Polygraph Evidence: Post-Daubert

Paul C. Giannelli

Follow this and additional works at: https://repository.uchastings.edu/hastings_law_journal

Part of the Law Commons

Recommended Citation
Available at: https://repository.uchastings.edu/hastings_law_journal/vol49/iss4/1

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Law Journal by an authorized editor of UC Hastings Scholarship Repository.
Polygraph Evidence: Post-\textit{Daubert}

\textit{by}

PAUL C. GIANNELLI

Table of Contents

Introduction ........................................................................................................... 896

I. \textit{Pre-Daubert} Decisions .............................................................................. 900
A. 1970's Cases ................................................................................................. 900
B. 1980's Cases ................................................................................................ 901

II. Polygraph Technique ...................................................................................... 903
A. Theory ........................................................................................................... 903
B. Instrument ..................................................................................................... 904
C. Procedure ....................................................................................................... 904
D. Examiner ........................................................................................................ 905
E. Types of Examinations .................................................................................. 906
F. Types of Formats ........................................................................................... 906
G. Methods of Evaluation .................................................................................. 909
H. Quality Control Procedures .......................................................................... 910

III. Validity ........................................................................................................... 910
A. Purpose of Study .......................................................................................... 911
B. Field & Laboratory Studies .......................................................................... 911
C. Selected Studies ............................................................................................ 912
D. Computer Scoring ........................................................................................ 914
E. Countermeasures ........................................................................................... 914
F. Friendly Polygrapher Issue ............................................................................ 915

IV. Evidentiary Standard ...................................................................................... 915
A. \textit{Frye} Standard .......................................................................................... 916
B. \textit{Daubert} Standard .................................................................................... 917

V. Polygraph Cases ............................................................................................. 919
A. Lack of Research .......................................................................................... 919
B. Specificity of Procedure ................................................................................. 921
C. Examiner Qualifications ................................................................................. 923

Conclusion .......................................................................................................... 923

*Albert J. Weatherhead III & Richard W. Weatherhead Professor of Law, Case Western Reserve University. J.D., 1970, LL.M., 1975, University of Virginia; M.S. in Forensic Science, 1973, George Washington University. I would like to thank Dale E. Olsen, Ph.D., for his comments on an earlier draft of this article.
Aldrich Ames, the CIA agent who sold secrets to the Soviet Union, "did not 'beat' the polygraph tests; instead, he simply talked the CIA out of believing the results, which showed deception every time he took the test. Since the false positive error rate is approximately 10 percent . . . Ames could have argued with some force that the positive results were false."

Introduction

After the United States Supreme Court's decision in Daubert v. Merrell Dow Pharmaceuticals, Inc. the Fifth Circuit in United States v. Posado stated that "the rationale underlying this circuit's per se rule against admitting polygraph evidence did not survive Daubert." The court went on to comment that "[t]here can be no doubt that tremendous advances have been made in polygraph instrumentation and technique in the years since Frye [1923] . . . Current research indicates that, when given under controlled conditions, the polygraph technique accurately predicts truth or deception between seventy and ninety percent of the time." The court, however, limited its ruling, commenting that:

[W]e do not now hold that polygraph examinations are scientifically valid or that they will always assist the trier of fact . . . We merely remove the obstacle of the per se rule against admissibility which was based on antiquated concepts about the technical ability of the polygraph and legal precepts that have been expressly overruled by the Supreme Court.

The Ninth and Tenth Circuits adopted the same position. Two other circuits had embraced this position prior to Daubert; the Seventh Circuit had long abandoned the per se rule, and in 1989 the Eleventh

3. 293 F. 1013 (D.C. Cir. 1923).
4. 57 F.3d 428 (5th Cir. 1995).
5. Id. at 429.
6. Id. at 434.
7. Id. See also United States v. Pettigrew, 77 F.3d 1500, 1515 (5th Cir. 1996) (trial court did not abuse its discretion in excluding polygraph evidence).
8. See United States v. Cordoba, 104 F.3d 225, 227-28 (9th Cir. 1997) (noting that its former per se rule of exclusion is inconsistent with Daubert); United States v. Call, 129 F.3d 1402, 1405 (10th Cir. 1997) (same).
9. See United States v. Pulido, 69 F.3d 192, 205 (7th Cir. 1995) ("Our decisions acknowledge the considerable scientific and legal debate over polygraph testing and recognize that a trial court deciding whether to admit polygraph evidence 'must engage in a delicate
Circuit followed suit.\textsuperscript{10}

Other circuits took a more cautious approach.\textsuperscript{11} Indeed, the Second Circuit wrote that “the ‘legal Pandora’s box’ which the Fifth Circuit opened in \textit{United States v. Posado}... is not yet agape in this Circuit.”\textsuperscript{12} In addition, the state cases have not been influenced by this development; several have reaffirmed the rule of categorical inadmissibility,\textsuperscript{13} even while citing \textit{Daubert}.\textsuperscript{14} Other courts have continued to require stipulations.\textsuperscript{15} Nevertheless, \textit{Daubert} has dramatically changed balancing of many factors including probative value, prejudicial effect, confusion of the issues, misleading the jury, and undue delay.”\textsuperscript{16} (quoting \textit{United States v. Olson}, 978 F.2d 1472, 1480 (7th Cir. 1992)).

10. \textit{See United States v. Piccinonna}, 885 F.2d 1529 (11th Cir. 1989). \textit{See also} Barker v. Jackson Nat’l Life Ins. Co., 896 F. Supp. 1159, 1161-62 (N.D. 1995) (“Surprisingly, in the six years since \textit{Piccinonna}, the Eleventh Circuit appears not to have had opportunity to elaborate on the decision, other than to decide that it was not to be applied retroactively.”).

11. \textit{See} Conti v. Commissioner, 39 F.3d 658, 663 (6th Cir. 1994) (noting that “unilaterally obtained polygraph evidence is almost never admissible under Evidence Rule 403”); United States v. Sherlin, 67 F.3d 1208, 1216-17 (6th Cir. 1995) (“A privately commissioned polygraph test, which was unknown to the government until after its completion, is of extremely dubious probative value.”); United States v. Williams, 95 F.3d 723, 729 (8th Cir. 1996) (defendant failed to lay a foundation for polygraph admissibility).


13. \textit{See, e.g.}, People v. Gard, 632 N.E.2d 1026, 1032 (Ill. 1994) (“\textit{T}he use of polygraph evidence... is no less repugnant to and no less an affront to the integrity of the judicial process when the examination has been given to a witness... than it is when the examination has been given to the defendant himself.”).

14. \textit{See, e.g.}, State v. Porter, 698 A.2d 739 (Conn. 1997) (adopting \textit{Daubert} and excluding polygraph results under Rule of Evidence 403); State v. Beard, 461 S.E.2d 486, 493 (W. Va. 1995) (“\textit{W}e remain convinced that the reliability of such examinations is still suspect and not generally accepted within the relevant scientific community. Therefore, any speculation that our position in \textit{Frazier} regarding polygraph admissibility is in question due to the \textit{Daubert}/\textit{Wilt} rulings is put to rest today.”).

