Polygraph

VOLUME 27 1998 NUMBER 1

Contents

Published Quarterly

©American Polygraph Association, 1998

P.O. Box 8037, Chattanooga, Tennessee 37414-0037

Editorial

In most behavioral science journals, the new editor traditionally begins his or her tenure by outlining for the readership the editor's views regarding the particular discipline. This is to prepare readers for some measure of inevitable style changes in the publication, and to provide some vision as to where the publication is going. Such is the case here. Your new editor has stepped into some very large shoes by taking over the American Polygraph Association publications from Norman Ansley, a man who has influenced our discipline as few have. The journal will continue to benefit from Norm in his new role as Associate Editor, as well as from the other distinguished Associate Editors. I am grateful for their selfless service. I, indeed we, are indebted to each of these folks.

Polygraph is unique publication. It serves the needs not only of practicing examiners, but also of scientists, lawyers, elected leaders, investigators, law enforcement officials, and government programs. With such a varied readership, the content of the journal is necessarily broad in scope. We will maintain the tradition of diversity in the material published. The categories of papers we would consider publishing include; research reports, legal briefs and reviews, theoretical papers, book reviews, literature reviews, technical and instructional papers, foreign reports, case histories, taxonomies, monologues and commentaries, short reports, and bibliographies. If you think you have something you would like to submit, we'd welcome a call, e-mail, or a manuscript. We'll offer guidance where we can. Each and every paper will be carefully read, and the authors will be notified of editorial actions and decisions.

With regard to editorial decisions, there are two key questions we will ask ourselves with every submission: is this work of interest to a segment of the readership, and does the writing meet the standards of clarity, accuracy, and style? The ordering of these questions is significant. If a paper does not address the needs of the discipline, it will be rejected irrespective of literary eloquence. However, we will strive to help writers meet the second criterion if their papers would add to the body of knowledge on psychophysiological detection of deception, and related areas. The standards of style appear regularly in this journal.

What is on the horizon? We live in a computerized age, and we will move toward more automation of the editorial process. For example, we encourage writers to send electronic versions of their manuscripts in addition to the hard copies so minor editing can be accomplished more quickly. We will also be available on e-mail to correspond with writers, to offer suggestions, to help them locate sources, or to fine tune papers. The style and appearance of this journal will be updated, and language standardized as much as possible. We will explore options for peer reviews for some papers. There will be special issues on individual topics. We will also approach experts in other disciplines who have something to contribute to polygraphy to write papers by invitation. And we will be responsive to suggestions by the membership for other enhancements.

One simple theme will be central to all of our decisions: this journal represents the members of the American Polygraph Association, and must continue to be the best we can offer. I am honored to serve as your editor, and look forward to your comments, suggestions, and input.

Donald Krapohl

United States, Petitioner

v.

Edward G. Scheffer

No. 96-1133

United States Supreme Court.

[March 31, 1998]

JUSTICE THOMAS announced the judgment of the Court and delivered the opinion of the Court with respect to Parts I, II-A, and II-D, and an opinion with respect to Parts II-B and II-C, in which THE CHIEF JUSTICE, JUSTICE SCALIA and JUSTICE SOUTER joined.

This case presents the question whether Military Rule of Evidence 707, which makes polygraph evidence inadmissible in court-martial proceedings, unconstitutionally abridges the right of accused members of the military to present a defense. We hold that it does not.

I

In March 1992, respondent Edward Scheffer, an airman stationed at March Air Force Base in California, volunteered to work as an informant on drug investigations for the Air Force Office of Special Investigations (OSI). His OSI supervisors advised him that, from time to time during the course of his undercover work, they would ask him to submit to drug testing and polygraph examinations. In early April, one of the OSI agents supervising respondent requested that he submit to a urine test. Shortly after providing the urine sample, but before the results of the test were known, respondent agreed to take a polygraph test administered by an OSI examiner. In the opinion of the examiner, the test "indicated no deception" when respondent denied using drugs since joining the Air Force.[1]

On April 30, respondent unaccountably failed to appear for work and could not be found on the base. He was absent without leave until May 13, when an Iowa state patrolman arrested him following a routine traffic stop and held him for return to the base. OSI agents later learned that respondent's urinalysis revealed the presence of methamphetamine.

Respondent was tried by general court-martial on charges of using methamphetamine, failing to go to his appointed place of duty, wrongfully absenting himself from the base for 13 days, and, with respect to an unrelated matter, uttering 17 insufficient funds checks. He testified at trial on his own behalf, relying upon an "innocent ingestion" theory and denying that he had knowingly used drugs while working for OSI. On cross-examination, the prosecution attempted to impeach

---

1

The OSI examiner asked three relevant questions: (1) 'Since you've been in the [Air Force], have you used any illegal drugs?'; (2) 'Have you lied about any of the drug information you've given OSI?'; (3) 'Besides your parents, have you told anyone you're assisting OSI?' Respondent answered "no' to each question. App. 12.

respondent with inconsistencies between his trial testimony and earlier statements he had made to OSI.

Respondent sought to introduce the polygraph evidence in support of his testimony that he did not knowingly use drugs. The military judge denied the motion, relying on Military Rule of Evidence 707, which provides, in relevant part:

"(a) Notwithstanding any other provision of law, the results of a polygraph examination, the opinion of a polygraph or any reference to an offer to take, failure to take, or taking of a polygraph examination, shall not be admitted into evidence."

The military judge determined that Rule 707 was constitutional because "the President may, through the Rules of Evidence, determine that credibility is not an area in which a fact finder needs help, and the polygraph is not a process that has sufficient scientific acceptability to be relevant."[2] App. 28. He further reasoned that the factfinder might give undue weight to the polygraph examiner's testimony, and that collateral arguments about such evidence could consume "an inordinate amount of time and expense." *Ibid.*

Respondent was convicted on all counts and was sentenced to a bad-conduct discharge, confinement for 30 months, total forfeiture of all pay and allowances, and reduction to the lowest enlisted grade. The Air Force Court of Criminal Appeals affirmed in all material respects, explaining that Rule 707 "does not arbitrarily limit the accused's ability to present reliable evidence." 41 M.J. 683, 691 (1995) (en banc).

By a 3-to-2 vote, the United States Court of Appeals for the Armed Forces reversed. 44 M.J. 442 (1996). Without pointing to any particular language in the Sixth Amendment, the Court of Appeals held that "[a] per se exclusion of polygraph evidence offered by an accused to rebut an attack on his credibility, ... violates his Sixth Amendment right to present a defense." *Id.*, at 445.[3] Judge Crawford, dissenting, stressed that a defendant's right to present relevant evidence is not absolute, that relevant evidence can be excluded for valid reasons, and that Rule 707 was supported by a number of valid justifications. *Id.*, at 449-451. We granted certiorari, 520 U.S. ____ (1997), and we now reverse.

II

A defendant's right to present relevant evidence is not unlimited, but rather is subject to reasonable restrictions.[4] *See, Taylor v. Illinois*, 484 U.S. 400, 410 (1988); *Rock v. Arkansas*, 483 U.S. 44, 55 (1987); *Chambers v. Mississippi*, 410 U.S. 284, 295 (1973). A defendant's interest in presenting such evidence

---

2

Article 36 of the Uniform Code of Military Justice authorizes the President, as Commander in Chief of the Armed Forces, *see* U.S. Const., Art. II, §2, to promulgate rules of evidence for military courts: "Pretrial, trial, and post-trial procedures, including modes of proof, ... may be prescribed by the President by regulations which shall, so far as he considers practicable, apply the principles of law and the rules of evidence generally recognized in the trial of criminal cases in the United States district courts." 10 U.S.C. §836(a).

3

In this Court, respondent cites the Sixth Amendment's Compulsory Process Clause as the specific constitutional provision supporting his claim. He also briefly contends that the "combined effect" of the Fifth and Sixth Amendments confers upon him the risk to a "meaningful opportunity to present a complete defense," *Crane v. Kentucky*, 476 U.S. 683, 690 (1986) (citations omitted), and that this right in turn encompasses a constitutional right to present polygraph evidence to bolster his credibility.

4

The words "defendant" and "jury" are used throughout in reference to general principles of law and in discussing nonmilitary precedents. In reference to this case or to the military specifically, the terms "court," "court members," or "court-martial" are used throughout, as is the military term, "accused," rather than the civilian term, "defendant."

may thus "bow to accommodate other legitimate interests in the criminal trial process." *Rock, supra*, at 55 (quoting Chambers, *supra*, at 295); accord *Michigan v. Lucas*, 500 U.S. 145, 149 (1991). As a result, state and federal rulemakers have broad latitude under the Constitution to establish rules excluding evidence from criminal trials. Such 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." *Rock, supra*, at 56; accord *Lucas, supra*, at 151. 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. *See Rock, supra*, at 58; *Chambers, supra*, at 302; *Washington v. Texas*, 388 U.S. 14, 22-23 (1967).

Rule 707 serves several legitimate interests in the criminal trial process. These interests include ensuring that only reliable evidence is introduced at trial, preserving the jury's role in determining credibility, and avoiding litigation that is collateral to the primary purpose of the trial.[5] The rule is neither arbitrary nor disproportionate in promoting these ends. Nor does it implicate a sufficiently weighty interest of the defendant to raise a constitutional concern under our precedents.

A

State and federal governments unquestionably have a legitimate interest in ensuring that reliable evidence is presented to the trier of fact in a criminal trial. Indeed, the exclusion of unreliable evidence is a principal objective of many evidentiary rules. *See, e.g., Fed. Rule Evid.* 702; *Fed. Rule Evid.* 802; *Fed. Rule Evid.* 901; see also *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 589 (1993).

The contentions of respondent and the dissent notwithstanding, there is simply no consensus that polygraph evidence is reliable. To this day, the scientific community remains extremely polarized about the reliability of polygraph techniques. 1 D. Faigman, D. Kaye, M. Saks, & J. Sanders, Modern Scientific Evidence 565, n. 7; 14-2.0, and § 14-3.0 (1997); see also 1 P. Giannelli & E. Imwinkelried, Scientific Evidence §8-2(C), pp. 225-227 (2d ed. 1993) (hereinafter Giannelli & Imwinkelried); 1 J. Strong, McCormick on Evidence §206, p. 909 (4th ed. 1992) (hereinafter McCormick). Some studies have concluded that polygraph tests overall are accurate and reliable. *See, e.g.*, S. Abrams, The Complete Polygraph Handbook 190-191 (1968) (reporting the overall accuracy rate from laboratory studies involving the common "control question technique" polygraph to be "in the range of 87 percent"). Others have found that polygraph tests assess truthfulness significantly less accurately that scientific field studies suggest the accuracy rate of the "control question technique" polygraph is "little better than could be obtained by the toss of a coin," that is, 50 percent. *See* Iacono & Lykken, The Scientific Status of Research on Polygraph Techniques: The Case Against Polygraph Tests, in 1 Modern

5

These interests, among others, were recognized by the drafters of Rule 707, who justified the Rule on the following grounds: the risk that court members would be misled by polygraph evidence; the right that the traditional responsibility of court members to ascertain the facts and adjudge guilt or innocence would be usurped; the danger that confusion of the issues "could result in the court-martial degenerating into a trial of the polygraph machine;" "the likely waste of time on collateral issues, and the fact that the" "reliability of polygraph evidence has not been sufficiently established." See 41 M.J. 683, 686 (USAF Cr. Crim. App. 1995) (citing *Manual for Courts-Martial, the United States App.*, pp. A22-A46 (1994 ed)).

