1994

Polygraph Evidence: Part I

Paul C. Giannelli

Forensic Science: Polygraph Evidence: Part I

By Paul C. Giannelli*

In 1989, two noteworthy polygraph cases were decided. In *United States v. Piccinonna*[1] 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."[2] According to the court, there "is no question that in recent years polygraph testing has gained increasingly widespread acceptance."[3]

The second case, *Commonwealth v. Mendes*,[4] was decided by the Supreme Judicial Court of Massachusetts. In 1974, that court had decided a landmark case admitting polygraph evidence without stipulation.[5] In *Mendes* the court abruptly changed directions 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."[6]

Thus, within the span of three months two courts reached diametrically opposed views on the reliability of polygraph evidence, both basing their respective opinions on "recent scientific research." Sixty years after polygraph evidence was first excluded in *Frye v. United States*,[7] the controversy continues. Indeed, *Frye* itself was overturned by the U.S. Supreme Court in 1993.[8]

This is the first of a two-part article on polygraph evidence in criminal trials.[9] This part focuses on the scientific issues and procedures. The second examines the legal issues, such as admissibility.

---

* Albert J. Weatherhead III & Richard W. Weatherhead Professor of Law, Case Western Reserve University. This column is based in part on P. Giannelli & E. Imwinkelried, *Scientific Evidence* (2d ed. 1993). Reprinted by permission. Part II of this column will appear in an upcoming issue of the *Criminal Law Bulletin*.

[1] 885 F.2d 1529 (11th Cir. 1989).

[2] *Id.* at 1533.

[3] *Id.* at 1535.

[4] 406 Mass. 201, 547 N.E.2d 35 (1989).

[5] Commonwealth v. A Juvenile, 365 Mass. 421, 313 N.E.2d 120 (1974).

[6] 406 Mass. at 201, 547 N.E.2d at 35-36.

[7] 293 F. 1013 (D.C. 1923).

[8] Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1986).

[9] Polygraph testing, however, raises additional legal issues, such as preemployment screening. These issues are governed by the Employee Polygraph Protection Act of 1988. 29 U.S.C. §§ 2001-2009 (1991). With limited exceptions, the act prohibits the use of polygraph tests for preemployment screening or during the course of employment. See Note, "The Employee Polygraph Protection Act of 1988—Should the Federal Government Regulate the Use of Polygraphs in the Private Sector?," 58 U. Cin. L. Rev. 559 (1989); Note, "The Employee Polygraph Protection Act of 1988: A Balance of Interests," 75 Iowa L. Rev. 539 (1990).

Underlying Theory

Modern polygraph procedures developed over a long period of time, commencing around the turn of the century.[10] The most common polygraph examination (the control question technique) is based upon two premises: (1) The psychological stress caused by the fear of detection produces involuntary physiological responses and (2) a polygraph examiner, based on these responses as recorded by a polygraph machine, can detect deception. A report by the Office of Technology Assessment (OTA) explained:

> The basic theory of polygraph testing is only partially developed. The testing process is complex and not amenable to easy understanding. The most commonly accepted theory at present is that, when the person being examined fears detection, that fear produces a measurable physiological reaction when the person responds deceptively. Thus, in this theory, the polygraph instrument is measuring the fear of detection rather than deception per se. And the examiner infers deception when the physiological response to questions about the crime or unauthorized activity is greater than the response to other questions.[11]

The "fear of detection" theory, however, is but one of a number of theories that have been proposed to explain a subject's reactions.[12]

Critics of the Theory

Critics argue that the physiological responses caused by the fear of detection have not been shown to be different from physiological responses caused by other emotions:

> [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 turmoil, but we experience a similar turmoil when we are falsely accused of a crime, when we are anxious about having to defend outselves against accusations, when we are questioned about sensitive topics—and, for that matter, when we are elated or otherwise emotionally stirred.[13]

Proponents, however, do not claim that there is a special physiological response that indicates deception. Rather, they believe that *changes* in physiological reactions in response to

[10] For a history of the development of the polygraph and the extent of its present day use, see Barland, "The Polygraph Test in the USA and Elsewhere," in *The Polygraph Test: Lies, Truth and Science* 73 (A. Gale ed. 1988).