15. \textit{See, e.g.}, State v. Webber, 918 P.2d 609, 619-20 (Kan. 1996) (citing “the ironclad rule that the results of such examinations are inadmissible absent a stipulation between the parties”); State v. Cosby, 927 P.2d 638, 642 (Utah 1996) (reaffirming the need for a stipulation).
the legal landscape. In this altered climate, some district courts have admitted polygraph evidence— for example, *United States v. Crumby* 17 and *United States v. Galbreth*. 18 In short, the “trend appears to be moving toward admissibility.” 19

Recently, in *United States v. Scheffer*, 20 the United States Supreme Court held that the per se rule of exclusion is not unconstitutional. The Court’s ruling was not surprising, given the limited nature of the constitutional right at issue. 21 What was unexpected was Justice Kennedy’s concurring opinion, in which three other justices joined. He wrote that:

I doubt, though, that the rule of per se exclusion is wise, and some later case might present a more compelling case for introduction of the testimony than this one does. Though the considerable discretion given to the trial court in admitting and excluding scientific evidence is not a constitutional mandate, there is some tension between that rule and our holding today. And, as Justice Stevens points out [in dissent], there is much inconsistency between the Government’s extensive use of polygraphs to make vital security determinations and the argument it makes here, stressing the inaccuracy of these tests. 22

In addition to the admissibility and constitutional issues, 23 the

---

20. 118 S. Ct. 1261 (1998). The U.S. Court of Appeals for the Armed Forces had held that Military Rule 707, which excludes polygraph evidence, was unconstitutional as applied in that case. See *United States v. Scheffer*, 44 M.J. 442, 445 (C.C.A.F. 1996). In that court’s view:

A per se exclusion of polygraph evidence offered by an accused to rebut an attack on his credibility, without giving him an opportunity to lay a foundation under Mil. R. Evid. 702 and Daubert, violates his Sixth Amendment right to present a defense. We limit our holding to exculpatory evidence arising from a polygraph examination of an accused, offered to rebut an attack on his credibility.

Id.

21. In the Court’s view, evidence “rules do not abridge an accused’s right to present a defense so long as they are not ‘arbitrary’ or ‘disproportionate to the purposes they are designed to serve.’ Moreover, we have found the exclusion of evidence to be unconstitutionally arbitrary or disproportionate only where it has infringed upon a weighty interest of the accused.” *Scheffer*, 118 S. Ct. at 1264 (citations omitted).
22. *Id.* at 1269 (citations omitted).
23. Several constitutional arguments have been advanced to support the admissibility of polygraph evidence. First, defendants have argued, as in *Scheffer*, that the constitutional right to present a defense includes the right to introduce favorable polygraph results. Second, defendants also have argued for a constitutional right to impeach prosecution witnesses who have failed government-administered examinations. See 1 GIANNELLI & IMWINKELRIED, *supra* note 2, § 8-3(D) (discussing constitutional arguments).
polygraph technique is worth examining for other reasons. First, the polygraph cases provide a rich history of how courts have grappled with scientific evidence. No other scientific technique has been subject to judicial review for such a long period; it has been 75 years since Frye was decided. For the most part, other techniques came to the courts only once: they were accepted or rejected, albeit often over a tortuous path. In contrast, the polygraph fades into the background and then erupts back to the forefront as in the 1970's, 1980's, and now the post-Daubert cases.

Second, in many ways the courts' treatment of polygraph evidence has differed radically from their approach to other scientific techniques. The paradoxes are striking. For example, many jurisdictions admit polygraph evidence upon stipulation, even though the stipulation does nothing to enhance the reliability of the evidence, which is the principal reason for exclusion.24 Similarly, courts have admitted polygraph evidence in suppression hearings, sentencing hearings, motions for new trial proceedings, and prison disciplinary hearings.25 Moreover, some courts have enforced plea bargains based on polygraph evidence.26

Third, the extensive use of the polygraph by the government raises another paradox—especially for prosecutors challenging its admissibility. In 1996, the Department of Defense (DoD) conducted 12,548 polygraph examinations. Sixty-three percent (7,945) involved the DoD Counterintelligence-Scope Polygraph (CSP) Program. The other categories include 21.5% criminal investigations (2,696), 4.6% exculpatory (579), and 10.6% miscellaneous (1,328). The latter includes security investigations, counterintelligence and intelligence operations, and assistance to non-DoD federal agencies.27 Further, the Department of Defense Polygraph Institute (DoDPI) trains 100 federal examiners a year in a masters level program28 and allocates funds for polygraph research.

24. See id. at § 8-3(B) (discussing admissibility by stipulation).
25. See id. at § 8-3(E) (discussing admissibility in proceedings other than trials).
26. In some cases prosecutors have gone beyond stipulating to the admissibility of test results and have agreed to dismiss charges if the defendant passes a polygraph examination. See id. at § 8-4(C) (discussing pretrial agreement cases). See also United States v. Santiago-Gonzales, 66 F.3d 3, 6 n.1 (1st Cir. 1995) (polygraph used to measure defendant's requirement to be truthful under a plea agreement).
27. See DEPARTMENT OF DEFENSE, POLYGRAPH PROGRAM, ANNUAL REPORT TO CONGRESS 1 (1996) [hereinafter DODPI 1996 ANNUAL REPORT].
28. The DoDPI trains all federal polygraph examiners. See also Ronald M. Furgerson, Perspectives on Polygraphs: A Guide to Survival, 21 POLYGRAPH 164, 164 (1992) (from 1977 to 1992, 115 FBI agents attended the DoD Polygraph Institute (or its predecessor); they have conducted over 40,000 polygraph exams).
I. Pre-Daubert Decisions

The admissibility of polygraph evidence was first considered and rejected in *Frye v. United States*, the 1923 case in which the D.C. Circuit established the "general acceptance" test for the admissibility of scientific evidence. According to the court, a precursor of the modern polygraph had not gained general acceptance in the fields of psychology and physiology. From *Frye* until the 1970's, the courts overwhelmingly rejected polygraph evidence.

A. 1970's Cases

In 1972, several trial courts departed from nearly fifty years of precedent and admitted the results of unstipulated polygraph examinations. In *United States v. Ridling*, a federal district court found that "the theory of the polygraph is sound" and "directly relevant" to the issue being litigated (perjury). The court further held that the results of a polygraph examination conducted by a court-appointed expert would be admissible under certain conditions. In *United States v. Zeiger*, another district court held that the "polygraph has been accepted by authorities in the field as being capable of producing highly probative evidence in a court of law when properly used by competent, experienced examiners." In addition, in *People v. Cutter*, a California trial court admitted polygraph evidence during a suppression hearing after finding that the "polygraph now enjoys general acceptance among authorities, including psychologists and researchers ... as well as polygraph examiners." In the next year, a third district court, in an extensive opinion, stated that polygraph results would have been admitted if the court had not been bound by precedent. Several other

29. 293 F. 1013 (D.C. Cir. 1923).
30. The instrument used in *Frye* measured only one physiological response (blood pressure), whereas the modern polygraph measures respiration and galvanic skin resistance in addition to blood pressure. The machine in *Frye* also did not record continuously, as does the modern machine. The technique also has been improved through the development of control questions, the pretest interview, and stimulation methods.
31. See *Frye*, 293 F. at 1014.
33. *Id.* at 95.
35. *Id.* at 690.
37. *Id.* at 2134.
courts also admitted polygraph evidence at this time, in civil cases, and, under special circumstances, in criminal litigation. In 1974, the Supreme Judicial Court of Massachusetts admitted polygraph evidence without stipulation in Commonwealth v. A Juvenile. In 1975, the New Mexico Supreme Court in State v. Dorsey held that polygraph results are admissible if (1) the operator was qualified, (2) the testing procedures were reliable, and (3) the test of the particular subject was valid. An Ohio trial court in 1977 and a New York trial court in 1979 reached similar results.

The trend in favor of admissibility, which these cases seemed to forecast, never developed. Zeiger was reversed per curiam, while Ridling and Cutter were never appealed, thus precluding appellate approval. Nevertheless, the judicial approach to polygraph evidence seems to have been altered by these decisions and the attention that they received in the legal literature. In particular, after these decisions were rendered a number of courts admitted polygraph results upon stipulation. By the end of the decade the cases could be divided into three groups. The first consisted of those courts that adhered to the traditional position, holding polygraph evidence per se inadmissible. A second group of courts admitted polygraph evidence upon stipulation. Finally, a few courts entrusted the admissibility of polygraph evidence to the discretion of the trial court.