Scientific Evidence, *supra*, §14-5.3, p. 629 (hereinafter Iacono & Lykken).[6]

This lack of scientific consensus is reflected in the disagreement among state and federal courts concerning both the admissibility and the reliability of polygraph evidence.[7] Although some Federal Courts of Appeal have abandoned the per se rule excluding polygraph evidence, leaving its admission or exclusion to the discretion of district courts under *Daubert, see, e.g., United States v. Posado*, 57 F.3d 428, 434 (CA5 1995); *United States v. Cordoba*, 104 F.3d 225, 228 (CA9 1997), at least one Federal Circuit has recently reaffirmed its per se ban, *see United States v. Sanchez*, 118 F.3d 192, 197 (CA4 1997), and another recently noted that it has "not decided whether polygraphy has reached a sufficient state of reliability to be admissible." *United States v. Messina*, 131 F.3d 36, 42 (CA2 1997). Most States maintain per se rules excluding polygraph evidence. *See, e.g, State v. Porter*, 241 Conn. 57, 92-95, 698 A.2d 739, 758-759 (1995); *People v. Gard*, 158 Ill.2d 191, 202-204, 632 N.E.2d 1026, 1032 (1994); *In re Odell*, 672 A.2d 457, 459 (RI 1996)(per curiam); *Perkins v. State*, 902 S.W.2d 88, 94-95 (Ct.App. Tex. 1995). New Mexico is unique in making polygraph evidence generally admissible without the prior stipulation of the parties and without significant restriction. *See N.M. Rule Evid.* §11-707.[8] Whatever their approach, state and federal courts continue to express doubt about whether such evidence is reliable. *See, e.g., United States v. Messina, supra*, at 42; *United States v. Posado, supra*, at 434; *State v. Porter, supra*, at 126-127, 698 A.2d at 774; *Perkins v. State, supra*, at 94; *People v. Gard, supra*, at 202-204, 632 N.E.2d at 1032; *In re Odell, supra*, at 459.

6

The United States notes that in 1983 Congress' Office of Technology Assessment evaluated all available studies on the reliability of polygraphs and concluded that " [o]verall, the cumulative research evidence suggests that when used in criminal investigations, the polygraph test detects deception better than chance, but with error rates that could be considered significant'" Brief for United States 21 (quoting U.S. Congress, Office of Technology Assessment, Scientific Validity of Polygraph Testing: A Research Review and Evaluation-A Technical Memorandum 5 (OTA-TM-H-15, Nov. 1983)). Respondent, however, contends current research shows polygraph testing is reliable more than 90 percent of the time. Brief for Respondent 22 (citing J. Matte, Forensic Psychophysiology Using the Polygraph, 121-129 (1996)). Even if the basic debate about the reliability of polygraph technology itself were resolved, however, there would still be controversy over the efficacy of countermeasures, or deliberately adopted strategies that a polygraph examinee can employ to provoke physiological responses that will obscure accurate readings and thus "fool" the polygraph machine and the examiner. *See, e.g.*, Iacono & Lykken §14-3.0.

7

Until quite recently, federal and state courts were uniform in categorically ruling polygraph evidence inadmissible under the test set forth in *Frye v. United States*, 293 F. 1013 (CADC 1923), which held that scientific evidence must gain the general acceptance of the relevant expert community to be admissible. In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), we held that *Frye* had been superseded by the Federal Rules of Evidence and that expert testimony could be admitted if the district court deemed it both relevant and reliable.

Prior to *Daubert*, neither federal nor state courts found any Sixth Amendment obstacle to the categorical rule. *See, e.g., Bashor v. Risley*, 730 F.2d 1228, 1238 (CA9), cert. denied, 469 U.S. 838 (1984); *People v. Price*, 1 Cal. 4th 324, 419-420, 821 P.2d 610, 663 (1991), cert. denied, 506 U.S. 851 (1992). Nothing in *Daubert* foreclosed, as a constitutional manner, per se exclusionary rules for certain types of expert or scientific evidence. It would be an odd inversion of our hierarchy of laws if altering or interpreting a rule of evidence worked a corresponding change in the meaning of the Constitution.

8

Respondent argues that because the Government--and in particular the Department of Defense--routinely uses polygraph testing, the Government must consider polygraphs reliable. Governmental use of polygraph tests, however, is primarily in the field of personnel screening, and to a lesser extent as a tool in criminal and intelligence investigations, but not as evidence at trials. See Brief for United States 34, n. 17; Barland, The Polygraph Test in the USA and Elsewhere, in The Polygraph Test 76 (A. Gale ed. 1988). Such limited out of court uses of polygraph techniques obviously differ in character from, and carry less severe consequences than, the use of polygraphs as evidence in a criminal trial. They do not establish the reliability of polygraphs as trial evidence, and they do not invalidate reliability as a valid concern supporting Rule 707's categorical ban.

The approach taken by the President in adopting Rule 707--excluding polygraph evidence in all military trials--is a rational and proportional means of advancing the legitimate interest in barring unreliable evidence. Although the degree of reliability of polygraph evidence may depend upon a variety of identifiable factors, there is simply no way to know in a particular case whether a polygraph examiner's conclusion is accurate, because certain doubts and uncertainties plague even the best polygraph exams. Individual jurisdictions therefore may reasonably reach differing conclusions as to whether polygraph evidence should be admitted. We cannot say, then, that presented with such widespread uncertainty, the President acted arbitrarily or disproportionately in promulgating a per se rule excluding all polygraph evidence.

B

It is equally clear that Rule 707 serves a second legitimate governmental interest: Preserving the jury's core function of making credibility determinations in criminal trials. A fundamental premise of our criminal trial system is that "the jury is the lie detector." *United States v. Barnard,* 490 F.2d 907, 912 (CA9 1973) (emphasis added), cert. denied, 416 U.S. 959 (1974). Determining the weight and credibility of witness testimony, therefore, has long been held to be the "part of every case [that] belongs to the jury, who are presumed to be fitted for it by their natural intelligence and their practical knowledge of men and the ways of men" *Aetna Life Ins. Co. v. Ward,* 140 U.S. 76, 88 (1891).

By its very nature, polygraph evidence may diminish the jury's role in making credibility determinations. The common form of polygraph test measures a variety of physiological responses to a set of questions asked by the examiner, who then interprets these physiological correlates of anxiety and offers an opinion to the jury about whether the witness--often, as in this case, the accused--was deceptive in answering questions about the very matter at issue in the trial. See 1 McCormick §206.[9] Unlike other expert witnesses who testify about factual matters outside the jurors' knowledge, such as the analysis of fingerprints, ballistics, or DNA found at a crime scene, a polygraph expert can supply the jury only with another opinion, in addition to its own, about whether the witness was telling the truth. Jurisdictions, in promulgating rules of evidence, may legitimately be concerned about the risk that juries will give excessive weight to the opinions of a polygrapher, clothed as they are in scientific expertise and at times offering, as in respondent's case, a conclusion about the ultimate issue in the trial. Such jurisdictions may legitimately determine that the aura of infallibility attending polygraph evidence can lead jurors to abandon their duty to assess credibility and guilt. Those jurisdictions may also take into account the fact that a judge cannot determine, when ruling on a motion to admit polygraph evidence, whether a particular polygraph expert is likely to influence the jury unduly. For these reasons, the President is within his

---

9

The examiner interprets various physiological responses of the examinee, including blood pressure, perspiration, and respiration, while asking a series of questions, commonly in three categories: direct accusatory questions concerning the matter under investigation, irrelevant or neutral questions, and more general "control" questions concerning wrongdoing by the subject in general. The examiner forms an opinion of the subject's truthfulness by comparing the physiological reactions to each set of questions. See generally Giannelli & Imwinkelried 219-222; Honts & Quick, The Polygraph in 1995: Progress in Science and the Law, 71 N.D.L.Rev. 987, 990-992 (1995).

constitutional prerogative to promulgate a per se rule that simply excludes all such evidence.

C

A third legitimate interest served by Rule 707 is avoiding litigation over issues other than the guilt or innocence of the accused. Such collateral litigation prolongs criminal trials and threatens to distract the jury from its central function of determining guilt or innocence. Allowing proffers of polygraph evidence would inevitably entail assessments of such issues as whether the test and control questions were appropriate, whether a particular polygraph examiner was qualified and had properly interpreted the physiological responses, and whether other factors such as countermeasures employed by the examinee had distorted the exam results. Such assessments would be required in each and every case.[10] It thus offends no constitutional principle for the President to conclude that a per se rule excluding all polygraph evidence is appropriate. Because litigation over the admissibility of polygraph evidence is by its very nature collateral, a per se rule prohibiting its admission is not an arbitrary or disproportionate means of avoiding it.[11]

D

The three of our precedents upon which the Court of Appeals principally relied, *Rock v. Arkansas*, *Washington v. Texas*, and *Chambers v. Mississippi*, do not support a right to introduce polygraph evidence, even in very narrow circumstances. The exclusions of evidence that we declared unconstitutional in those cases significantly undermined fundamental elements of the accused's defense. Such is not the case here.

In *Rock*, the defendant, accused of a killing to which she was the only eyewitness, was allegedly able to remember the facts of the killing only after having her memory hypnotically refreshed. *See Rock v. Arkansas*, 483 U.S., at 46. Because Arkansas excluded all hypnotically refreshed testimony, the defendant was unable to testify about certain relevant facts, including whether the killing had been accidental. *See id.*, at 47-49. In holding that the exclusion of this evidence violated the defendant's "right to present a defense," we noted that the rule deprived the jury of the testimony of the only witness who was at the scene and had firsthand knowledge of the facts. *See id.*, at 57. Moreover, the rule infringed upon the accused's interest in testifying in her own defense--an interest that we deemed particularly significant, as it is the defendant who is the target of any criminal prosecution. *See id.*, at 52. For this reason, we stated that an accused ought to be allowed 'to present his own version of events in his own words.' *Ibid.*

In *Washington*, the statutes involved prevented co-defendants or co-participants in a crime from testifying for one another and thus

10

Although some of this litigation could take place outside the presence of the jury, at the very least a foundation must be laid for the jury to assess the qualifications and skill of the polygrapher and the validity of the exam, and significant cross-examination could occur on these issues.