[11] U.S. Congress, Office of Technology Assessment, Scientific Validity of Polygraph Testing: A Review and Evaluation—A Technical Memorandum, OTA-TM-H-15 (1983) (hereinafter OTA Report), reprinted in 12 Polygraph 196, 201 (1983).

[12] See Davis, "Physiological Responses as a Means of Evaluating Information," in *The Manipulation of Human Behavior* 142, 160–165 (A. Biderman & Zimmer eds. 1961) (discussing conditioning theories (the conditioned response theory, the conflict theory), and the threat-of-punishment theory).

[13] Kleinmuntz & Szucko, "On the Fallibility of Lie Detection," 17 Law & Soc'y Rev. 85, 87 (1982). See also Lykken, "The Lie Detector and the Law," 8 Crim. Def. 19, 21 (May-June 1981) ("But people do not all react in the same way when they are lying and, more important, any reaction that you might display when answering deceptively you might also display another time, when you are being truthful").

different types of questions indicates deception.

Instrument

The physiological responses used in polygraph testing are changes in blood pressure-pulse, respiration, and galvanic skin resistance.[14] The polygraph machine simultaneously and continuously measures and records these physiological reactions on a graph or chart (i.e., polygram). Blood pressure-pulse is measured by a sphygmomanometer (i.e., blood pressure cuff that is placed on the subject's arm; respiration is measured by pneumograph tubes that are fastened around the subject's abdomen and chest; and galvanic skin response is measured by electrodes that are attached to the subject's fingertips.[15]

There seems little question that a quality polygraph machine can accurately measure and record these responses.[16] The machine, 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.

Role of the Examiner

The examiner's role is critical because it is the examiner who decides whether there is sufficient indication of deception. The OTA report states:

> [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.[17]

Even the proponents of the polygraph technique agree that the examiner, and not the machine, is the crucial factor in arriving at reliable results.[18] 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. One authority has commented that "a substantial proportion of those who conduct tests in the public and private sectors lack adequate training and compe-

---

[14] Some machines are also equipped to record muscular activity. These tracings may reveal efforts to "beat" the machine and in some cases provide independent deception criteria. J. Reid & F. Inbau, *Truth and Deception* 262 (2d ed. 1977).

[15] The galvanic skin resistance or electrodermal response involves the measurement of changes in the flow of electrical current. *Id.* at 275–291.

[16] See State v. Dean, 103 Wis. 2d 228, 235, 307 N.W.2d 628, 632 (1981) ("A quality machine accurately measures and records these body responses").

[17] OTA Report, *supra* note 11, reprinted in 12 Polygraph at 196 (statement of John Gibbons, Director of Office of Technology Assessment).

[18] J. Reid & F. Inbau, *supra* note 14, at 5 ("[T]he most important factor involved in the use of any such instrument is the ability, experience, education, and integrity of the examiner himself").

tence."[19] Others have observed that "[p]olygraph examiners in the United States, as a whole, are poorly trained."[20]

The American Polygraph Association accredits polygraph schools. The private schools conduct a seven-week course, while the Department of Defense (DOD) Polygraph Institute runs a 14-week course. Many states also have licensing statutes. In addition, one authority recommends three years of full-time experience, including 250 forensic cases (not employment screening).[21] A number of state licensing statutes as well as the Department of Defense require continuing education (typically 20 hours per year). The DOD Polygraph Institute, the American Polygraph Association, the American Association of Police Polygraphists, and various regional and state polygraph associations conduct specialized and annual courses.

Procedure

The polygraph technique involves several steps, the most important of which are the pretest interview and the examination of the subject while attached to the machine. These steps are preceded by a preliminary investigation[22] and followed by a posttest interview.[23]

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.[24] 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,[25] low intelligence,[26] or the use of medication,[27] that may affect the test results. Third, test questions are formulated with the subject's assistance during the interview.

Types of Examinations

There are several different types of polygraph examinations. Originally, examiners used the "relevant-irrelevant question" (RI) test. Relevant questions are incriminating questions, and the response to such questions are compared with the subject's response to irrelevant or neutral questions. The main criticism of the RI test is its underlying assumption that an innocent person will not react to the relevant questions: "Because relevant questions may produce relatively strong reactions, even when answered truthfully, many innocent subjects would be expected to produce deceptive outcomes on the

[19] Raskin, "The Polygraph in 1986: Scientific, Professional and Legal Issues Surrounding Application and Acceptance of Polygraph Evidence," 1986 Utah L. Rev. 29, 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.