B. 1980's Cases

Developments in the 1980's were mixed. In 1982, a California ap-

---

41. 313 N.E.2d 120 (Mass. 1974).
42. 539 P.2d 204, 205 (N.M. 1975).
44. See People v. Daniels, 422 N.Y.S. 2d 832, 837 (Sup. Ct. 1979).
47. The leading stipulation case, State v. Valdez, 371 P.2d 894 (Ariz. 1962), was decided in 1962 but most of the decisions adopting Valdez were decided in the late 1970's and 1980's.
48. See GIANNELLI & IMWINKELRIED, supra note 2, § 8-3(A).
49. Id. at § 8-3(B).
50. Id. at § 8-3(C).
pellate court opened the door to the possible admission of polygraph evidence. In 1983, New Mexico adopted Evidence Rule 707, which permits the admission of polygraph evidence. In 1986, the Idaho Supreme Court ruled polygraph evidence admissible in proceedings under the Child Protective Act. In 1987, the Court of Military Appeals ruled polygraph evidence admissible, as did the Eleventh Circuit in 1989.

There were also major setbacks. The California legislature closed the door opened by Witherspoon, enacting a statute that allows admission only of stipulated polygraph results. The Supreme Judicial Court of Massachusetts, which had decided a landmark case admitting polygraph evidence without stipulation, abruptly changed direction and excluded polygraph evidence. In 1991, Military Rule of Evidence 707 abrogated the Gipson decision, making polygraph results per se inadmissible. This was the basis of Scheffer, in which the United States Supreme Court upheld the Military Rule of Evidence. In addition, several courts that had recognized the admissibility of stipulated polygraph results retreated, reverting back to the per se exclusion rule.

The Wisconsin Supreme Court’s decision in State v. Dean may be the most significant. That court jettisoned its earlier opinion after the Seventh Circuit held in McMorris v. Israel that the prosecution in a stipulation state was constitutionally required to justify its refusal to stipulate.

54. United States v. Piccinonna, 885 F.2d 1529 (11th Cir. 1989) (en banc).
55. CAL. EVID. CODE § 351.1.
56. See note 41 and accompanying text (discussing A Juvenile).
60. 307 N.W.2d 628, 653 (Wis. 1981).
61. 643 F.2d 458 (7th Cir. 1981).
62. Although stipulated polygraph results were admissible under Wisconsin law at that time, the prosecutor, without offering any reasons, refused to stipulate. In granting habeas corpus relief, the Seventh Circuit wrote: “Where credibility is as critical as in the instant case, the circumstances are such as to make the polygraph evidence materially exculpatory within the meaning of the Constitution.” Id. at 462. The court, however, rested its decision on narrower grounds: The prosecution's refusal to stipulate without offering a valid ground for the refusal deprived the defendant of due process. “From all that appears, [the prosecutor] was
Finally, Congress passed the Employee Polygraph Protection Act in 1988. With limited exceptions, the Act prohibits the use of polygraph tests for preemployment screening or during the course of private employment. Governmental entities are exempt. Although the Act does not apply in criminal cases, it had the effect of putting many polygraphers out of business, including many incompetent ones.

In 1993, the Supreme Court decided *Daubert*, and the lower courts, once again, were forced to address polygraph evidence. To understand the post-*Daubert* cases, some appreciation of the fundamental elements of polygraph testing is required.

**II. Polygraph Technique**

**A. Theory**

Modern polygraph procedures developed over a long time, commencing around the turn of the century. The most common polygraph examination today, the control question technique (CQT), is based upon two premises: (1) the psychological stress caused by the fear of detection produces involuntary physiological responses in the autonomic nervous system, and (2) a polygraph examiner, based on these responses as recorded by a polygraph instrument, can detect this fear.

acting solely for tactical reasons in the belief that a test would not be helpful to his case. If the prosecutor refuses and states reasons, it then becomes the duty of the court to determine whether the reasons offered rise above the purely tactical considerations present in a given case." *Id.* at 466. Chief Justice Rehnquist characterized *McMorris* as a "dubious constitutional holding." *Israel v. McMorris*, 455 U.S. 967, 970 (1982) (Rehnquist, J., dissenting from denial of certiorari).


64. The technical discussion in this section is taken from 1 GIANNELLI & IMWINKELRIED, *supra* note 2, ch. 8. See also 1 DAVID FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE ch. 14 (1996).

65. A different examination, the "Guilty Knowledge Test" (GKT) or "Concealed Knowledge Test," may be used when important information about a crime has not been disclosed to the public. Consequently, only the perpetrator, and not other suspects, will possess this knowledge and react to it during the test. The theory of this test differs markedly from the CQT, which is a deception test; the GKT is an "information" test. The requirement of concealed information, however, greatly limits the use of this examination. *See* David C. Raskin, *The Polygraph in 1986: Scientific, Professional and Legal Issues Surrounding Application and Acceptance of Polygraph Evidence*, 1986 UTAH L. REV. 29, 31-32 ("[I]t is employed infrequently because the special information necessary to construct a valid concealed information test is typically lacking.")

66. A report by the Office of Technology Assessment (OTA) explained:

The basic theory of polygraph testing is only partially developed. The testing proc-
B. Instrument

The physiological responses used in polygraph testing are changes in blood pressure-pulse, respiration, and galvanic skin resistance. The polygraph instrument simultaneously and continuously measures and records these physiological reactions on a graph or chart (polygram).\(^67\) A quality polygraph instrument can accurately measure and record these responses. The modern instruments are lap top computers. The instrument, however, detects neither deception nor the fear of detection; it provides only a recording of physiological responses. It is the examiner who, based on these recordings, infers deception.

C. Procedure

The CQT involves several steps, the most important of which are the pretest interview, the examination of the subject while attached to the instrument, and chart interpretation. A post-test interview is considered important by some authorities.\(^68\) However, its principal function is to elicit a confession from subjects considered deceptive, and it is quite effective for this purpose.

The pretest interview serves a variety of critical functions. First, it is used to acquaint the subject with the effectiveness of the technique; this will allay the apprehensions of the truthful subject and stimulate the deceptive subject's concern about the prospect of detection. Second, the interview is used to assess the suitability of the subject for testing. The examiner may be alerted to some condition, such as a physical ailment, low intelligence, or the use of medication, that may affect the test results. Third, test questions are formulated with the

---

\(^67\) Blood pressure and pulse are measured by a sphygmomanometer (blood pressure cuff) that is placed on the subject's arm, respiration is measured by pneumograph tubes fastened around the subject's abdomen and chest, and galvanic skin response is measured by electrodes attached to the subject's fingertips.

\(^68\) See John E. Reid & Fred E. Inbau, Truth and Deception 4 (2d ed. 1977). Unlike the other stages, the examiner need not remain objective in the post-test interview.
subject's assistance during the interview. The examiner must define and clarify the issues of the test.

D. Examiner

The examiner's role cannot be overstated, because it is the examiner who decides whether there is sufficient indication of deception. The examiner's expertise is critical in (1) determining the suitability of the subject for testing, (2) formulating proper test questions, (3) establishing the necessary rapport with the subject, (4) detecting attempts to mask or create chart reactions, or other countermeasures, (5) stimulating the subject to react, and (6) interpreting the charts. Even though the examiner is the linchpin of the procedure, proponents acknowledge that there are serious problems on this score. Dr. David Raskin has commented that "a substantial proportion of those who conduct tests in the public and private sectors lack adequate training and competence." Others have observed that "[p]olygraph examiners in the United States, as a whole, are poorly trained."