11

Although the Court of Appeals stated that it had merely remove[d] the obstacle of the per se rule against admissibility of polygraph evidence in cases where the accused wishes to proffer an exculpatory polygraph to rebut an attack on his credibility, 44 M.J. 442, 446 (1996), and respondent thus implicitly argues that the Constitution would require collateral litigation only in such cases, we cannot see a principled justification whereby a right derived from the Constitution could be so narrowly contained.

precluded the accused from introducing his accomplice's testimony that the accomplice had in fact committed the crime. *See Washington v. Texas*, 388 U.S., at 16-17. In reversing Washington's conviction, we held that the Sixth Amendment was violated because "the State arbitrarily denied [the accused] the right to put on the stand a witness who was physically and mentally capable of testifying to events that he had personally observed." *Id.* At 23.[12]

In *Chambers*, we found a due process violation in the combined application of Mississippi's common law "voucher rule," which prevented a party from impeaching his own witness, and its hearsay rule that excluded the testimony of three persons to whom that witness had confessed. *See Chambers v. Mississippi*, 410 U.S., at 302. *Chambers* specifically confined its holding to the "facts and circumstances" presented in that case; we thus stressed that the ruling did not "signal any diminution in the respect traditionally accorded to the States in the establishment and implementation of their own criminal trial rules and procedures," *Id.*, at 302-303. *Chambers* therefore does not stand for the proposition that the accused is denied a fair opportunity to defend himself whenever a state or federal rule excludes favorable evidence.

Rock, Washington, and *Chambers* do not require that Rule 707 be invalidated, because, unlike the evidentiary rules at issue in those cases, Rule 707 does not implicate any significant interest of the accused. Here, the court members heard all the relevant details of the charged offense from the perspective of the accused, and the Rule did not preclude him from introducing any factual evidence.[13] Rather, respondent was barred merely from introducing expert opinion testimony to bolster his own credibility. Moreover, in contrast to the rule at issue in *Rock*, Rule 707 did not prohibit respondent from testifying on his own behalf; he freely exercised his choice to convey his version of the facts to the court-martial members. We therefore cannot conclude that respondent's defense was significantly impaired by the exclusion of polygraph evidence. Rule 707 is thus constitutional under our precedents.

* * *

For the foregoing reasons, Military Rules of Evidence 707 does not unconstitutionally abridge the right to present a defense. The judgment of the Court of Appeals is reversed.

It is so ordered.

12

In addition, we noted that the State of Texas could advance no legitimate interests in support of the evidentiary rules at issue, and those rules burdened only the defense and not the prosecution. See 388 U.S., at 22-23. Rule 707 suffers from neither of these defects.

13

The dissent suggests, post, at 13, that polygraph results constitute "factual evidence." The raw results of a polygraph exam-the subject's pulse, respiration, and perspiration rates-may be factual data, but these are not introduced at trial, and even if they were, they would not be "facts" about the alleged crime at hand. Rather, the evidence introduced is the expert opinion testimony of the polygrapher about whether the subject was truthful or deceptive in answering questions about the alleged crime. A *per se* rule excluding polygraph results therefore does not prevent an accused-just as it did not prevent respondent here-from introducing factual evidence or testimony about the crime itself, such as alibi witness testimony, see post. at 12. For the same reasons, an expert polygrapher's interpretation of polygraph results is not evidence of "the accused's whole conduct." see *post*, at 18, to which Dean Wigmore referred. It is not evidence of the "accused's ... conduct" at all, much less "conduct"' concerning the actual crime at issue. It is merely the opinion of a witness with no knowledge about any of the facts surrounding the alleged crime, concerning whether the defendant spoke truthfully or deceptively on another occasion.

JUSTICE KENNEDY, with whom JUSTICE O'CONNOR, JUSTICE GINSBURG, and JUSTICE BREYER join, concurring in part and concurring in the judgment.

I join Parts I, II-A, and II-D of the opinion of the Court.

In my view it should have been sufficient to decide this case to observe, as the principle opinion does, that various courts and jurisdictions 'may reasonably reach differing conclusions as to whether polygraph evidence should be admitted.' *Ante*. At 8. The continuing, good-faith disagreement among experts and courts on the subject of polygraph reliability counsels against our invalidating a *per se* exclusion of polygraph results or of the fact an accused has taken or refused to take a polygraph examination. If we were to accept respondent's position, of course, our holding would bind state courts, as well as military and federal courts. Given the ongoing debate about polygraphs, I agree the rule of exclusion is not so arbitrary or disproportionate that it is unconstitutional.

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 or excluding scientific evidence is not a constitutional mandate, *see Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 587 (1993), there is some tension between that rule and our holding today. And, as JUSTICE STEVENS points out, 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.

With all respect, moreover, it seems the principal opinion overreaches when it rests its holding on the additional ground that the jury's role in making credibility determinations is diminished when it hears polygraph evidence. I am in substantial agreement with JUSTICE STEVENS' observation that the argument demeans and mistakes the role and competence of jurors in deciding the factual question of guilt or innocence. *Post*. At 18. In the last analysis the principal opinion says it is unwise to allow the jury to hear 'a conclusion about the ultimate issue in the trial.' *Ante*. At 10. I had thought this tired argument had long since been given its deserved repose as a categorical rule of exclusion. Rule 704(a) of the Federal Rules of Evidence states: "Except as provided in subdivision (b), testimony in the form of an opinion or inference otherwise admissible in not objectionable because it embraces an ultimate issue to be decided by the trier of fact.' The Advisory Committee's Notes state"

> The older cases often contained strictures against allowing witnesses to express opinions upon ultimate issues, as a particular aspect of the rule against opinions. The rule was unduly restrictive, difficult of application, and generally served only to deprive the trier of fact of useful information. 7 Wigmore §§1920, 1921; McCormick §12. The basis usually assigned for the rule, to prevent the witness from "usurping the province of the jury,"

> is aptly characterized as "empty rhetoric." 7 Wigmore §1920, p. 17.' Advisory Committee's Notes on Fed. Rule Evid. 704, 28 U.S.C., p. 888.

The principal opinion is made less convincing by its contradicting the rationale of Rule 704 and the well considered reasons the Advisory Committee recited in support of its adoption.

The attempt to revive this outmoded theory is especially inept in the context of the military justice system; for the one narrow exception to the abolition of the ultimate issue rule still surviving in the Federal Rules of Evidence has been omitted from the corresponding rule adopted for the military. The ultimate issue exception in the Federal Rules of Evidence is as follows:

> No expert witness testifying with respect to the mental state or condition of a defendant in a criminal case may state an opinion or inference as to whether the defendant did or did not have the mental state or condition constituting an element of the crime charged or of a defense thereto. Such ultimate issues are matters for the trier of fact alone. Fed. Rule Evid. 704(b).

The drafting committee for the Military Rules of Evidence renounced even this remnant. It said: "The statutory qualifications for military court members reduce the risk that military court members will be unduly influenced by the presentation of ultimate opinion testimony from psychiatric experts." Manual for Courts-Martial, United States, Analysis of the Military Rules of Evidence, p. A22-48 (1995 ed.). Any supposed need to protest the role of the finder of fact is diminished even further by this specific acknowledgment that members of military courts are not likely to give excessive weight to opinions of experts or otherwise to be misled or confused by their testimony. Neither in the federal system nor in the military courts, then, is it convincing to say that polygraph test results should be excluded because of some lingering concern about usurping the jury's responsibility to decide ultimate issues.

JUSTICE STEVENS, dissenting.

The United States Court of Military Appeals held that the President violated the Constitution in June, 1991, when he promulgated Rule 707 of the Military Rules of Evidence. Had I been a member of that Court, I would not have decided that question without first requiring the parties to brief and argue the antecedent question whether Rule 707 violates Article 36(a) of the Uniform Code of Military Justice, 10 U.S.C. §836(a). As presently advised, I am persuaded that the Rule does violate the statute and should be held invalid for that reason. I also agree with the Court of Appeals that the Rule is unconstitutional. This Court's contrary holding rests on a serious undervaluation of the importance of the citizen's constitutional right to present a defense to a criminal charge and an unrealistic appraisal of the importance of the governmental interests that undergird the Rule. Before discussing the constitutional issue, I shall comment briefly on the statutory question.

I

Rule 707 is a blanket rule of exclusion.[14] No matter how reliable and how probative the results of a polygraph test may be, Rule 707 categorically denies the defendant any opportunity to persuade the court that the evidence should be received for any purpose. Indeed, even if the parties stipulate in advance that the results of a lie detector test may be admitted, the Rule requires exclusion.

The principal charge against the respondent in this case was that he had knowingly used methamphetamine. His principal defense was "innocent ingestion"; even if the urinalysis tests conducted on April 7, 1992, correctly indicated that he did ingest the substance, he claims to have been unaware of that fact. The results of the lie detector test conducted three days later, if accurate, constitute factual evidence that his physical condition at that time was consistent with the theory of his defense and inconsistent with the theory of the prosecution. The results were also relevant because they tended to confirm the credibility of his testimony. Under Rule 707, even if the results of the polygraph test were more reliable than the results of the urinalysis, the weaker evidence is admissible and the stronger evidence is inadmissible.

Under the now discredited reasoning in a case decided 75 years ago, *Frye v. United States*, 54 App. D.C. 46, 293 F. 1013 (1923), that anomalous result would also have been reached in non-military cases tried in the federal courts. In recent years, however, we have not only repudiated *Frye*'s general approach to scientific evidence, but the federal courts have also been engaged in the process of rejecting the once-popular view that all lie detector evidence should be categorically inadmissible.[15] Well reasoned opinions are concluding, consistently with this Court's decision, in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), and *General Electric Co. V. Joiner*, 522 U.S. ___ (1997), that the federal rules wisely allow district judges to exercise broad discretion when evaluating the admissibility of scientific evidence.[16] Those opinions correctly observe that the rules of evidence generally recognized in the trial of civil and criminal cases to the federal courts do not contain any blanket prohibition against the admissibility of polygraph evidence.

14

Rule 707 states, in relevant part:

"Notwithstanding any other provision of law, the results of a polygraph examination, the opinion of a polygraph examiner, or any reference to an offer to take, failure to take, or taking of a polygraph examination, shall not be admitted into evidence." Mil. Rule Evid. 707(a).

