evidence. Wellborn's article has been correctly praised for its sensitivity in the use of empirical evidence,[105] while more recent discussions have drawn stronger conclusions on weaker evidence. In this context, our bias should be toward caution.

---

[104] DePaulo's meta-analysis reviewed a number of studies that provided incentives to lie successfully and found increases in some of the traditional deception cues. While unmotivated liars do not differ significantly in their eye contact, motivated senders in fact do avert their gaze when lying more frequently than they do when telling the truth. Furthermore, motivated liars appeared tenser and more likely to increase their vocal pitch. DePaulo concludes that "for studies in which there was no special incentive for succeeding, cues to deception were generally weak. Overall, the size of the effects increased somewhat when some incentive was provided." DePaulo et al., *supra* note 42, at 97. This effect was increased when the cues related to transgressions by the liar or involved subjects that related to the speaker's self-identity. *Id.* at 104.

[105] Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1539-41 (2001).

APPENDIX ON LEVINE/PARK RESULTS

This appendix provides a formal statement of the claims outlined *supra* in the text accompanying note 74. Below I present a modified version of the Levine/Park results in support of the conclusions claimed in the main text.[106]

Adopting their notation, define $P(H)$ as the probability that a message is judged as honest, $P(-H) = 1 - P(H)$ as the probability a message is judged as a lie, $P(T)$ as the probability that a message is true (called the "truth base rate"), and $P(-T) = 1 - P(T)$ as the probability a message is false. Further let $P(H \mid T)$ and similar constructions have their standard meaning as conditional probabilities. Subjects are defined as exhibiting a veracity effect if they more accurately classify true statements than false ones, represented formally as $P(H \mid T) > P(-H \mid -T)$.

Levine/Park suggest that truth accuracy, lie accuracy, and total accuracy can be modeled using a simple Bayesian construction. That is, truth accuracy is the probability that a message is both true and judged as honest.

$$P(T)P(H \mid T) = P(T \bigcap H) = P(T \mid H)P(H)$$

Similarly, lie accuracy is the probability that the message is false and is judged as dishonest.

$$P(-T)P(-H \mid -T) = P(-T \bigcap -H) = P(-T \mid -H)P(-H)$$

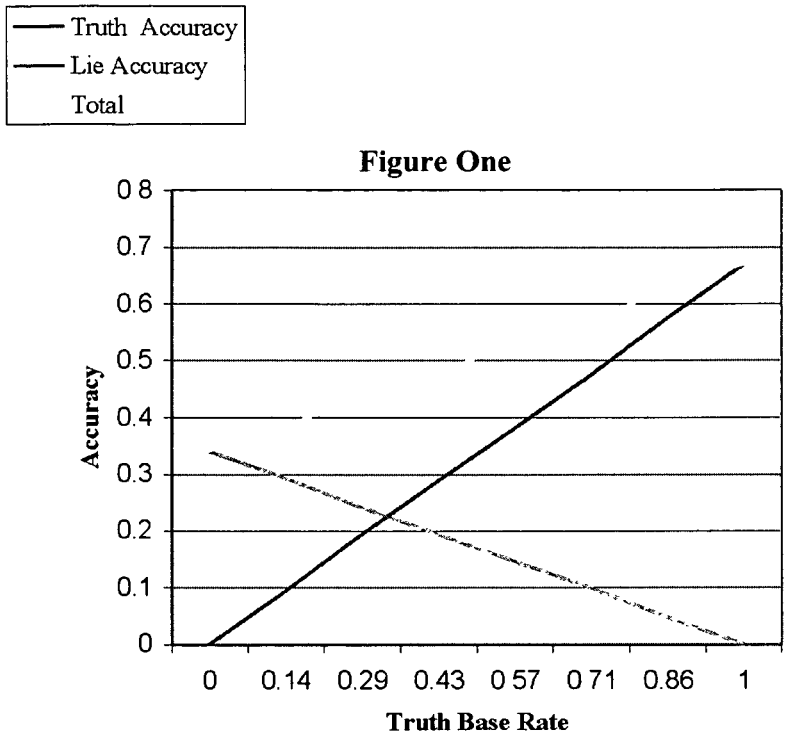
Levine/Park demonstrate that accuracy rates can be modeled as a simple linear function of the truth base rate, with slope $P(H \mid T) - P(-H \mid -T)$ and intercept $P(-H \mid -T)$.

$$P(H \bigcap T) + P(-H \bigcap -T) = P(H \mid T)P(T) + P(-H \mid -T)P(-T) = P(-H \mid -T) + (P(H \mid T) - P(-H \mid -T))P(T) \text{ (Equation 1).}$$

Figure One reproduces a modified version of Figure 1 from Park/Levine 2004 showing the data in support of their result.

---

106 *See* Levine et al., *supra* note 74, at 243; Park & Levine, *supra* note 74, at 201; Levine et al., *supra* note 72, at 125.



The main text makes two primary claims. First, the concern of the legal system should be the fit between bias and the base rate, i.e. whether the accuracy of the base rate matches well with observer bias. Equation 1 and Figure 1 demonstrate that total accuracy increases as the base rate increases (is upward-sloping) if and only if there is a veracity effect—truth judgments are more accurate than lie judgments. This leads directly to the notion of "fit" described in the main text. Overall accuracy rates are relatively high if observers are truth-biased when they are confronted with mostly truths and lie-biased if they are confronted with mostly lies.

Second, the main text claims that background beliefs about truthfulness matter more than the information received by observers. Park/Levine's result that accuracy is a linear function of the base rate $P(T)$ also implies that the relevant conditional probabilities, e.g. $P(-H \mid -T)$, are independent of the base rate. This leads easily to the conclusion that $P(H \mid T) \propto P(H)$. That is, the truth bias, $P(H)$, is highly correlated with $P(H \mid T)$, the conditional probability that a statement will be judged as honest given that it is in fact true. The prior expectation of honesty is more important than the data received.