The American Polygraph Association (APA) accredits polygraph schools. The private schools conduct a seven-week course,

69. The OTA report stated:
   [T]he polygraph is not simply a machine or instrument that signals whether a person is being truthful or deceptive. The instrument cannot itself detect deception. A polygraph test is very complex and depends heavily on the interaction between the examiner and the individual being tested, and requires that the examiner infer deception or truthfulness based on a comparison of the person's physiological responses to various questions. The quality of the questions asked depends in part on what information the examiner already has about the person being questioned. OTA REPORT, supra note 66, reprinted in 12 POLYGRAPH at 196 (statement of John Gibbons, Director of Office of Technology Assessment).

70. Raskin, supra note 65, at 66-67. "As bad as the situation is in the federal sector, it is generally worse in the local law enforcement agencies and in the private sector." Id. at 68.


72. The APA was founded in 1966. It publishes a quarterly journal entitled Polygraph and a bimonthly newsletter. "The American Polygraph Association (APA) is an international association which represents the polygraph profession in programs which establish standards of ethical practice, techniques, instrumentation, and training. The APA has approximately 2,000 members, of which the majority are in the United States." Polygraph: Issues and Answers, 25 POLYGRAPH 134, 145 (1996).

73. The APA has accredited fourteen polygraph schools. One school is in Canada (Canadian Police College Polygraph School). The others are located in the United States, including two government-sponsored schools: DoD Polygraph Institute and the Texas De-
while the DoD Polygraph Institute runs a 14-week course. Twenty-nine states have licensing statutes.

E. Types of Examinations

Polygraph examinations can be divided into two categories based on the purpose of the examination. One type focuses on a specific incident, typically a crime or other occurrence involving misconduct. The second type is a "background" examination, involving past conduct over a period of time. The typical pre-employment screening test is an example. Although the use of this type of examination has been greatly limited by the 1988 federal act, screening tests are still used extensively by the federal, state, and local governments, especially in law enforcement. Indeed, in "case studies carried out in Illinois, Ohio, Maryland and Florida it was found that out of 3,576 police applicants given pre-employment polygraph examinations, 58% (2,068) were identified as high risk candidates for police work in that they had behavioral histories of involvement in activities such as felony thefts, burglaries, robberies, the use and sale of illegal drugs, bribery, car thefts, and various sexual offenses." Similar tests are used in the intelligence and national security communities.

F. Types of Formats

Examiners can use several different types of formats for investigating a specific incident. The early examiners used the "relevant-irrelevant question" test (RIT). Relevant questions concern the subject matter under investigation. For example: "Did you take $100 from your employer's safe?" Irrelevant or neutral questions are used to ob-

---

74. See infra note 176 and accompanying text.
75. See Polygraph: Issues and Answers, supra note 72, at 142.
76. Id. at 135 ("According to a recent survey over 60% of the large police departments in the United States use polygraph testing in the pre-employment screening process. Police agencies have found that polygraph testing is a very effective means by which to identify high risk candidates.").
77. The DoD 1994 report stated: "During fiscal year 1994, DoD investigations obtained unique and significant information from interviews conducted with the aid of the polygraph. In all illustrated instances, the polygraph examination process produced significant security or criminal information which would not otherwise have been secured for the specific investigation. The polygraph examination process was also valuable in helping to establish the innocence of persons charged with serious infractions." DEPARTMENT OF DEFENSE POLYGRAPH PROGRAM, 1994 REPORT TO CONGRESS, reprinted at 24 POLYGRAPH 19 (1995). See also DEPARTMENT OF DEFENSE POLYGRAPH PROGRAM, 1993 REPORT TO CONGRESS, reprinted at 23 POLYGRAPH 61 (1994).
tain a subject's normal truthful reactions and chart tracings. Examples of irrelevant questions are: "Is your name [subject's name]?" and "Are you over 21 years of age?" The response to the relevant questions are compared with the subject's response to irrelevant questions. The main criticism of the RIT in this context is its underlying assumption that an innocent person will not react to the relevant questions.\(^7\)

The control question test was developed in the 1940's as an improvement of the relevant-irrelevant exam.\(^7\) Control questions concern "an act of wrongdoing of the same general nature as the main incident under investigation, and one to which the subject, in all probability, will lie or to which his answer will be of dubious validity in his own mind."\(^8\) For this reason, they are sometimes called "probable lie" questions and are typically general and vague—for example, "Before the age of 25, did you ever steal anything?" Control questions are designed as a stimulus for the truthful subject. Generally, the truthful person will respond more to the control questions than to the relevant questions because they represent a greater threat to that person. For the same reason, the deceptive person will respond more to the relevant questions than to the control questions. Therefore, the subject's comparative responses to the control and relevant questions are the key in the CQT.\(^8\) A better name would be the "comparison" question test because the CQT does not involve "controls" in the scientific sense of that word.

Questions are formulated to elicit either a "yes" or "no" response. There are no surprise questions; the examiner reviews the questions with the subject during the pretest interview to ensure understanding. The examination typically consists of ten to twelve questions. While the subject knows the questions, she does not know the order in which they will be asked. The examination lasts several minutes and is repeated at least one more time; often two or three more examinations are conducted.\(^8\) Formulating adequate control questions is not an easy

\(^7\) See Honts & Perry, supra note 71, at 359 ("Almost all of the scientists involved in detection of deception research reject the notion that the relevant-irrelevant test could be a useful discriminator of truth and deception."). The RIT is most often used for background or pre-employment screening but is sometimes used in criminal cases.

\(^8\) See John E. Reid, A Revised Questioning Technique in Lie Detection Tests, 37 J. CRIM. L., CRIMINOLOGY & POLICE SCI. 542 (1947).
task.  

There are several different types of CQT. "[T]he two most widely used formats are the zone comparison and the modified general question test (MGQT)." Some formats contain additional types of questions. For example, a "sacrifice relevant question" may be used as the first relevant question—e.g., "In regard to the theft, do you intend to answer the questions truthfully?" This question is not evaluated because it is assumed that the first introduction of the subject under investigation will trigger a response, even from an innocent person. A "symptomatic" question may also be used—e.g., "Are you convinced that I will not ask you a question that has not already been reviewed?" This question is intended to identify a subject who is concerned with some topic other than the one under investigation.

A recent refinement is the "directed lie" control question. In this procedure, the subject is told to lie to the directed control question—e.g., "Before age 25, did you ever tell a lie?" Subjects are instructed that if their reaction to this question is not sufficiently strong, the test will be inconclusive. In theory, guilty subjects will be unconcerned about this question and focus on the relevant questions. In contrast, innocent subjects will react because they are concerned about an inconclusive result. Nevertheless, this format is not recognized as an improvement by some federal agencies.
G. Methods of Evaluation

There are three methods of evaluation: global, numerical, and computerized. Global evaluation, the oldest method, involves an overall impression of the charts plus other factors. The most controversial of these other factors is the examiner’s “clinical impressions” of the subject during the pretest interview and the examination. In other words, the examiner considers the subject’s demeanor (“body language”) as well as the recorded reactions of the instrument. Critics contend that such a judgment is “a highly subjective and hence speculative interpretation about the meaning of a complex series of verbal, behavioral and physiological responses.”

The numerical approach was developed about 1960. The subject’s behavioral reactions are not considered, only the recorded chart reactions. There are several different scoring systems. The DoD Polygraph Institute and the University of Utah have developed similar systems. The comparative reaction to each pair of relevant and control questions is scored. The scores range from 3 for a dramatic reaction to a control question to -3 for the same type of reaction to a relevant question. Noticeable but smaller reactions are scored 1 or -1. A lack of a significant reaction is scored 0. Total scores of 6 or higher indicate truthfulness, while -6 or lower indicate deception. Scores that fall in between are considered inconclusive. The primary advantage of the numerical approach is that it “helps to ensure a rigorous, semi-objective evaluation of the physiological information contained in the charts.”