15

"There is no question that in recent years polygraph testing has gained increasingly widespread acceptance as a useful and reliable scientific tool. Because of the advances that have been achieved in the field which have led to the greater use of polygraph examination, coupled with a lack of evidence that juries are unduly swayed by polygraph evidence, we agree with those courts which have found that a per se rule disallowing polygraph evidence is not longer warranted. ... Thus, we believe the best approach in this area is one which balances the need to admit all relevant and reliable evidence against the danger that the admission of the evidence for a given purpose will be unfairly prejudicial." *United States v. Piccinonna*, 885 F.2d 1529, 1535 (CA11 1989). "[W]e do not now hold that polygraph examinations are scientifically valid or that they will always assist the trier of fact, in this or any other individual case. 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." *United States v. Posado*, 57 F.3d 428, 434 (CA5 1995).

16

"The per se ... rule excluding unstipulated polygraph evidence is inconsistent with the "flexible inquiry" assigned to the trial judge by *Daubert*. This is particularly evident because *Frye*, which was overruled by *Daubert*, involved the admissibility of polygraph evidence." *United States v. Cordoba*, 104 F.3d 225, 227 (CA9 1997).

In accord with the modern trend of decisions on this admissibility issue, in 1987 the Court of Military Appeals held that an accused was "entitled to attempt to lay" the foundation for admission of favorable polygraph evidence. *United States v. Gipson*, 24 M.J. 246, 253 (1987). The President responded to Gipson by adopting Rule 707. The governing statute authorized him to promulgate evidentiary rules "which shall, so far as he considers practicable, apply the principles of law and the rules of evidence generally recognized in the trial of criminal cases in the United States district courts." 10 U.S.C., §836(a).[17] Thus, if there are military concerns that warrant a special rule for military tribunals, the statute gives him ample authority to promulgate special rules that take such concerns into account.

Rule 707 has no counterpart in either the Federal Rules of Evidence or the Federal Rules of Criminal Procedure. Moreover, to the extent that the use of the lie detector plays a special role in the military establishment, military practices are more favorable to a rule of admissibility than is the less structured use of lie detectors in the civilian sector of our society. That is so because the military carefully regulates the administration of polygraph tests to ensure reliable results. The military maintains 'very stringent standards for polygraph examiners"[18] and has established its own Polygraph Institute, which is "generally considered to be the best training facility for polygraph examiners in the United States."[19] The military has administered hundreds of thousands of such tests and routinely

17

"Pretrial, trial, and post-trial procedures, including modes of proof, for cases arising under this chapter triable in courts-martial, military commissions and other military tribunals, and procedures for courts of inquiry, may be prescribed by the President by regulations which shall, so far as he considers practicable, apply the principles of law and the rules of evidence generally recognized in the trial of criminal cases in the United States district courts, but which may not be contrary to or inconsistent with this chapter." 10 U.S.C. §836(a).

18

According to the Department of Defense's 1996 Report to Congress: "The Department of Defense maintains very stringent standards for polygraph examiners. The Department of Defense Polygraph Institute's basic polygraph program 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. Candidates selected for the Department of Defense polygraph position must meet the following minimum requirements:

"1. Be a United States citizen.

"2. Be at least 25 years of age.

"3. Be a graduate of an accredited four-year college or have equivalent experience that demonstrates the ability to master graduate-level academic courses.

"4. Have two years of experience as an investigator with a Federal or other law enforcement agency ...

"5. Be of high moral character and sound emotional temperament, as confirmed by a background investigation.

"6. Complete a Department of Defense-approved course of polygraph instruction.

"7. Be adjudged suitable for the position after being administered a polygraph examination designed to ensure that the candidate realizes, and is sensitive to, the personal impact of such examinations.

"All federal polygraph examiners receive their basic polygraph training at the Department of Defense Polygraph Institute. After completing the basic polygraph training, DoD personnel must serve an internship consisting of a minimum of six months on-the-job-training and conduct at least 25 polygraph examinations under the supervision of a certified polygraph examiner before being certified as a Department of Defense polygraph examiner. In addition, DoD polygraph examiners are required to complete 80 hours of continuing education every two years." Department of Defense Polygraph Program, Annual Polygraph Report to Congress, Fiscal Year 1996, pp. 14-15; see also Yankee, The Current Status of Research in Forensic Psychophysiology and Its Application in the Psychophysiological Detection of Deception, 40 J. Forensic Sciences 63 (1995).

19

Honts & Perry, Polygraph Admissibility: Changes and Challenges, 16 Law and Human Behavior 357, 359, n. 1 (1992) (hereinafter Honts & Perry).

uses their results for a wide variety of official decisions.[20]

The stated reasons for the adoption of Rule 707 do not rely on any special military concern. They merely invoke three interests: (1) the interest in excluding unreliable evidence; (2) the interest in protecting the trier of fact from being misled by an unwarranted assumption that the polygraph evidence has "an aura of near infallibility"; and (3) the interest in avoiding collateral debates about the admissibility of particular test results.

It seems clear that those interests pose less serious concerns in the military than in the civilian context. Disputes about the qualifications of the examiners, the equipment, and the testing procedures should seldom arise with respect to the tests conducted by the military. Moreover, there surely is no reason to assume that military personnel who perform the fact-finding function are less competent than ordinary jurors to assess the reliability of particular results, or their relevant to the issues.[21] Thus, there is no identifiable military concern that justifies the President's promulgation of a special military rule that is more burdensome to defendants in military trials than the evidentiary rules applicable to the trial of civilians.

It, therefore, seems fairly clear that Rule 707 does not comply with the statute. I do not rest on this ground, however, because briefing might persuade me to change my views, and because the Court has decided only the constitutional question.

II

The Court's opinion barely acknowledges that a person accused of a crime has a constitutional right to present a defense. It is not necessary to point to "any particular language in the Sixth Amendment," *ante*, at 3, to support the conclusion that the right is firmly established. It is, however, appropriate to comment on the importance of that right before discussing the three interests that the Government relies upon to justify Rule 707.

The Sixth Amendment provides that "the accused shall enjoy the right ... to have compulsory process for obtaining witnesses in his favor." Because this right "is an essential attribute of the adversary system itself," we have repeatedly stated that few rights "are more fundamental than that of an accused to present witnesses in his own defense."[22] According to Joseph Story, that provision was included in the Bill of Rights in reaction to a

---

20

Between 1981 and 1997, the Department of Defense conducted over 400,000 polygraph examinations to resolve issues arising in counterintelligence, security, and criminal investigations. Department of Defense Polygraph Program, Annual Polygraph Report to Congress. Fiscal Year 1997, p. 1; *id.*, Fiscal Year 1996, p. 1, *id.*; Fiscal Year 1995, p. 1, *id.*; Fiscal Year 1994, p. 1; *id.*, Fiscal Year 1993, App. A; id., Fiscal Year 1992, App. A; *id.*, Fiscal Year 1991, App. A-1 (reporting information for 1981-1991).

21

When the members of the court-martial are officers, as was true in this case, they typically have at least a college degree as well as significant military service. See 10 U.S.C. §825(d)(2); see also, *e.g.*, *United States v. Carter*, 22 M.J. 771, 776 (A.C.M.R. 1986).

22

"Few rights are more fundamental than that of an accused to present witnesses in his own defense, see, *e.g.*,*Chambers v. Mississippi*, 410 U.S. 284, 302 (1973). Indeed, this right is an essential attribute of the adversary system itself. ... The right to compel a witness' presence in the courtroom could not protect the integrity of the adversary process if it did not embrace the right to have the witness' testimony heard by the trier of fact. The right to offer testimony is thus grounded in the Sixth Amendment. ..." *Taylor v. Illinois*, 484 U.S. 400, 408-409 (1988).

notorious common-law rule categorically excluding defense evidence in treason and felony cases.[23] Our holding in *Washington v. Texas*, 388 U.S. 14 (1967), that this right is applicable to the States, rested on the premises that it "is in plain terms the right to present a defense" and that it "is a fundamental element of due process of law."[24] Consistent with the history of the provision, the Court in that case held that a state rule of evidence that excluded "whole categories" of testimony on the basis of a presumption of unreliability was unconstitutional.[25]

The blanket rule of inadmissibility held invalid in *Washington v. Texas* covered the testimony of alleged accomplices. Both before and after that decision, the Court has recognized the potential injustice produced by rules that exclude entire categories of relevant evidence that is potentially unreliable. At common law interested parties such as defendants,[26] their spouses, [27] and their co-conspirators[28] were not competent witnesses. "Nor were those named the only grounds of exclusion from the witness stand; conviction of crime, want of religious belief, and other matters were held sufficient. Indeed, the theory of the common law was to admit to the witness stand only those presumably honest, appreciating the sanctity of an oath, unaffected as a party by the result, and free from

---

23

"Joseph Story, in his famous Commentaries on the Constitution of the United States, observed that the right to compulsory process was included in the Bill of Rights in reaction to the notorious common-law rule that in cases of treason or felony the accused was not allowed to introduce witnesses in his defense at all. Although the absolute prohibition of witnesses for the defense had been abolished in England by statute before 1787, the Framers of the Constitution felt it necessary specifically to provide that defendants in criminal cases should be provided the means of obtaining witnesses so that their own evidence, as well as the prosecution's, might be evaluated by the jury." *Washington v. Texas*, 388 U.S. 14, 19-20 (1967) (footnotes omitted).

24

"The right to offer the testimony of witnesses, and to compel their attendance, if necessary, is in plain terms the right to present a defense, the right to present the defendant's version of the facts as well as the prosecution's to the jury so it may decide where the truth lies. Just as an accused has the right to confront the prosecution's witnesses for the purpose of challenging their testimony, he has the right to present his own witnesses to establish a defense. This right is a fundamental element of due process of law." *Id.*., at 19.

25

"It is difficult to see how the Constitution is any less violated by arbitrary rules that prevent whole categories of defense witnesses from testifying on the basis of a priori categories than presume them unworthy of belief.
"The rule disqualifying an alleged accomplice from testifying on behalf of the defendant cannot even be defended on the ground that it rationally sets apart a group of persons who are particularly likely to commit perjury." *Id.*, at 22.

26

"It is familiar knowledge that the old common law carefully excluded from the witness stand parties to the record, and those who were interested in the result; and this rule extended to both civil and criminal cases. Fear of perjury was the reason for the rule." *Benson v. United States*, 146 U.S. 325, 335 (1892).