[20] Honts & Perry, "Polygraph Admissibility," 16 Law & Hum. Behav. 357, 375 (1992).

[21] Barland, "Standards for the Admissibility of Polygraph Results as Evidence," 16 U.West L.A. L. Rev. 37, 44 (1984).

[22] The preliminary investigation is designed to provide the examiner with as much information as possible about the incident under investigation and the subject of the examination. J. Reid & F. Inbau, *supra* note 14, at 11.

[23] Unlike the pretest interview and the examination itself, the examiner need not remain objective in the posttest interview. Indeed, its principal function is usually to elicit a confession from those subjects considered deceptive. *Id.* at 4.

[24] *Id.* at 13–14.

[25] *Id.* at 233.

[26] *Id.* at 247.

[27] *Id.* at 236.

test."[28] Other authorities disagree, arguing that RI test, if properly used, may be useful.[29] Despite the criticism, the RI test is still used today.

Later, the control question technique was developed as an improvement of the relevant-irrelevant exam.[30] It is the most common type of examintion in criminal cases and is discussed below.

A third examination, the "Guilty Knowledge" or "Concealed Knowledge" test, is 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 RI and control question technique (CQT) tests, which are deception tests. The requirement of concealed information greatly limits its use.[31]

Control Question Technique

The most important type of examination is the CQT.[32] 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 that the subject understands them.

Several different types of questions are used in the CQT. Irrelevant or neutral questions are used to obtain a subject's normal truthful reactions and chart tracings. Examples of irrelevant questions are: "Is your name [subject's name]?" "Are you over 21 years of age?" Relevant questions concern the subject matter under investigation. For example: "Did you take $100 from your employer's safe?" The third type of question is the control question. 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."[33] For this reason, they are sometimes called "probable lie" questions. An example would be: "Did you ever steal anything in your life?" 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

---

[28] Raskin, *supra* note 19, 1986 Utah L. Rev. at 33. See also Honts & Perry, "Polygraph Admissibility," 16 Law & Hum. Behav. 357, 359 (1992) ("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").

[29] See Barland, "The Polygraph Test in the USA and Elsewhere," in *The Polygraph Test: Lies, Truth and Science* 73, 80 (A. Gale ed. 1988).

[30] See Reid, "A Revised Questioning Technique in Lie Detection Tests," 37 J. Crim. L., Criminology & Police Sci. 542 (1947).

[31] See Raskin, *supra* note 19, 1986 Utah L. Rev. at 31–32 ("[I]t is employed infrequently because the special information necessary to construct a valid concealed information test is typically lacking").

[32] Further refinements of the control question technique have been reported. See Honts & Raskin, "A Field Study of the Validity of the Directed Lie Control Question," 16 J. Police Sci. & Admin. 56 (1988) (discussing the directed lie control question).

[33] J. Reid & F. Inbau, *supra* note 14, at 28. See also Raskin, "Science, Competence, and Polygraph Techniques," 8 Crim. Def. 11, 13 (May–June 1981) ("[T]he control question deals with similar subject matter, is very general in nature, covers a long span of time and a large number of possible acts, and it is almost impossible for most people to answer it with an unequivocal 'no' and with certainty that they are being completely truthful").

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.[34]

The examination typically consists of ten to twelve questions. The first one or two questions are irrelevant questions. Other irrelevant questions as well as the relevant and control questions are interspersed in the remaining questions. While the subject knows the questions, he does not know the order in which they will be asked. The examination lasts a few minutes and is repeated at least one more time; often two or three more examinations are conducted.