Computerized scoring is the third method. There are four algo-
rithms currently available: PolyScore™ (Applied Physics Laboratory. Johns Hopkins University), Computerized Polygraph System (University of Utah), Axciton, and Identifi. Computer scoring may be used as a quality control procedure or as the principal scorer; the latter method is only developmental at this time.

H. Quality Control Procedures

Typically, polygraph examinations conducted by federal agencies are independently reviewed by senior examiners. These agencies use numerical scoring, and the quality control reviews are “blind” examinations; the charts are evaluated without viewing the subject, or knowing the field examiner’s conclusion. As noted above, computers can also be used for quality control. An automated system reduces the risk of human error and minimize disagreements among examiners. Unfortunately, many examinations are conducted without any quality control safeguards.

III. Validity

The validity of polygraph testing in criminal investigations remains controversial. The question is extremely complex. Critics argue that the physiological responses caused by the fear of detection have not been shown to differ from physiological responses caused by other emotions. Proponents, however, do not claim that there is a special

96. See Olsen et al., supra note 84, at 61.
97. "Dichotomous computer classifications of subjects in the standardization sample were 93% correct. Blind numerical evaluations of the same data by an expert interpreter were 89% correct." Kircher & Raskin, supra note 81, at 291. This study used mock crime experiments and thus its application to field conditions cannot be assumed. See id. at 301. In a later field study, using U.S. Secret Service examiners, the computer evaluations proved reliable: “[T]he accuracy of human and computer interpretations was higher than the blind interpretations, and it ranged from 95-96% on confirmed truthful subjects and 83-96% on confirmed deceptive subjects.” DAVID C. RASKIN ET AL., A STUDY OF THE VALIDITY OF POLYGRAPH EXAMINATIONS IN CRIMINAL INVESTIGATION (May 1988) (Nat'l Inst. Of Justice, Grant No. 85-IJ-CX-0040).
98. “Disagreements among polygraph examiners are common and limit the validity and utility of the technique in applied settings.” Kircher & Raskin, supra note 81, at 292.
100. See Benjamin Kleinmuntz & Julian J. Szucko, On the Fallibility of Lie Detection, 17 LAW & SOC'y REV. 85, 87 (1982):

"[T]here is no reason to believe that lying produces distinctive physiological changes that characterize it and only it . . . [T]here is no set of responses—physiological or otherwise—that humans emit only when lying or that they produce only when telling the truth. . . . No doubt when we tell a lie many of us experience an inner tur-
physiological response that indicates deception. Rather, they believe that certain changes in physiological reactions in response to different types of questions indicate deception.

A. Purpose of Study

Some polygraph studies test for validity, while others test for reliability. The term "validity" refers to a test procedure’s accuracy, that is, its ability to measure what it is supposed to measure. The term “reliability” refers to the test’s consistency: whether the same results are obtained each time it is performed. Validity includes reliability, but the converse is not necessarily true. Understanding the literature on this subject also requires an appreciation of a number of additional issues. First, some tests such as the CQT appear to be “more accurate at detecting the deception of the guilty person than detecting the truthfulness of the innocent person.” Second, the subject of the test—whether suspect or victim—may also be important. Apparently, false positives are more frequent with victims than with suspects. Third, the type of issue involved is important: polygraph examinations involving specific factual issues produce more valid results than those involving mental state issues.

B. Field & Laboratory Studies

There are two different kinds of polygraph validity studies: (1) field studies of actual cases, and (2) mock crime experiments, which are moil, but we experience a similar turmoil when we are falsely accused of a crime, when we are anxious about having to defend ourselves against accusations, when we are questioned about sensitive topics—and, for that matter, when we are elated or otherwise emotionally stirred.

101. Another purpose (albeit non-scientific) involves an evaluation of the “utility” of the polygraph. Even if polygraph evidence is not admitted at trial, it may nevertheless be sufficiently reliable for other important uses, such as preemployment screening for police departments, national security purposes, and criminal investigations.


103. The latter would include a study designed to determine whether a single examiner reaches consistent results over a period of time, or whether several different examiners reach the same conclusion when the same subject is tested.


106. See Barland, supra note 104, at 83-84; Raskin, supra note 65, at 46-47.
laboratory simulations. Both have drawbacks. Field studies depend on establishing a valid criterion for determining guilt or innocence—sometimes called "ground truth." Some studies use panels of trial attorneys to determine guilt, an approach with obvious problems. Dr. Raskin argued that the best criterion is confirmation by a subject's later confession. Others, however, have pointed out that the confession criterion introduces a number of problems of sampling bias. The best available base line is probably confessions that have been independently corroborated to some degree.

Laboratory experimentation has different limitations. There are important differences between the laboratory and forensic environments that may undermine the validity of these experiments. The principal difference is that fear of detection is not as strong for experimental subjects. In addition, some of the laboratory studies fail to replicate field conditions; they use neither experienced examiners nor general population samples as subjects.

C. Selected Studies

The validity of polygraph testing is hotly disputed. In the fall of 1983, the Office of Technology Assessment of the U.S. Congress submitted a report in which it reviewed and evaluated the research on polygraph validity. The report includes the following passage as part of its findings:

OTA found meaningful scientific evidence of polygraph validity only

107. See Raskin, supra note 65, at 44 (stating that the best available method uses cases in which suspects confess after the polygraph examination, after which the charts are evaluated blindly by independent examiners).
108. See Christopher J. Patrick & William G. Iacono, Validity of the Control Question Polygraph Test: The Problem of Sampling Bias, 76 J. APPLIED PSYCHOL. 229, 229 (1991) ("A deceptive polygraph test outcome provides the incentive for an examiner to interrogate a subject, and if the subject confesses, the polygraph outcome is confirmed. (In the ultimate version of this self-fulfilling prophecy, an examiner may accept a minor admission from the subject—one that does not relate specifically to the issue of the polygraph test—as evidence of the subject's guilt.").
109. See David Lykken, The Lie Detector and the Law, 8 CRIM. DEF. 19, 23 (May-June 1981) ("Since the emotional impact of such artificial simulations, as well as the importance to the individual of the outcome, is inevitably very different than in real life situations, such laboratory assessments provide no valid basis for estimating the accuracy of the lie test in the field."). Researchers attempt to solve this problem by using substantial cash bonuses. See Raskin, supra note 105, at 99.
110. See Norman Ansley, Simulations and Polygraph Research, 24 POLYGRAPH 275. 282 (1995) ("If researchers want to know about the validity of polygraph examinations or find data on other aspects of field testing, simulations in the laboratory where they can control the factors under study need to be as much like field conditions as possible.").
in the area of criminal investigations. However, even here, there is a wide divergence in the results of the relevant research. Six prior research reviews showed average validity ranging from a low of 64 percent to a high of 98 percent. OTA's own review of 28 studies meeting minimum acceptable scientific criteria found that, for example, correct guilty detections ranged from 17 to 100 percent. Overall, the cumulative research evidence suggests that when used in criminal investigations, the polygraph test detects deception better than chance, but with significant error rates.\textsuperscript{111}

In contrast, a 1984 Department of Defense report reached far more favorable conclusions about polygraph validity:

Thus, the Bersh study, experience of investigators and quality control personnel, and mock crime laboratory studies give different estimates of the accuracy of control question tests in criminal investigations, ranging from about 80\% to 95\%. ... [T]here are no data suggesting that the various polygraph techniques and applications in [the Department of Defense] have high false positive or high false negative error rates.\textsuperscript{112}

A significant problem in interpreting the various reports concerns which studies used scientifically valid methodologies.\textsuperscript{113} In 1986, Dr. Raskin cited five mock crime studies that he considered valid because they used the control question technique, trained examiners, field techniques, and non-student populations. He concluded: "The combined accuracy of decisions was 95\%, with an inconclusive rate of 8\%. It should be noted that the majority of errors were false positive errors of diagnosing deception in subjects who were actually truthful. The evaluations misdiagnosed 8\% of innocent subjects as deceptive and only 3\% of guilty subjects as truthful (false negative errors)."\textsuperscript{114} In 1988, Dr. Raskin and his colleagues reported the results of a field study on the control question technique as administered by Secret Service personnel. In addition to field examiners, blind interpretation of charts by quality control examiners and computer interpretation were studied. The report concluded: "The accuracy of human and computer interpretations was very high. Decision by the original examiners on in-

\textsuperscript{111} OTA REPORT, supra note 66, reprinted in 12 POLYGRAPH 196, 200.