27

"The common-law rule, accepted at an early date as controlling in this country, was that husband and wife were incompetent as witnesses for or against each other. ..."
"The Court recognized that the basic reason underlying th[e] exclusion [of one spouse's testimony on behalf of the other] had been the practice of disqualifying witnesses with a personal interest in the outcome of a case. Widespread disqualifications because of interest, however, had long since been abolished both in this country and in England in accordance with the modern trend which permitted interested witnesses to testify and left it for the jury to assess their credibility. Certainly, since defendants were uniformly allowed to testify in their own behalf, there was no longer a good reason to prevent them from using their spouses as witnesses. With the original reason for barring favorable testimony of spouses gone the Court concluded that this aspect of the old rule should go too." *Hawkins v. United States*, 358 U.S. 74, 75-76 (1958).

28

See *Washington v. Texas*, 388 U.S., at 20-21.

any of the temptations of interest. The courts were afraid to trust the intelligence of jurors." *Benson v. United States*, 146 U.S. 325, 326 (1892). And, of course, under the regime established by *Frye v. United States*, scientific evidence was inadmissible unless it met a stringent "general acceptance" test. Over the years, with respect to category after category, strict rules of exclusion have been replaced by rules that broaden the discretion of trial judges to admit potentially unreliable evidence and to allow properly instructed juries to evaluate its weight. While that trend has included both rulemaking and non-constitutional judicial decisions, the direction of the trend has been consistent and it has been manifested in constitutional holdings as well.

Commenting on the trend that had followed the decision in Benson, the Court in 1918 observed that in the

"Years which have elapsed since the decision of the *Benson* Case, the disposition of courts and of legislative bodies to remove disabilities from witnesses has continued, as that decision shows it had been going forward before, under dominance of the conviction of our time that the truth is more likely to be arrived at by hearing the testimony of all persons of competent understanding who may seem to have knowledge of the facts involved in a case, leaving the credit and weight of such testimony to be determined by the jury or by the court, rather than by rejecting witnesses as incompetent, with the result that this principle has come to be widely, almost universally, accepted in this country and in Great Britain." *Rosen v. United States*, 245 U.S. 467, 471 (1918).

See also, *Funk v. United States*, 290 U.S. 371, 377-378 (1933). It was in a case involving the disqualification of spousal testimony that Justice Stewart stated: 'Any rule that impedes the discovery of truth in a court of law impedes as well the doing of justice.' *Hawkins v. United States*, 358 U.S. 74, 81 (1958) (Stewart, J., concurring).

State evidentiary rules may so seriously impede the discovery of truth, "as well as the doing of justice," that they preclude the "meaningful opportunity to present a complete defense" that is guaranteed by the Constitution, *Crane v. Kentucky*, 476 U.S. 683, 690 (1986) (internal quotation marks omitted).[29] In *Chambers v. Mississippi*, 410 U.S. 284, 302 (1973), we concluded that 'where constitutional rights directly affecting the ascertainment of guilt are implicated, the hearsay rule may not be applied mechanistically to

---

29

"Whether rooted directly in the Due Process Clause of the Fourteenth Amendment, *Chambers v. Mississippi*, [410 U.S. 284 (1973)], or in the Compulsory Process or Confrontation clauses of the Sixth Amendment, *Washington v. Texas*, 388 U.S. 14, 23 (1967); *Davis v. Alaska*, 415 U.S. 308 (1974), the Constitution guarantees criminal defendants 'a meaningful opportunity to present a complete defense." *California v. Trombetta*, 467 U.S. [479, 485 (1984)]; cf. *Strickland v. Washington*, 466 U.S. 668, 684-685 (1984) ("The Constitution guarantees a fair trial through the Due Process Clauses, but it defines the basic elements of a fair trial largely through the several provisions of the Sixth Amendment"). We break no new ground in observing that an essential component of procedural fairness is an opportunity to be heard. *In re Oliver*, 333 U.S. 257, 273 (1948); *Grannis v. Ordean*, 234 U.S. 385, 394 (1914). That opportunity would be an empty one if the State were permitted to exclude competent, reliable evidence bearing on the credibility of a confession when such evidence is central to the defendant's claim of innocence. In the absence of any valid state justification, exclusion of this kind of exculpatory evidence deprives a defendant of the basic right to have the prosecutor's case encounter and "survive the crucible of meaningful adversarial testing." *United States v. Cronic*, 466 U.S. 648, 656 (1984). See also *Washington v. Texas*, *supra*, at 22-23.' *Crane v. Kentucky*, 476 U.S. 683, 690-691 (1986).

defeat the ends of justice.'[30] As the Court notes today, restrictions on the "defendant's right to present relevant evidence," *ante*, at 4, must comply with the admonition in *Rock v. Arkansas*, 483 U.S. 44, 56 (1987), that they "may not be arbitrary or disproportionate to the purposes they are designed to serve." Applying that admonition to Arkansas' blanket rule prohibiting the admission of hypnotically refreshed testimony, we concluded that a "State's legitimate interest in barring unreliable evidence does not extent to per se exclusions that may be reliable in an individual case." *Id.*, at 61. That statement of constitutional law is directly relevant to this case.

III

The constitutional requirement that a blanket exclusion of potentially unreliable evidence must be proportionate to the purposes served by the rule obviously makes it necessary to evaluate the interests on both sides of the balance. Today the Court all but ignores the strength of the defendant's interest in having polygraph evidence admitted in certain cases. As the facts of this case illustrate, the Court is quite wrong in assuming that the impact of Rule 707 on respondent's defense was not significant because it did not preclude the introduction of any "factual evidence" or prevent him from conveying "his version of the facts to the court-martial members." *Ante*, at 13. Under such reasoning, a rule that excluded the testimony of alibi witnesses would not be significant as long as the defendant is free to testify himself. But given the defendant's strong interest in the outcome-an interest that was sufficient to make his testimony presumptively untrustworthy and therefore inadmissible at common law-his uncorroborated testimony is certain to be less persuasive than that of a third-party witness. A rule that bars him "from introducing expert opinion testimony to bolster his own credibility," *ibid.*, unquestionably impairs any "meaningful opportunity to present a complete defense"; indeed, it is sure to be outcome-determinative in many cases.

Moreover, in this case the results of the polygraph test, taken just three days after the urinalysis, constitute independent factual evidence that is not otherwise available and that strongly supports his defense of "innocent ingestion." Just as flight or other evidence of "consciousness of guilt" may sometimes be relevant, on some occasions evidence of "consciousness of innocence" may also be relevant to the central issue at trial. Both the answers to the questions propounded by the examiner, and the physical manifestations produced by those utterances, were probative of an innocent state of mind shortly after he ingested the drugs. In Dean Wigmore's view, both "conduct" and "utterances" may constitute factual evidence of a "consciousness of

30

"Few rights are more fundamental than that of an accused to present witnesses in his own defense. *E.g., Webb v. Texas*, 409 U.S. 95 (1972); *Washington v. Texas*, 388 U.S. 14, 19 (1967); *In re Oliver*, 333 U.S. 257 (1948). In the exercise of this right, the accused, as is required of the State, must comply with established rules of procedure and evidence designed to assure both fairness and reliability in the ascertainment of guilt and innocence. Although perhaps no rule of evidence has been more respected or more frequently applied in jury trials than that applicable to the exclusion of hearsay, exceptions tailored to allow the introduction of evidence which in fact is likely to be trustworthy have long existed. The testimony rejected by the trial court here bore persuasive assurances of trustworthiness and thus was well within the basic rationale of the exception for declaration against interest. That testimony also was critical to Chambers' defense. In these circumstances, where constitutional rights directly affecting the ascertainment of guilt are implicated, the hearsay rule may not be applied mechanistically to defeat the ends of justice." *Chambers v. Mississippi*, 410 U.S. 284, 302 (1973).

innocence."[31] As the Second Circuit has held, when there is a serious factual dispute over the "basic defense [that defendant] was unaware of any criminal wrongdoing," evidence of his innocent state of mind is "critical to a fair adjudication of criminal charges."[32] The exclusion of the test results in this case cannot be fairly equated with a ruling that merely prevented the defendant from encumbering the record with cumulative evidence. Because the Rule may well have affected the outcome of the trial, it unquestionably "infringed upon a weighty interest of the accused." *Ante*, at 4-5.

The question, then, is whether the three interests on which the Government relies are powerful enough to support a categorical rule excluding the results of all polygraph tests no matter how unfair such a rule may be in particular cases.

Reliability

There are a host of studies that place the reliability of polygraph tests at 85% to 90%.[33] While critics of the polygraph argue that accuracy is much lower, even the studies cited by the critics place polygraph accuracy at 70%.[34] Moreover, to the extent that the polygraph errs, studies have repeatedly shown that the polygraph is more likely to find innocent people guilty than vice versa.[35] Thus, exculpatory polygraphs-like the one in this case-are likely to be more reliable than inculpatory ones.

Of course, within the broad category of lie detector evidence, there may be a wide variation in both the validity and the relevance[36] of

---

31

"Moreover, there are other principles by which a defendant may occasionally avail himself of conduct as evidence in his favor-in particular, of conduct indicating consciousness of innocence. ... of utterances asserting his innocence ..., and, in sedition charges, of conduct indicating a loyal state of mind. ..." 1 A.J. Wigmore, Evidence §56.1, p. 1180 (Tiller rev. ed. 1983); see *United States v. Reifsteck*, 841 F.2d 701, 705 CA6 1988).

32

"Mariotta's basic defense was that he was unaware of any criminal wrongdoing at Wedtech, that he was an innocent victim of the machinations of the sophisticated businessmen whom he had brought into the company to handle its financial affairs. That defense was seriously in issue as to most of the charges against him, drawing considerable support from the evidence. ...
"With the credibility of the accusations about Mariotta's knowledge of wrongdoing seriously challenged, evidence of his denial of such knowledge in response to an opportunity to obtain immunity by admitting it and implicating others became highly significant to a fair presentation of his defense. ...
"Where evidence of a defendant's innocent state of mind, critical to a fair adjudication of criminal charges, is excluded, we have not hesitated to order a new trial." *United States v. Biaggi*, 909 F.2d 662, 691-692 (CA2 1990); see *United States v. Bucur*, 194 F.2d 297 (CA7 1952); *Herman v. United States*, 48 F.2d 479 (CA5 1931).

33

Raskin, Honts, & Kircher, The Scientific Status of Research on Polygraph Techniques: The Case for Polygraph Tests, in 1 Modern Scientific Evidence 572 (D. Faigman, D. Kaye, M. Saks, & J. Sanders, eds. 1997) (hereinafter Faigman) (compiling eight laboratory studies that place mean accuracy at approximately 90%); id., at 575 (compiling four field studies, scored by independent examiners, that place mean accuracy at 90.5%); Raskin, Honts, & Kircher, A Response to Professor Iacono and Lykken, in Faigman 627 (compiling six field studies, scored by original examiners, that place mean accuracy at 97.5%); S. Abrams, The Complete Polygraph Handbook 190-191 (1989) (compiling 13 laboratory studies that, excluding inconclusive results, place mean accuracy at 87%).