The following example has been used to illustrate the control question sequence:

1. (Neutral) Do you understand that I will ask only the questions we have discussed?
2. (Pseudo-Relevant) Regarding whether you took that ring, do you intend to answer all of the questions truthfully?
3. (Neutral) Do you live in the United States?
4. (Control) During the first twenty-four years of your life, did you ever take something that did not belong to you?
5. (Relevant) Did you take a ring from the Behavioral Sciences Building on July 1, 1985?
6. (Neutral) Is your name Joanne?
7. (Control) Between the ages of ten and twenty-four, did you ever do anything dishonest or illegal?
8. (Relevant) Did you take that diamond ring from a desk in the Behavioral Sciences Building on July 1?
9. (Neutral) Were you born in the month of February?
10. (Control) Before 1984 did you ever lie to get out of trouble or to cause a problem for someone else?
11. (Relevant) Were you in any way involved in the theft of that diamond ring from the Behavioral Sciences Building last July?[35]

As part of the examination, a stimulation test is often administered.[36] It is sometimes known as the number test, card test, or stim test. A wide variety of stimulation tests are used. The purpose of the test is to impress the subject with the efficacy of the technique. Reid and Inbau describe a card test that is based on deceiving the subject.[37] However, most stimulation tests, such as those used by federal examiners, do not involve trickery.[38]

---

[34] "[I]nnocent subjects are expected to show stronger reactions to the control questions than to the relevant questions, whereas guilty subjects are expected to show the opposite." Kircher & Raskin, "Human Versus Computerized Evaluations of Polygraph Data in a Laboratory Setting," 73 J. Applied Psychol. 291 (1988).

[35] Raskin, *supra* note 19, 1986 Utah L. Rev. at 36.

[36] An issue of *Polygraph* is devoted to the subject. 7 Polygraph 173–214 (1978).

[37] In the card test, the subject is asked to select a card from a deck. The examiner then goes through all the cards, one at a time, asking if each was the one selected. The subject is instructed to answer "no" each time, even when the correct card is shown. The examiner, supposedly based on the polygraph technique, then identifies the correct card. Often the identification is not made through the polygraph technique but because the cards are marked. J. Reid & F. Inbau, *supra* note 14, at 42 & 85.

[38] Decker, "The Army Stimulation Test–A Control Procedure," 7 Polygraph 176, 176 (1978) ("There is absolutely no trickery in this test").

Formulating adequate control questions is not an easy task. As one writer has noted, "it is extremely difficult to devise control questions that would ensure the eliciting of stronger reactions in an innocent person than would the relevant questions relating to the crime of which they had been accused."[39] This may explain why there are more false positives than false negatives when the CQT is used.[40]

Methods of Evaluation

There are three methods of evaluation: global evaluation, numerical scoring, and computerized scoring. 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 as well as the recorded reactions of the machine.[41] 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."[42]

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 systems developed by the DOD Polygraph Institute and by the University of Utah are similar.[43] 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 the relevant question. Noticeable but small reactions are scored +1 or −1. No 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 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."[44] Moreover, some research indicates that numerical scoring systems are more reliable.[45] However, since the subject's behavior is not considered, a higher number of inconclusive conclusions are reached in this approach. Computerized scoring is a specific application of numerical scoring.

Quality Control Procedures

Typically, polygraph examinations conducted by federal agencies are independently reviewed by other examiners. The quality control reviews are "blind" examinations; the charts are evaluated without viewing the subject, or knowing the field examiner's conclusion.

---

[39] Bull, "What is the Lie-Detection Test?," in *The Polygraph Test: Lies, Truth and Science* 14 (A. Gale ed. 1988).

[40] *Id.*

[41] "For example, [the examiner] must look at the polygraph charts, the suspect's demeanor and behavior, the case facts, and whatever other indicators there may be." Barland, *supra* note 21, 16 U.West.L.A. L. Rev. at 39.

[42] Kleinmuntz, "The Polygraph as Credible Court Evidence," The Champion 14, 16 (Sept.-Oct. 1984).

[43] Raskin, *supra* note 19, 1986 Utah L. Rev. at 37–38.

[44] Barland, *supra* note 21, 16 U.West L.A. L. Rev. at 40.

[45] See Barland & Podlesny, "Validity and Reliability of Detection of Deception," 6 Polygraph 1, 18 (1977); Weaver, "The Numerical Evaluation of Polygraph Charts: Evaluation and Comparison of Three Major Systems," 9 Polygraph 94 (1980).

> 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.[46]

Computers can be used for quality control.[47] An automated system reduces the risk of human error and minimizes disagreements among examiners.[48]

In the absence of a quality control program, review by an independent examiner is critical: "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."[49] Unfortunately, many examinations are conducted without any quality control safeguards.