\textsuperscript{113} See John C. Kircher et al., Meta-Analysis of Mock Crime Studies of the Control Polygraph Technique, 12 LAW & HUM. BEHAV. 79 (1988) (differences in subjects, incentives, and decision policies may account for as much as 65\% of the observed variance in detection rates).

\textsuperscript{114} Raskin, supra note 65, at 42 ("These error rates indicate the limitations of the control question technique, even when it is performed under carefully controlled conditions by highly skilled examiners with extensive psychological training and expertise.").
dividual relevant questions ranged from 91-96% correct on confirmed truthful answers and 85-95% correct on confirmed deceptive answers. Blind interpretation produced somewhat lower accuracies, ranging from 63-85% on truthful answers and 84-94% on deceptive answers.”

Nevertheless, the controversy continued. Iacono and Patrick concluded in 1988 that the “best defense one can offer for the continued use of the CQT is that its accuracy is indeterminate.” In 1989, Raskin responded: “The voluminous scientific literature indicates that they can be highly accurate when properly employed in appropriate circumstances, but they are also subject to abuse and misinterpretation.” A decade later, the debate continues unabated.

D. Computer Scoring

A significant recent development involves computer software programs to evaluate polygraph charts. For example, PolyScoreTM, developed by John Hopkins University Applied Physics Laboratory, is a computerized algorithm for evaluating zone comparison examinations and is based on digitized polygraph data from criminal investigations. The program’s evaluations differ from those used by examiners.

E. Countermeasures

Another important research issue involves countermeasures or techniques to “beat” the test by a guilty subject. The OTA report commented that the “research on countermeasures has been limited and

115. RASKIN ET AL., supra note 97 (“However, the accuracy of the computer interpretations was higher than the blind interpretations, and it ranged from 95-96% on confirmed truthful suspects and 83-96% on confirmed deceptive subjects. The results provide considerable support for the accuracy of decisions made by the original examiners concerning the outcomes of polygraph tests.”).


118. See 1 DAVID L. FAIGMAN ET AL., supra note 64, ch. 14.

119. See Olsen et al., supra note 84.

120. The lab started collecting data in 1989. The 1997 article concluded: (1) “Because subject status, for many of the cases in the database, was determined by examiner decision, these results demonstrate that the algorithm is consistent with experienced examiner decisions.” (2) “The results indicate that an algorithm has been built that can separate the confirmed truthful and deceptive subjects in this database.” (3) “The algorithm does not mimic traditional chart interpretation.” Id. at 69-70.
the results conflicting." Countermeasures can be divided into two categories: (1) those that change the examinee's general physiological state such as drugs and biofeedback, and (2) those that produce effects at specific points in the examination such as mental imagery and physical countermeasures.

Physical countermeasures appear to pose the greatest threat to the control question technique. To be effective, the subject must produce stronger physiological responses to the control questions than to the relevant questions. Biting the tongue or pressing toes against the floor have produced significant false negative results in laboratory studies. One study concluded that the principal threat are subjects who "have received systematic training in countermeasures." Spontaneous countermeasures were not effective. Another study concluded that a polygraph instrument with an activity sensor would detect most attempts at physical countermeasures.

F. Friendly Polygrapher Issue

The friendly polygrapher hypothesis, first raised by Dr. Martin Orne as an issue worth investigating, suggests that a polygraph examination privately conducted by the defense may not be reliable because the fear of detection is not sufficiently realistic; the defendant knows that failing the test does not result in any adverse consequence. Thus, the possibility of a false negative will increase. This hypothesis, however, has never been established, and "[a]t present, the only research bearing upon this hypothesis does not support it."

IV. Evidentiary Standard

A number of objections to polygraph evidence can be labelled as
"Rule 403 issues." These include the danger that the jury will overvalue the expert's testimony,126 the danger that an opinion concerning the truthfulness of a witness will intrude too much into the jury's historic function of assessing credibility,129 and the possibility that the trial will degenerate into a time-consuming trial of the technique.130 None of these issues are insurmountable, and other commentators have recently written on them.131 Accordingly, this article focuses on the reliability issue"or rather, how courts have dealt with this issue.

A. Frye Standard

An examination of the polygraph cases illustrates the advantage of the Daubert approach over the Frye approach. Applying the Frye test raises several issues. According to Frye, psychology and physiology are the fields in which general acceptance must be achieved.132 Several decisions have expanded the "field" to include polygraph examiners.133 In United States v. Alexander,134 however, the Eighth Circuit rejected this view. A related issue concerns the extent to which the widespread use of the polygraph in law enforcement, intelligence, and industrial activities may be considered as evidence of general acceptance. Some courts have accorded such evidence considerable weight,135 while others have

128. See, e.g., United States v. Alexander, 526 F.2d 161, 168 (8th Cir. 1975) ("When polygraph evidence is offered . . . it is likely to be shrouded with an aura of near infallibility, akin to the ancient oracle of Delphi."); State v. Catanese, 368 So. 2d 975, 981 (La. 1979) ("trier of fact is apt to give almost conclusive weight to the polygraph expert's opinion").

129. See, e.g., People v. Baynes, 430 N.E.2d 1070, 1079 (Ill. 1981) ("A potential trial by polygraph is an unwarranted intrusion into the jury function."); State v. Davis, 407 So. 2d 702, 706 (La. 1981) ("usurps the jury's prerogative on a question involving credibility")

130. See, e.g., People v. Barbara, 255 N.W.2d 171, 196 (Mich. 1977) ("possibility of bogging down trials with collateral matters, perhaps resulting in a trial of the polygraph, or a battle of experts"); State v. Grier, 300 S.E.2d 351, 359-60 (N.C. 1983) ("possibility that the criminal proceeding may degenerate into a trial of the polygraph machine").


132. See Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923).


134. 526 F.2d 161, 164 n.6 (8th Cir. 1975) ("Experts in neurology, psychiatry and physiology may offer needed enlightenment upon the basic premises of polygraphy. Polygraphists often lack extensive training in these specialized sciences.").

ignored it. Finally, the ultimate *Frye* proof—a Gallup poll—has been cited as evidence of general acceptance and has spawned its own controversy.

These issues take courts further away from the critical issue of reliability. I have written elsewhere: "Perhaps the most important flaw in the *Frye* test is that by focusing attention on the general acceptance issue, the test obscures critical problems in the use of a particular technique." In contrast, *Daubert* requires courts to confront directly the validity issue.