34

Iacono & Lykken, The Scientific Status of Research on Polygraph Techniques: The Case Against Polygraph Tests, in Faigman 608 (compiling three studies that place mean accuracy at 70%).

35

E.g., Iacono & Lykken, The Case Against Polygraph Tests, in Faigman 608-609; Raskin, Honts, & Kircher, A Response to Professor Iacono and Lykken, in Faigman 621; Honts & Perry 362; S. Abrams, The Complete Polygraph Handbook, at 187-188, 191.

36

See, *e.g.*, Judge Gonzalez's careful attention to the relevance inquiry in the proceedings on remand from the Court of Appeals decision in *Piccinonna*, 729 F.Supp. 1336 (SD Fla 1990).

particular test results. Questions about the examiner's integrity, independence, choice of questions, or training in the detection of deliberate attempts to provoke misleading physiological responses may justify exclusion of specific evidence. But such questions are properly addressed in adversary proceedings; they fall far short of justifying a blanket exclusion of this type of expert testimony.

There is no legal requirement that expert testimony must satisfy a particular degree of reliability to be admissible. Expert testimony about a defendant's "future dangerousness" to determine his eligibility for the death penalty, even if wrong "most of the time," is routinely admitted. *Barefoot v. Estelle*, 463 U.S. 880, 898-901 (1983). Studies indicate that handwriting analysis, and even fingerprint identifications, may be less trustworthy than polygraph evidence in certain cases.[37] And, of course, even highly dubious eyewitness testimony is, and should be, admitted and tested in the crucible of cross-examination. The Court's reliance on potential unreliability as a justification for a categorical rule of inadmissibility reveals that it is "overly pessimistic about the capabilities of the jury and of the adversary system generally. Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence." *Daubert*, 509 U.S., at 596.[38]

37

One study compared the accuracy of fingerprinting, handwriting analysis, polygraph tests, and eyewitness identification. The study consisted of 80 volunteers divided into 20 groups of 4. Fingerprints and handwriting samples were taken from all of the participants. In each group of four, one person was randomly assigned the role of "perpetrator"' The perpetrator was instructed to take an envelope to a building doorkeeper (who knew that he would later need to identify the perpetrator), sign a receipt, and pick up a package. After the"'crime," all participants were given a polygraph examination. The fingerprinting expert (comparing the original fingerprints with those on the envelope), the handwriting expert (comparing the original samples with the signed receipt), and the polygrapher (analyzing the tests) sought to identify the perpetrator of each group. In addition, two days after the "crime," the doorkeeper was asked to pick the picture of the perpetrator out of a set of four pictures. The results of the study demonstrate that polygraph evidence compares favorably with other types of evidence. Excluding "inconclusive" results from each test, the fingerprinting expert resolved 100% of the case correctly, the polygrapher resolved 95% of the cases correctly, the handwriting expert resolved 94% of the cases correctly, and the eyewitness resolved only 64% of the cases correctly. Interestingly, when "inconclusive" results were included, the polygraph test as more accurate than any of the other methods: The polygrapher resolved 90% of the cases correctly, compared with 85% for the handwriting expert, 35% for the eyewitness, and 20% for the fingerprinting expert. Widacki & Horvath, An Experimental Investigation of the Relative Validity and Utility of the Polygraph Technique and Three Other Common Methods of Criminal Identification, 23 J. Forensic Sciences 596, 596-600 (1978); see also Honts & Perry 365.

38

The Government argues that there is a widespread danger that people will learn to "fool" the polygraph, and that this possibility undermines any claim of reliability. For example, the Government points to the availability of a book called "Beat the Box: The Insider's Guide to Outwitting the Lie Detector." Tr. Of Oral Arg. 53; Brief for United States 25, n. 10. 'Beat the Box,' however, actually cuts against a per se ban on polygraph evidence. As the preface of the book states:

"Dr. Kalashnikov [the author] is a polygraph professional. If you go up against him, or someone like him, he'll probably catch you at your game. That's because he knows his work and does it by the book.

"What most people don't realize is that there are a lot of not so professional polygraph examiners out there. It's very possible that you may be tested by someone who is more concerned about the number of tests he will run this week (and his Christmas bonus) than he is about the precision of each individual test.

"Remember, the adage is that you can't beat the polygraph system but you can beat the operator. This book is gleefully dedicated to the idea of a sporting chance." V. Kalashnikov, Beat the Box: The Insider's Guide to Outwitting the Lie Detector (1983) (preface): *id.*, at 9. ([W]hile the system is all but unbeatable, you can surely beat the examiner.'); Thus, "Beat the Box" actually supports the notion that polygraphs are reliable when conducted by a highly trained examiner-like the one in this case. Nonetheless, some research has indicated that people can be trained to use "countermeasures" to fool the polygraph. See, *e.g.*, Honts, Raskin, & Kircher, Mental and Physical Countermeasures Reduce the Accuracy of Polygraph Tests, 79 J. Applied Psychology 252 (1994). This possibility, however, does not justify a *per se* ban. First, research indicates that individuals must receive specific training before they can fool the polygraph (*i.e.*, information alone is not enough). Honts, Hodes, & Raskin, Effects of

The Role of the Jury

It is the function of the jury to make credibility determinations. In my judgment evidence that tends to establish either a consciousness of guilt or a consciousness of innocence may be of assistance to the jury in making such determinations. That also was the opinion of Dean Wigmore:

"Let the accused's whole conduct come in; and whether it tells for consciousness of guilt or for consciousness of innocence, let us take it for what it is worth, remembering that in either case it is open to varying explanations and is not to be emphasized. Let us not deprive an innocent person, falsely accused, of the inference which common sense draws from a consciousness of innocense and its natural manifestations." 2 J. Wigmore, Evidence §293, p. 232 (J. Chadbourn rev. ed. 1979).

There is, of course, some risk that some "juries will give excessive weight to the opinions of a polygrapher, clothed as they are in scientific expertise," *ante*, at 10. In my judgment, however, it is much more likely that juries will be guided by the instructions of the trial judge concerning the credibility of expert as well as lay witnesses. The strong presumption that juries will follow the court's instructions, see, *e.g.*, *Richardson v. Marsh*, 481 U.S. 200, 211 (1987), applies to exculpatory as well as inculpatory evidence. Common sense suggests that the testimony of disinterested third parties that is relevant to the jury's credibility determination will assist rather than impair the jury's deliberations. As with the reliance on the potential unreliability of this type of evidence, the reliance on a fear that the average jury is not able to assess the weight of this testimony reflects a distressing lack of confidence in the intelligence of the average American.[39]

Collateral Litigation

The potential burden of collateral proceedings to determine the examiner's qualifications is a manifestly insufficient justification for a categorical exclusion of expert testimony. Such proceedings are a routine predicate for the admission of any expert testimony, and may always give rise to searching cross-examination. If testimony that is critical to a fair determination of guilt or innocence could be excluded for that reason, the right to a meaningful opportunity to present a defense would be an illusion.

It is incongruous for the party that selected the examiner, the equipment, the testing procedures, and the questions asked of the defendant to complain about the examinee's burden of proving that the test was properly conducted. While there may well be a need for substantial collateral proceedings

Physical Countermeasures on the Physiological Detection of Deception, 70 J. Applied Psychology 177, 185 (1985); see also Honts, Raskin, Kircher, & Hodes, Effects of Spontaneous Countermeasures on the Physiological Detection of Deception, 16 J. Police Science and Administration 91, 93 (1988) (spontaneous countermeasures ineffective). Second, as countermeasures are discovered, it is fair to assume that polygraphers will develop ways to detect these countermeasures. See, *e.g.*, Abrams & Davidson, Counter-Countermeasures in Polygraph Testing, 17 Polygraph 16, 17-19 (1988); Raskin, Honts, & Kircher, The Case for Polygraph Tests, in Faigman 577-578. Of course, in any trial, jurors would be instructed on the possibility of countermeasures and could give this possibility its appropriate weight.

39

Indeed, research indicates that jurors do not "blindly" accept polygraph evidence, but that they instead weigh polygraph evidence along with other evidence. Cavoukian & Heslegrave, The Admissibility of Polygraph Evidence in Court: Some Empirical Findings, 4 Law and Human Behavior 117, 123, 127-128, 130 (1980) (hereinafter Cavoukian & Heslegrave); see also Honts & Perry 366-367. One study found that expert testimony about the limits of the polygraph 'completely eliminated the effect of the polygraph evidence' on the jury. Cavoukian & Heslegrave 128-129 (emphasis added).

when the party objecting to admissibility has a basis for questioning some aspect of the examination, it seems quite obvious that the Government is in no position to challenge the competence of the procedures that it has developed and relied upon in hundreds of thousands of cases.

In all events the concern about the burden of collateral debates about the integrity of a particular examination, or the competence of a particular examiner, provides no support for a categorical rule that requires exclusion even when the test is taken pursuant to a stipulation and even when there has been a stipulation resolving all potential collateral issues. Indeed, in this very case there would have been no need for any collateral proceedings because respondent did not question the qualifications of the expert who examined him, and surely the Government is in no position to argue that one who has successfully completed its carefully developed training program[40] is unqualified. The interest in avoiding burdensome collateral proceedings might support a rule prescribing minimum standards that must be met before any test is admissible,[41] but it surely does not support the blunderbuss at issue.[42]

IV

The Government's concerns would unquestionably support the exclusion of polygraph evidence in particular cases, and may well be sufficient to support a narrower rule designed to respond to specific concerns. In my judgment, however, those concerns are plainly insufficient to support a categorical rule that prohibits the admission of polygraph evidence in all cases, no matter how reliable or probative the evidence may be. Accordingly, I respectfully dissent.

* * * * * *

40

See n. 18, *supra.*

41

See N.M. Rule Evid. §11-707.

42

It has been suggested that if exculpatory polygraph evidence may be adduced by the defendant, the prosecutor should also be allowed to introduce inculpatory test results. That conclusion would not be dictated by a holding that vindicates the defendant's Sixth Amendment right to summon witnesses. Moreover, as noted above, studies indicate that exculpatory polygraphs are more reliable than inculpatory ones.

See n. 35, *supra.* In any event, a concern about possible future legal developments is surely not implicated by the narrow issue presented by the holding of the Court of Military Appeals in this case. Even if it were, I can see nothing fundamentally unfair about permitting the results of a test taken pursuant to stipulation being admitted into evidence to prove consciousness of guilt as well as consciousness of innocence.