One authority's account of the *Fay* case illustrates the importance of quality control procedures and examiner qualifications:

> In the celebrated case of Floyd Fay, who was wrongly convicted of murder in 1978 and served two years in prison before the actual perpetrators were apprehended, the five interpreters used different methods to analyze the polygraph tests. The two field polygraph examiners who testified at the trial diagnosed him as deceptive, one using the United States Army numerical scoring system and the other using a global evaluation. Another examiner, a professor of criminology at Michigan State University who employed a global evaluation, said that the test was inconclusive. A private polygraph examiner with a doctoral degree in psychology scored the polygraph charts at +1 (inconclusive) using the United States Army system, and a psychiatry professor at the University of Minnesota said that polygraph tests are of no value. The author [Dr. Raskin] interpreted the charts as truthful (+7).[50]

Validity

The validity of polygraph testing in criminal investigations remains controversial. The question is ex-

---

[46] Furgerson, "Polygraph Policy Model for Law Enforcement," 56 F.B.I. Law Enforcement Bull. 7, 14–19 (June 1987).

[47] "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, "Human Versus Computerized Evaluations of Polygraph Data in a Laboratory Setting," 73 J. Applied Psychol. 291 (1988). This study used mock crime experiments and thus its application to field conditions cannot be assumed. *Id.* at 301. In a later field study using U.S. Secret Service examiners the computer evalutions 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." D. Raskin, J. Kircher, C. Honts & S. Horowitz, A Study of the Validity of Polygraph Examinations in Criminal Investigation (May 1988) (Nat'l Inst. of Justice, Grant No. 85-IJ-CX-0040).

[48] "[D]isagreements among polygraph examiners are common and limit the validity and utility of the techniques in applied settings." Kircher & Raskin, *id.* at 292.

[49] Barland, *supra* note 21, 16 U.West L.A. L. Rev. at 50.

[50] Raskin, *supra* note 19, 1986 Utah L. Rev. at 39–40.

tremely complex[51] but two preliminary points are not subject to dispute. First, error rates frequently cited by field examiners are suspect because they are often based on the assumption that polygraph results are correct unless proven otherwise. In many instances no systematic follow-up studies have been conducted to verify the examiner's conclusions, verification criteria are not specified, and improper procedures are used to compute the error rate.[52]

Second, polygraph research is an on-going process. A 1984 Department of Defense study noted that there "has been more scientific research conducted on lie detection in the last six years than in the previous 60 years."[53] In 1988, another authority wrote: "Only now are superior paradigms being developed which combine the ground truth of the laboratory with the realism of field applications."[54]

Types of Studies

There are two different kinds of polygraph studies: (1) field studies of actual cases and (2) mock crime experiments, which are laboratory simulations. Both have drawbacks. Field studies depend on establishing a valid criterion for determining guilt or innocence. Some studies use panels of trial attorneys to determine guilt, an approach with obvious problems. One expert argues that the best criterion is confirmation by a subject's later confession.[55] Others, however, have pointed out that "the use of a confession criterion introduces a number of problems of sampling bias that in turn raise questions about the usefulness of confession studies."[56]

Laboratory experimentation has different limitations. There are important differences between the laboratory and forensic environments that may undermine the validity of these experiments.[57] The principal difference is that fear of detection is not as strong for experimental subjects.[58] In

---

[51] Orne, Thackray & Paskewitz, "On the Detection of Deception," in *Handbook of Psychophysiology* 743, 751 (N. Greenfield & R. Sternback eds. 1972) ("No fully satisfactory way is available at this time for evaluating the overall effectiveness of the technique, and it is probable that no such answer will be forthcoming in the near future from real life situations").

[52] "Validity Panel," in *Legal Admissibility of the Polygraph* 155 (N. Ansley ed. 1974) (statement of Gordon Barland) (hereinafter Validity Panel).

[53] Department of Defense, "The Accuracy and Utility of Polygraph Testing" (1984), reprinted in 13 Polygraph 1, 58 (1984).

[54] G. Barland, "The Polygraph Test in the USA and Elsewhere," *The Polygraph Test: Lies, Truth and Science* 76 (A. Gale ed. 1988).

[55] Raskin, *supra* note 19, 1986 Utah L. Rev. 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).

[56] Honts & Perry, "Polygraph Admissibility," 16 Law & Hum. Behav. 357, 361 (1992) (citing Patrick & Iacono, "Validity of the Control Question Polygraph Test: The Problem of Sampling Bias," 76 J. Applied Psychology 229 (1991)).