**B. *Daubert* Standard**

Although *Daubert* asks the right question, the admissibility decision under the *Daubert* reliability test is not necessarily easier to answer. Some courts view *Daubert* as a less stringent standard than *Frye*, and the abandonment of the per-se exclusion rule in recent polygraph cases supports this view. It is not my view, however. *Daubert*, properly understood, adopts a different, but not a less stringent, standard of admissibility. Cases limiting the admissibility of previously well-established techniques such as handwriting and hair comparisons buttress this position.

---


137. See FAIGMAN ET AL., supra note 64.


139. See, e.g., Borawick v. Shay, 68 F.3d 597, 610 (2d Cir. 1995) ("[B]y loosening the strictures on scientific evidence set by *Frye*, *Daubert* reinforces the idea that there should be a presumption of admissibility."); *cert. denied*, 116 S. Ct. 1869 (1996); United States v. Bonds, 12 F.3d 540, 568 (6th Cir. 1993) ("We find that the DNA testimony easily meets the more liberal test set out by the Supreme Court in *Daubert*.").

140. See United States v. Kwong, 69 F.3d 663, 668-69 (2d Cir. 1995) ("The Federal Rules of Evidence, although concededly more liberal than the *Frye* test, still require a determination that the proffered scientific evidence is both relevant and reliable.").

141. See, e.g., United States v. Starzecpyzel, 880 F. Supp. 1027, 1038 (S.D.N.Y. 1995) (concluding that the "testimony at the *Daubert* hearing firmly established that forensic document examination, despite the existence of a certification program, professional journals and other trappings of science, cannot, after *Daubert*, be regarded as 'scientific... knowledge'").

142. See, e.g., Williamson v. Reynolds, 904 F. Supp. 1529, 1558 (E.D. Okl. 1995) (excluding hair comparison while observing that the court had "been unsuccessful in its attempts to locate any indication that expert hair comparison testimony meets any of the requirements of *Daubert*.") Further, "[a]lthough the hair expert may have followed procedures
Prior to *Daubert*, courts had identified several reliability factors:\(^{143}\) (1) the lack of empirical validation,\(^{144}\) (2) the numerous uncontrollable factors involved in the examination,\(^{145}\) (3) the subjective nature of the deception determination,\(^{146}\) and (4) the absence of adequate standards for assessing the qualifications of examiners.\(^{147}\) The *Daubert* Court cited a nonexhaustive list of factors, which trial courts may use to determine reliability: (1) whether the scientific technique has been tested, (2) whether the technique has been subjected to peer review and publication, (3) the technique's rate of error, (4) the existence and maintenance of standards controlling the technique's operation, and (5) the technique's "general acceptance." These factors are often difficult to apply, as illustrated by the Court's questionable use of these factors in *Daubert* itself.\(^{148}\) An argument can be made for the validity of all these

---

\(^{143}\) See, e.g., United States v. Alexander, 526 F.2d 161, 166 (8th Cir. 1975) ("We are still unable to conclude that there is sufficient scientific acceptability and reliability to warrant the admission of the results of such tests in evidence."); People v. Baynes, 430 N.E.2d 1070, 1076 (Ill. 1981) ("[T]he primary obstacle in admission of polygraph evidence, stipulated to or not, has continually and consistently been the instrument's disputed scientific reliability.").

\(^{144}\) See, e.g., United States v. Wilson, 361 F. Supp. 510, 514 (D. Md. 1973) ("incipient stage of experimental research"); People v. Barbara, 255 N.W.2d 171, 189 (Mich. 1977) ("[W]e find there is difficulty in empirically verifying polygraphic findings.... some questions concerning the accuracy figures.... and an apparent lack of a rational scientific explanation for this phenomenon.").

\(^{145}\) See, e.g., People v. Anderson, 637 P.2d 354, 359 (Colo. 1981) ("Several uncontrollable or unascertainable physiological and psychological responses may cause difficulty or error...."); Baynes, 430 N.E.2d at 1075 ("abnormal blood pressure, heart and respiratory irregularities, fatigue, intoxication, rationalization, extreme nervousness, meditative abstraction, controlled breathing, instrument discomfort, uncomfortable room temperature, obesity, hidden muscular contractions, self-inflicted pain, mental incompetence, psychopathy or schizophrenia, ingestion of a sedative, an attempt to suppress a cough, fear induced by suspicion or accusation, or lack of fear of detection").

\(^{146}\) See, e.g., People v. Anderson, 637 P.2d 354, 360 (Colo. 1981) (The polygraph technique, "albeit based on a scientific theory, remains an art with unusual responsibility placed on the examiner."); People v. Monigan, 390 N.E.2d 562, 569 (Ill. App. 1979) ("almost total subjectiveness surrounding the use of the polygraph and the interpretation of the results").

\(^{147}\) See *Anderson*, 637 P.2d at 360 ("The absence of adequate qualification standards for the polygraph profession heighten[s] the possibility for grave abuse...."); State v. Catanese, 368 So. 2d 975, 982 (La. 1979) (lack of judicial and legislative control over competence of examiners).

factors, but they are often different ways of looking at the same problem "the underlying validity of polygraph evidence."

V. Polygraph Cases

Several aspects of the favorable polygraph cases are striking. The first is the failure of the courts to appreciate the lack of basic scientific research in this area. The second is the courts' failure to understand the differences in polygraph techniques—assuming incorrectly that all polygraph tests are alike. The third is the failure to specify the requisite qualifications for examiners.

A. Lack of Research

The polygraph has been around since Frye (1923), and courts often cite the tremendous advances that have been made over the last 70 years. Nevertheless, good scientific research on the polygraph is a recent development, and these courts have not rigorously scrutinized this issue.

In Ridling, the landmark 1972 case, the court wrote that the "scientific psychological basis for the polygraph examination is well established." The court did not, however, cite any validity studies to support this statement. The court did cite a polygraph text by Reid and Inbau, the leading polygraph experts of their time. Reid and Inbau wrote: "This estimate accords to the lie-detector technique, when applied under the most favorable conditions, an accuracy of 95 percent, with a 4 percent margin of indefinite determinations and a 1 percent margin of possible error." This error rate is suspect because it is based on the assumption that polygraph results are correct unless proven otherwise. No systematic follow-up studies were conducted to verify the examiner's conclusions, verification criteria are not specified, and improper procedures may have been used to compute the error rate.

149. See e.g., United States v. Posado, 57 F.3d 428, 434 (5th Cir. 1995) ("There can be no doubt that tremendous advances have been made in polygraph instrumentation and technique in the years since Frye."); United States v. Piccionnna, 885 F.2d 1529, 1530 (11th Cir. 1989) (commenting upon "the significant progress made in the field of polygraph testing over the past forty years"); United States v. Crumby, 895 F. Supp. 1354, 1358 (D. Ariz. 1995) ("[T]here has been a significant increase in the reliability of polygraph evidence over recent years.").


152. See Orne, supra note 126, at 103-04; Validity Panel, in LEGAL ADMISSIBILITY OF
As noted earlier, in 1989 two noteworthy polygraph cases were decided. In *Piccinonna*, the Eleventh Circuit concluded that polygraph evidence was admissible even in the absence of a stipulation. The court based its decision in part on “new empirical evidence and scholarly opinion which have undercut many of the traditional arguments against admission of polygraph evidence.” According to the court, there “is no question that in recent years polygraph testing has gained increasingly widespread acceptance.” The second case, *Mendes*, was decided by the Supreme Judicial Court of Massachusetts. In 1974 that court had decided a landmark case admitting polygraph evidence without stipulation. In *Mendes* the court abruptly changed direction and excluded polygraph evidence, basing its opinion in part on “the failure of the basic theory of polygraphy to have gained general acceptance among physiological and psychological authorities.” Thus, within the span of three months two courts reached diametrically opposed views on the validity of polygraph evidence, both basing their respective opinions on “recent scientific research.” Actually, neither court cited the most recent research. Indeed, there are no citations to recent research in the majority opinion in *Piccinonna*; only in the dissent is polygraph research cited. In *Posado*, a 1995 decision, the Fifth Circuit cited as “current research” no scientific article published after 1988.