United States v. Scheffer: A Review of the Opinion of the United States Supreme Court

Gordon L. Vaughan, Esq., APA General Counsel

Introduction

On March 31, 1998, the United States Supreme Court announced its opinion in *United States v. Scheffer*, ___ U.S. ___, 1998 WL 141151 (1998). In this much anticipated opinion,[1] the Court reversed the United States Court of Appeals for the Armed Services determination that *per se* evidentiary rules excluding polygraph evidence in criminal proceedings violated a defendant's U.S. Constitutional Sixth Amendment right to present a defense.

The direct effect of the Court's decision is to uphold, at least in the face of a Sixth Amendment challenge, a *per se* rule of evidence, imposed by the appropriate rule-making authority, excluding polygraph evidence in criminal proceedings. While this decision is a setback for proponents of the admission of polygraph evidence in criminal cases, the opinion of the Court does not foreclose the admission of polygraph evidence in those jurisdictions where rule-making authorities have not imposed a *per se* rule of exclusion of polygraph evidence. Further, though finding that there is an ongoing debate in the scientific community regarding the validity of polygraph evidence, the Court did not conclude, as has some lower courts, that polygraph evidence was not reliable. Rather, the Court held that, given the ongoing scientific debate, a *per se* rule excluding polygraph evidence is not unconstitutional as it cannot be said that such rule arbitrarily or disproportionately infringes upon a sufficient weighty interest of a criminal defendant.

Case Facts

In *Scheffer*, a general court-martial composed of officer members at March Air Force Base, California, convicted Airman Scheffer of writing bad checks, use of methamphetamine, and being absent without leave. Evidence presented at the court-martial established that Airman Scheffer was an informant for OSI. He reported to OSI that two civilians were dealing in significant quantities of drugs. Airman Scheffer was asked to and provided a voluntary urine sample and submitted to a polygraph examination. The relevant questions posed to Airman Scheffer were: (1) had he used drugs while in the Air Force, (2) had he lied regarding any drug information given to OSI, and (3) had he ever told anyone, other than his parents, that he was working for the OSI. His answer was no to each of these questions. The examiner concluded that no deception was indicated.[2]

After his urinalysis came back positive, Airman Scheffer was charged and court-martialed for, in part, use of methamphetamine. Airman Scheffer offered an unintentional ingestion defense. Although Airman Scheffer

1

Copies of the briefs submitted to the Supreme Court by Petitioner and Respondent and APA's *Amicus Curiae* brief are published in 26 *Polygraph* 130 (1997).

2

Although not set forth in the opinion, the author has been advised that the testing technique used by the examiner was an MGQT.

testified at the court-martial hearing, and his credibility was made a central issue by the prosecution, the court refused Scheffer's attempt to lay a foundation for the admission of his exculpatory polygraph.[3] The basis of the Court's refusal was Military Rule of Evidence 707, which provides:

"Notwithstanding any other provisions of law, the results of a polygraph examination, the opinion of the polygraph examiner, or any reference to an offer to take, failure to take, or taking of a polygraph examination, shall not be admitted into evidence."

The Court's Decision

Three separate opinions make up the Supreme Court's decision in *Scheffer*. The principal opinion was written by Justice Thomas and was fully joined in by Chief Justice Rehnquist, Justice Souter, and Justice Scalia. Justice Kennedy wrote an opinion concurring in part with the principal opinion and concurring in the judgment. Joining Justice Kennedy in this concurring in part opinion was Justice O'Connor, Justice Ginsburg, and Justice Breyer. Justice Stevens wrote a dissenting opinion.

The Principal Opinion

Justice Thomas' principal opinion is divided into six sections: I, II, II(a), II(B), II(C), and II(D). Section I presents the facts and procedural history of the case.

Section II sets out the test to be applied in determining whether an evidentiary rule excluding polygraph evidence in a criminal trial violates a defendant's constitutional rights. It is in Section II that the outcome of the case was fundamentally determined. Obviously the heavier the burden placed on the government to justify a per se evidentiary exclusion of polygraph evidence, the more likely such exclusion would be found to be unconstitutional. The Court adopted a comparatively light burden.

In the briefs and arguments of the parties, there was extensive debate regarding the burden the government must bear in justifying a per se exclusion of polygraph evidence. The case of *Rock v. Arkansas*, 483 U.S. 44 (1987), was central to this debate. In *Rock*, the Court had held that a per se rule excluding all hypnotically-refreshed testimony was unconstitutional. In so doing, the *Rock* Court had held that:

"... restrictions of a defendant's right to testify may not be arbitrary or disproportionate to the purposes they are designed to serve."

Rock at 55-56. In describing whether a *per se* ban of a defendant's hypnotically-refreshed testimony was arbitrary or disproportionate, the Court went on to hold that:

"...[w]holesale inadmissibility of a defendant's testimony is an arbitrary restriction on the right to testify in the absence of clear evidence by the State repudiating the validity of all post-hypnosis recollections." *Rock* at 61.

In *Scheffer*, Justice Thomas, citing *Rock* and other similar authority, found that:

"State and federal rulemakers therefore have broad latitude under the Constitution to establish rules excluding evidence. Such rules do

---

3

For a more detailed description of the facts and the lower court decision, *see* G. Vaughan, *United States v. Scheffer:* The United States Supreme Court Considers Admissibility of Polygraph Evidence, *26 Polygraph* 127 (1997).

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."'[4]

Justice Thomas declined to apply in *Scheffer* the more rigorous requirement of *Rock* that required that the government, to survive a constitutional challenge of a rule of per se inadmissibility of hypnotically-refreshed testimony, to repudiate the validity of all hypnotically-refreshed testimony. Justice Thomas observed that in *Rock*, the effect of excluding a defendant's hypnotically-refreshed testimony was to prevent the defendant from testifying on his own behalf as to relevant facts about the case. He further observed that excluding Airman Scheffer's exculpatory polygraph did not prevent Scheffer from testifying on his own behalf but prevented only the introduction of "expert opinion testimony to bolster his own credibility." Justice Thomas found that, unlike *Rock*, such restriction was not a significant impairment to a defendant's case. Based on this test, then, the government was required only to provide a legitimate governmental interest in promulgating a per se rule excluding all polygraph evidence.

Sections II(A), II(B), and II(C) set out three governmental interests that the principal opinion found to be legitimately served by Rule 707's per se exclusion of polygraph evidence. These were: II(A) - the rule insures the exclusion of unreliable evidence in that there is "no consensus that polygraph evidence is reliable"; II(B) - it preserves "the jury's core function of making credibility determinations"; and II(C) - it avoids "litigation over issues other than the guilt or innocence of the accused."

With regard to the governmental interest in ensuring the exclusion of unreliable evidence, Justice Thomas wrote:

"The contentions of respondent and the dissent notwithstanding, there is simply no consensus that polygraph evidence is reliable. To this day, the scientific community remains extremely polarized about the reliability of polygraph techniques. [Cites omitted.] Some studies have concluded that polygraph tests overall are accurate and reliable. *See, e.g.*, S. Abrams, The Complete Polygraph Handbook 190-191 (1968) (reporting the overall accuracy rate from laboratory studies involving the common "control question technique" polygraph to be "in the rage of 87 percent"). Others have found that polygraph tests assess truthfulness significantly less accurate - the scientific field studies suggest the accuracy rate of the "control question technique" polygraph is "little better than could be obtained by the toss of a coin," that is, 50 percent. *See* Iacono & Lykken, The Scientific Status of Research on Polygraph Techniques: The Case Against Polygraph Tests, in 1 Modern Scientific Evidence, *supra*, §14-5.3, P. 629."

With regard to the governmental interest in preserving the jury's function of making credibility determinations (often referred to in other cases as the "usurpation of jury function" argument), Justice Thomas wrote:

"A fundamental premise of our criminal trial system is that "the jury is the lie detector." [Cites omitted.] Determining the weight and credibility of witness testimony,

---

[4] At the time of the writing of this article publication page numbers for the *Scheffer* opinion were not available.

therefore, has long been held to be the "part of every case [that] belongs to the jury, who are presumed to be fitted for it by their natural intelligence and their practical knowledge of men and the ways of men." [Cites omitted.]

"By its very nature, polygraph evidence may diminish the jury's role in making credibility determinations."

Finally, with regard to the governmental interest in avoiding litigation over issues of guilt or innocence (often referred to in other cases as the "collateral litigation" argument), Justice Thomas wrote:

"Allowing proffers of polygraph evidence would inevitably entail assessments of such issues as whether the test and control questions were appropriate, whether a particular polygraph examiner was qualified and had properly interpreted the physiological responses, and whether other factors such as countermeasures employed by the examinee had distorted the exam results."

In section II(D), Justice Thomas concluded that, on the basis of these legitimate government interests, "Military Rule of Evidence 707 does not unconstitutionally abridge the right to present a defense."

Concurring in Part Opinion

In the concurring in part opinion, Justice Kennedy agreed that there was a continuing disagreement among the experts on the subject of polygraph reliability. On this basis alone, Justice Kennedy concurred with the principal opinion's judgment that Rule 707 was "not so arbitrary or disproportionate that it is unconstitutional." In so doing, however, Justice Kennedy observed that such rules of per se exclusions of polygraph evidence may not be wise. Leaving open the possibility that the Court might consider future cases on this issue, Justice Kennedy wrote:

"I doubt, though, that the rule of per se exclusion is wise, and for 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 or excluding scientific evidence is not a constitutional mandate, *see Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 587 (1993), there is some tension between that rule and our holding today. And, as JUSTICE STEVENS points out, 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."

Additionally, Justice Kennedy rejected as legitimate government interests the arguments that the introduction of polygraph evidence usurped jury responsibilities and introduced inappropriate collateral litigation. Regarding the jury usurpation argument, Justice Kennedy observed that:

"[I]t seems the principal opinion overreaches when it rests its holding on the additional ground that the jury's role in making credibility determination is diminished when it hears polygraph evidence. ... I had thought this tired argument had long since been given its deserved repose."

In beginning his dissenting opinion, Justice Stevens observed that the Court's holding

"Rests on a serious underevaluation of the importance of the citizen's constitutional right to present a defense to a criminal charge and an

unrealistic appraisal of the importance of the governmental interest that undergird the Rule."

With regard to the reliability of the polygraph, Justice Stevens wrote:

"There are a host of studies that place the reliability of polygraph tests at 85% to 90%. [Footnote omitted.] While critics of the polygraph argue that accuracy is much lower, even the studies cited by the critics place polygraph accuracy at 70%. [Footnote omitted.] Moreover, to the extent that the polygraph errs, studies have repeatedly shown that the polygraph is more likely to find innocent people guilty than vice versa. Thus, exculpatory polygraphs - like the one in this case - are likely to be more reliable than inculpatory ones."