[57] Validity Panel, *supra* note 52, at 160–162.

[58] Lykken, *supra* note 13, at 23 ("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. Raskin, "Does Science Support Polygraph Testing?," in *The Polygraph Test: Lies, Truth and Science* 96, 99 (A. Gale ed. 1988).

addition, some of the laboratory studies fail to replicate field conditions; they use neither experienced examiners nor general population samples as subjects.

A different issue concerns the purpose of a study, which turns on whether a study is testing for validity or reliability. The term "validity" refers to the ability of a test procedure to measure what it is supposed to measure—its accuracy. The term "reliability" refers to whether the same results are obtained each time the test is performed—its consistency.[59] Validity includes reliability, but the converse is not necessarily true. Some studies test for validity, while others test for reliability. 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 treated.

Other Issues

Understanding the literature on this subject also requires an appreciation of a number of additional issues.

First, the studies distinguish between false positives and false negatives. The former concerns a conclusion of deception for an innocent subject, while the latter involves erroneous exculpation of a guilty suspect. Some tests, such as the RI and CQT tests appear to be "more accurate at detecting the deception of the guilty person than detecting the truthfulness of the innocent person."[60]

Second, the subject of the test—whether a suspect or victim—may also be important. Apparently, false positives are more frequent with victims than with suspects.[61]

Third, the type of issue involved is important: polygraph examinations involving specific factual issues produce more valid results than those involving mental state issues.[62]

Selected Studies

A number of authorities have questioned the validity of polygraph testing. Although Dr. David Lykken is perhaps the most well-known critic,[63] he is by no means alone.[64] Other authorities, however, strongly support the validity of polygraph testing, at least under certain conditions.[65] A

---

[59] Barland, "The Reliability of Polygraph Chart Evaluation," in *Legal Admissibility of the Polygraph* 120, 121 (N. Ansley 1975).

[60] Barland, "The Polygraph Test in the USA and Elsewhere," *The Polygraph Test: Lies, Truth and Science* 73 (A. Gale ed. 1988).

[61] Raskin, "Does Science Support Polygraph Testing?," *The Polygraph Test: Lies, Truth and Science* 96, 101 (A. Gale ed. 1988).

[62] Raskin, *supra* note 19, 1986 Utah L. Rev. at 46-47.

[63] Dr. Lykken's writings include: D. Lykken, *A Tremor in the Blood: Uses and Abuses of the Lie Detector* (1981); Lykken, "The Lie Detector and the Law," 8 Crim. Def. 19 (May-June 1981); Lykken, "The Validity of Tests: Caveat Emptor," 27 Jurimetrics J. 263 (Spring 1987); Lykken, "The Case Against Polygraph Testing, in *The Polygraph Test: Lies, Truth and Science* 110 (A. Gale ed. 1988).

[64] Other critics include: Carroll, "How Accurate is Polygraph Lie Detection?," in *The Polygraph Test: Lies, Truth and Science* 19 (A. Gale ed. 1988); Kleinmuntz & Szucko, "A Field Study of the Fallibility of Polygraphic Lie Detection," 308 Nature 449 (1984) (the validity of polygraphic interrogation has yet to be established); Kleinmuntz & Szucko, *supra* note 13.

[65] See J. Reid & F. Inbau, *supra* note 14, at 304; D. Raski, G. Barland & J. Podlesny, *Validity and Reliability of Detection of Deception* (June 1978); Raskin, "Does Science Support Polygraph Testing?," *The Polygraph Test: Lies, Truth and Science* 96, 101 (A. Gale ed. 1988); Barland, "The Polygraph Test

comprehensive discussion of all the validation studies is beyond the scope of this article. Instead, selected excerpts of several reports are presented in chronological order.

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 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.[66]

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.[67]

A significant problem in interpreting the various reports concerns which studies used scientifically valid methodologies.[68] In 1986 an expert in this field cited five mock crime studies that he considered valid because they used the control question technique, trained examiners, field techniques, and nonstudent populations. He concluded:

> The combined accuracy of decisions was 95%, with an inconclusive rate of 8%. It should be noted

---

in the USA and Elsewhere," *id.* at 83; Raskin, *supra* 19, 1986 Utah L. Rev.; Raskin & Kircher, "The Validity of Lykken's Criticisms: Fact or Fancy?," 27 Jurimetrics J. 271 (Spring 1987).