The lack of quality research until quite recently has often been commented on. In 1988, Dr. Gordon Barland wrote: “Only now are superior paradigms being developed which combine the ground truth of the laboratory with the realism of field applications.” In 1992, the FBI’s top polygraph expert commented that “the greatest threat to polygraph is the shakiness of the scientific theory upon which it rests.” He added:

---

153. 885 F.2d at 1533.
154. *Id.* at 1535. On remand, the trial court excluded the evidence because the questions and answers were irrelevant, and any probative value was outweighed by the danger of misleading the jury. United States v. Piccinonna, 729 F. Supp. 1336 (S.D. Fla. 1990).
156. See supra note 41 and accompanying text (discussing *A Juvenile*).
158. United States v. Posado, 57 F.3d 428, 434 (5th Cir. 1995).
159. A 1984 DoD study noted that there “has been more scientific research conducted on lie detection in the last six years than in the previous 60 years.” *Department of Defense, The Accuracy and Utility of Polygraph Testing* (1984), *reprinted in 13 Polygraph* 1, 58 (1984).
160. Barland, supra note 104, at 76.
We define control questions a little differently now, minor instrumentation changes have occurred, and we score charts more objectively. This is little change. Drew Richardson expressed it this way. He observed that Robert Goddard would likely be amazed to see modern space flight. And Alexander Fleming would be affected similarly with modern medicine. But, were Dr. William Marston (of 1923 Frye case fame), their most famous contemporary in lie detection to return from the grave today, he would hardly feel he had skipped a beat. In 1995, the Director of the DoD Polygraph Institute noted: "The period between 1986 and the present has been one of unparalleled advances in the psychophysiological detection of deception testing procedures and processes.

In sum, the improvements in polygraph testing over the last 70 are significant but do not, by themselves, answer the Daubert question—whether the validity of the improvements has been scientifically established.

B. Specificity of Procedure

Courts admitting polygraph evidence often refer to "properly conducted" examinations without specifying what is a proper examination. For example, in Ridling, the court did not specify what polygraph technique it was sanctioning; the CQT is not mentioned. Similarly, in Posado, the Fifth Circuit was remarkably vague concerning the specifics of polygraph examinations, mentioning only "controlled conditions."

The two leading federal district court cases are United States v. Crumby and United States v. Galbreth. Dr. Raskin was the expert in both cases. In Crumby, the court mentioned the control question test but did not specify the method of evaluation (global or numerical). Galbreth provides more information about the testing procedures used. It identifies the "directed lie control question" (which Crumby

162. Id. at 165.
163. Yankee, supra note 95, at 63.
164. See e.g., United States v. Galbreth, 908 F. Supp. 877, 881 (D.N.M. 1995) ("[T]he validity of polygraph results in a particular case is absolutely dependent on certain conditions such as a properly conducted examination by a competent examiner.").
165. United States v. Posado, 57 F.3d 428, 434 (5th Cir. 1995).
168. The court also cited N.M. R. Evid. 707 for guidance. The Rule does require taping and disclosure of prior testing but little useful beyond this. It requires two relevant questions, quantitative scoring, and at least three charts. The federal statute requires a pretest interview, written questions, and a minimum of a 90 minute test. See supra note 63 and accompa-
did not) but does not provide citations to the research it discusses. Moreover, *Crumby* makes no reference to quality control, a significant issue in any scientific procedure. *Galbreth* mentions quality control but does not seem to realize its importance, or that it is rare except in federal examinations.

The *Galbreth* opinion also mentions computer programs. The court wrote:

The specific computer scoring method used in this case was developed by Dr. Raskin and has been subjected to scientific investigation. It works by an analytic solution using a standard computer program. The examiner simply runs the program and the computer makes tens of thousands of calculations within five to ten seconds. This scoring method is completely objective.

There are several problems with this passage. First, the method is not “completely objective.” As one published paper on computer scoring acknowledged: “We emphasize that PolyScore can only effectively evaluate properly collected data. If a subject is not a suitable candidate for a polygraph examination, if the issue is not clearly defined or of sufficient intensity, or if proper zone comparison procedures are not used, accuracy will suffer.” In short, the computer input depends on the examiner’s ability, a subjective aspect of the technique.

Second, if the computer program arrived at its conclusions independent of the examiner’s method of scoring, as some programs do, the court has missed a critical point. This method of evaluation is significantly different from examiner evaluation and should, therefore, be independently scrutinized under *Daubert*. In other words, research on the CQT as scored by examiners (even using a numerical scoring system) does not validate an evaluation using a computer algorithm. Moreover, the validity of one algorithm does not validate other algorithms.

As they did with DNA evidence, courts should specify the particu-


170. See Ronald M. Furgerson, *Polygraph Policy Model for Law Enforcement*, 56 F.B.I. LAW ENFORCEMENT BULL. 7, 19 (June 1987) (“Experience has shown the value of quality control as an integral part of law enforcement polygraph usage. In such a program, polygraph charts and documentation are reviewed ‘in the blind’ by another senior and well-qualified examiner to insure that they substantiate the conclusion of the testing examiner as to truth or deception.”); Barland, *supra* note 89, at 50 (“By far the most important single safeguard that should be required prior to admissibility of polygraph evidence is the review of the polygraph examination by an objective, disinterested expert polygraph examiner.”).


172. Olsen et al., *supra* note 84, at 70.

173. See *id.* at 69; *supra* note 120 and accompanying text.
lar procedure they are addressing and then ask whether this particular procedure has been scientifically validated.

C. Examiner Qualifications

Courts admitting polygraph evidence also talk about “qualified examiners,” again without specifying what qualifications are necessary. This is rather strange when even polygraph proponents state that examiners’ qualifications are often deplorable. Dr. Raskin was the expert in both *Crumby* and *Galbreth*. Very few examiners have his qualifications, which are primarily as a researcher.

As noted above, the DoDPI trains all federal polygraph examiners. It “is the only program known to base its curriculum on forensic psychophysiology, and conceptual, abstract, and applied knowledge that meet the requirements of a master’s degree-level of study.” The qualifications standards include a college degree (or equivalent), two years’ experience as an investigator, and a six-month internship that includes conducting at least 25 polygraph exams. In addition, 80 hours of continuing instruction are required annually. Outside of federally trained examiners, few have such qualifications.

Conclusion

The polygraph should not be treated as “junk science.” The current research on computer algorithms is an important, promising development. Moreover, the polygraph is now used extensively in criminal practice, albeit on the periphery of the trial. It often assists in the extrajudicial resolution of difficult credibility cases—for example, child sexual abuse cases. Paradoxically, these cases are often both too hard and too easy to prove. It is not surprising that the DoD annual reports include a disproportionate number of such cases, leading to confessions as well as to exculpation. The polygraph may also assist in motions to suppress (the issue in *Posado*).

Nevertheless, courts must scrutinize the admissibility issue with far more care, specifying required procedures and requisite qualifications.

174. See, e.g., United States v. Ridling, 350 F. Supp. 90, 93 (E.D. Mich. 1972) (“For a test to be successful, it is important . . . that a person skilled in the interpretation of the polygraph charts make the interpretation of the test results.”).
175. See notes 70-71 and accompanying text.
177. See N.M. R. EVID. 707 (Minimum qualifications include five years’ experience administering or interpreting examinations or equivalent academic training and at least twenty hours of continuing education during the twelve months prior to the examination offered in evidence).
Moreover, the use of court-appointed examiners to reduce subjective bias is necessary.