[66] OTA Report, *supra* note 11, reprinted in 12 Polygraph at 200. For other articles and reports on the validity issue, see Abrams, "Polygraph Validity and Reliability: A Review," 18 J. Forensic Sci. 313 (1973); Ansley, "A Compendium on Polygraph Validity," 12 Polygraph 53 (1983); Horvath, "Detection of Deception: A Review of Field and Laboratory Research," 5 Polygraph 107 (1976).

[67] Department of Defense, "The Accuracy and Utility of Polygraph Testing," *supra* note 53, reprinted in 13 Polygraph at 63.

Moreover, a 1982 Gallup poll survey of the Society for Psychological Research reported that 61 percent of the 155 members responding believed that the polygraph is a useful diagnostic tool when considered with other available information. Gallup Organization, "Survey of Members of the Society for Psychological Research Concerning Their Opinion of Polygraph Test Interpretation," 13 Polygraph 153, 157 (1984).

[68] See Kircher, Horowitz & Raskin, "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 percent of the observed variance in detection rates).

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). 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.[69]

In 1988 that writer 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 examinations, 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 individual 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. 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.[70]

Perhaps the field examiners, either consciously or unconsciously, used clinical impressions to minimize false positive errors.

Nevertheless, the controversy continues. Two other writers concluded in 1988 that the "best defense one can offer for the continued use of the CQT is that its accuracy is indeterminate."[71] The writer of the first study 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."[72]

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 the results conflicting."[73] Here again, a number of factors must be understood.

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

---

[69] Raskin, *supra* note 19, 1986 Utah L. Rev. at 42.

[70] D. Raskin, J. Kircher, C. Honts & S. Horowitz, A Study of the Validity of Polygraph Examinations in Criminal Investigation (May 1988) (Nat'l Inst. of Justice, Grant No. 85-IJ-CX-0040).

[71] Iacono & Patrick, "Assessing Deception: Polygraph Techniques," in *Clinical Assessment of Malingering and Deception*, 205, 233 (R. Rogers ed. 1988).

[72] Raskin, "Polygraph Techniques for the Detection of Deception," in *Psychological Methods in Criminal Investigations and Evidence* 247, 290 (D. Raskin ed. 1989).

[73] OTA Report, *supra* note 11, reprinted in 12 Polygraph at 201.

effects at specific points in the examination such as mental imagery and physical countermeasures.[74] In addition, countermeasures that may be successful against one type of examination, such as the CQT test, would not necessarily be effective against the concealed information test.

Physical countermeasures appear to pose the greatest threat to the control quesiton technique.[75] 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."[76] Spontaneous countermeasures were not effective. Another study concluded that a polygraph machine with an activity sensor would detect most attempts at physical countermeasures.[77]

The Friendly Polygrapher Issue

The friendly polygrapher hypothesis[78] 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 if he fails the test, he will suffer no 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."[79]

---

[74] Honts, "Interpreting Research on Polygraph Countermeasures," 15 J. Police Sci. & Admin. 204 (1987). See also Gudjonsson, "How to Defeat the Polygraph Tests," in *The Polygraph Test: Lies, Truth and Science* 126 (A. Gale ed. 1988).

[75] Raskin, *supra* note 19, 1986 Utah L. Rev. at 50–51.

[76] Honts, Raskin, Kircher & Hodes, "Effects of Spontaneous Countermeasures on the Physiological Detection of Decepton," 16 J. Police Sci. & Admin. 91, 93 (1988).

[77] Abrams & Davidson, "Counter-Countermeasures in Polygraph Testing," 17 Polygraph 16, 19 (1988).

[78] This hypothesis was first suggested by Dr. Martin Orne. See Orne, "Implications of Laboratory Research for the Detection of Deception," in *Legal Admissibility of the Polygraph* 94 (N. Ansley ed. 1975).

[79] Barland, *supra* note 21, 16 U.West L.A. L. Rev. at 49. See also Raskin, *supra* note 19, 1986 Utah at 63 ("When the 'friendly polygrapher' hypothesis is examined in light of all meaningful scientific data, no credible evidence supports the theory.").