"Of course, within the broad category of lie detector evidence, there may be a wide variation in both the validity and the relevance [Footnote omitted.] of particular test results. Questions about the examiner's integrity, independence, choice of questions, or training in the detection of deliberate attempts to provoke misleading physiological responses may justify exclusion of specific evidence. But such questions are properly addressed in adversary proceeding; they fall far short of justifying a blanket exclusion of this type of expert testimony."

With regard to the usurpation of jury function argument, Justice Stevens wrote that the

"fear that the average jury is not able to assess the weight of this testimony reflects a distressing lack of confidence in the intelligence of the average American." [Footnote omitted.]

As to the argument that polygraph evidence injects into the case inappropriate collateral litigation, Justice Stevens observed:

"The potential burden of collateral proceedings to determine the examiner's qualifications is a manifestly insufficient justification for a categorical exclusion of expert testimony."

"It is incongruous for the party that selected the examiner, the equipment, the testing procedures, and the questions asked of the defendant to complain about the examinee's burden of proving that the test was properly conducted. While there may well be a need for substantial collateral proceedings when the party objecting to admissibility has a basis for questioning some aspect of the examination, it seems quite obvious that the Government is in no position to challenge the competence of the procedures that it has developed and relied upon in hundreds of thousands of cases."

Justice Stevens concluded that Rule 707 did indeed violate Airman Scheffer's Sixth Amendment right to present a defense.

Analysis

Although the concurring in part opinion leaves open the possibility that "some later case might present a more compelling case for introduction of polygraph testimony," *Scheffer* leaves little room for a Sixth Amendment attack on evidentiary rules, established by the proper rule-making authority, excluding polygraph evidence in criminal cases. It is important to observe, however, that *Scheffer* leaves open and, in fact, may support the argument that judicially-imposed per se bans on polygraph evidence are not appropriate in the wake of the amendments to the Federal Rules of Evidence and the decision in *Daubert*

v. Merrell Dow Pharmaceuticals, Inc., supra.[5]

As observed by Justice Stevens in his dissenting opinion, Military Rule of Evidence 707 "has no counterpart in either the Federal Rules of Evidence or the Federal Rules of Criminal Procedure." Accordingly, until such time, if ever, that a similar rule of evidence ins imposed on the federal courts, the admissibility of polygraph evidence in federal courts continues to be governed by *Daubert v. Merrell Dow Pharmaceuticals, Inc.* Under *Daubert*, considerable discretion is placed in the trial court to determine whether to admit or exclude scientific evidence. Similarly, states which have adopted rules similar to the federal rules or which follow the test for admitting scientific evidence under *Daubert*, and whose rule-making authority has not imposed a per se ban of polygraph evidence, are not affected by this opinion.

It should be anticipated that opponents of polygraph will now encourage their respective rule-making authorities to impose polygraph exclusion rules similar to Military Rule of Evidence 707. It is at the rule-making level that much of the battle regarding the admissibility of polygraph evidence may now be expected to be fought. In the federal system, rules of evidence and procedure are promulgated by the United States Supreme Court and subject to approval by Congress. *See* 28 U.S.C. §2074. Given the sentiments expressed by a majority of the justices in the concurring in part and dissenting opinions, it is unlikely that the Supreme Court, as presently constituted, would endorse such a rule in the federal system.

It is also important to note that the combined effort of the concurring in part and dissenting opinions appear, finally, to put to rest the jury usurpation and collateral litigation arguments often championed by opponents of polygraph. Presumably, these arguments will carry little weight in future litigation involving polygraph admissibility.

Proponents for the use of polygraph evidence in the military criminal justice system should also remember that Military Rule of Evidence 707 is a presidentially-imposed rule of evidence. It is subject to rescission or modification at the direction of the president. Thus, further scientific evidence and/or the political winds may yet work to convince the present or some future administration to remove the per se ban of polygraph evidence in military court-martial proceedings.

Conclusion

As previously noted, *Scheffer* is a setback for proponents of the admission of polygraph evidence in criminal cases. It is not, however, as reported by many in the media, the death knell for the admission of polygraph evidence in criminal trials. Rather, other than constitutionally upholding rule-making authorities' evidentiary limitations of polygraph evidence in criminal cases, *Scheffer* does not impede and, in some ways, endorses the growing trend of the courts to consider, at the trial court level, the admissibility of polygraph evidence on a case-by-case basis.

While the courts will continue to struggle with issues of polygraph admissibility, much of the battle may now be expected to be fought at the rule-making level. Many jurisdictions

[5] Most per se bans of polygraph have not been imposed by rule but, rather, by judicial decision.

will no doubt, in the wake of *Scheffer*, impose their own versions of Rule 707. However, such rules are more easily changed than court precedent. As such, as the science and profession of polygraphy improve, rules imposing per se restrictions on polygraph evidence will become increasingly difficult to justify.

* * * * * *

Effectiveness of Detection of Deception Examinations Using the Computer Voice Stress Analyzer

Michael J. Janniro, Ph.D. & Victor L. Cestaro, Ph.D.

Abstract

The accuracy of the Computer Voice Stress Analyzer (CVSA) instrument and associated processes for the detection of deception was assessed using a mock theft scenario. One hundred nine subjects were randomly assigned to two groups and given detection of deception examinations using a CVSA instrument. Subjects in one group were programmed deceptive and participated in taking $100 from a metal box located in a scenario room. The non-deceptive group did not participate in the scenarios nor did they have knowledge of the mock theft. Four trained and certified CVSA examiners conducted the examinations using a CVSA technique called the Modified Zone of Comparison test. CVSA test chart evaluators, who had not taken part in the study and who were blind to subject programming, obtained an overall accuracy of 49.8% ($z = -.05$, $p = .96$). Administering CVSA examiners correctly identified 53 of the 109 (48.6%) subjects as either deceptive or non-deceptive ($z = -.21$, $p = .84$). More deceptive subjects were correctly identified by examiners than non-deceptive subjects (32 of 55 *vs.* 21 of 54). However, decisions were not significantly different from chance in either case.

Key words: Computer Voice Stress Analyzer, CVSA, detection of deception, voice stress

---

Acknowledgments:

The authors wish to express sincere gratitude to Detective Gary Noel and Robert Neubauer, and Investigators Terry Kilgore and Bernie Erwin for serving as the examiners in this study. Special appreciation goes to Special Agent Bob Tippett, Florida Department of Law Enforcement, for his technical advice on the Computer Voice Stress Analyzer (CVSA), and for thoroughly reviewing the study protocol. Also, special thanks to Brenda Smith, Charlene Stephens, Kathy Harris, Linda Landy, Johan Harrison-Woodard, and Sarah Tidwell.

This project was funded by the Department of Defense Polygraph Institute as DoDPI95-P-0016. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. For reprints, contact Dr. Victor Cestaro, DoDPI, Bldg. 3195 and 13th Avenue., Ft. McClellan, AL 36205.

In 1971 Dektor Counterintelligence and Security, Inc., (Savannah, Georgia), developed a device for detecting stress, which they called the Psychological Stress Evaluator (PSE). The National Institute for Truth Verification (NITV) *Certified Examiners Course Manual* (1995) states that the PSE detects subaudible micro tremors in the human voice, and that analysis of these stress related tremors has great utility for the detection of deception. Soon afterwards, advertisements in popular magazines, newspapers, and trade journals began comparing the accuracy and utility of the polygraph to voice stress analyzers (NITV, 1990; NITV, 1994; NITV, 1995). Claims have been made in newspaper articles that the CVSA is easier to use and more accurate than the polygraph (NITV, 1990, p. 18).

The PSE has recently been supplanted by an instrument called the Computer Voice Stress Analyzer (CVSA) manufactured by the NITV. Although the theoretical physiological basis of monitoring subaudible micro tremors is unchanged from the PSE, instrument design changes and ease of use are making the CVSA a popular tool. Periodic publications of the NITV's *Journal of Continuing Education* (*e.g.*, 1990) include several newspaper articles pronouncing the CVSA's effectiveness and acceptance by many police departments. Most testimonials cited in NITV's journals, regarding the efficacy of the CVSA, stress its utility in obtaining admissions and confessions. However, the manufacturer does not provide evidence of controlled laboratory studies that would support the high accuracy rates (97-100%) routinely claimed (NITV training registration form). Furthermore, no explanations are provided for how these accuracy rates are determined.

The effectiveness of the polygraph, has been the subject of a number of controlled scientific studies over the years (Ansley & Garwood, 1984; U.S. Congress, Office of Technology Assessment, 1983). According to Horvath (1982), many well-controlled laboratory studies and field studies support the polygraph and its associated procedures and processes. Horvath argues that even the most severe critics agree that the findings show an accuracy that justifies the use of polygraph testing under certain conditions. However, the CVSA and its testing procedures and processes have not been subjected to the same vigorous scientific evaluation as the polygraph and its procedures.

In 1993, the Inspector General of the Florida Department of Law Enforcement (FDLE, Tallahassee, FL) released a position paper recommending that FDLE prohibit the use of voice stress analysis as an investigative tool because of the lack of scientific evidence supporting its validity. Since the CVSA records physiological data from a response system (the voice) that the current polygraph is incapable of recording, it is possible that the combination of instruments and processes (polygraph and the CVSA) could increase the accuracy and reliability of the detection of deception. The purpose of this study was to evaluate the CVSA and its associated procedures to determine its efficacy in detecting deception.

Method

Subjects

One hundred nine subjects were recruited from a local contract agency and randomly assigned to deceptive and non-deceptive groups. Volunteers were male or female, literate, between the ages of 19 and 65 years, and had a minimum of a high school diploma or GED. Each subject

was determined to be in good health and slept at least six hours the evening before testing.

Apparatus

Four Computer Voice Stress Analyzers (National Institute for Truth Verification, W. Palm Beach, FL) were used to record and display voice response data on paper charts. Lapel microphones (Radio Shack, Fort Worth TX, Model 33-3003) were used for supplying subjects' verbal responses to the input jacks of the CVSAs.

Examiners

Four CVSA examiners, trained and certified by NITV, conducted the examinations. The examiners were blind to subject programming. The CVSA tests were also independently blind scored by three trained and certified CVSA examiners, hereinafter referred to as evaluators.

Procedures

Upon arrival at the Department of Defense Polygraph Institute (DoDPI) testing site, each participant was escorted by a research team assistance to the DoDPI library and asked to read a brief description of the research project. Subjects were programmed in groups of four, two groups in the morning and afternoon. Individuals willing to participate in the study were asked to read and sign a volunteer agreement affidavit. A brief biographical/medical questionnaire was completed to ensure that each participant was in good health and not taking medication that could interfere with examination results. Research team assistants than began programming deceptive and non-deceptive subjects according to the scenario instructions. All subjects were then given their appropriate written instructions. Random assignments of subjects to groups were made before the actual experiment. Nearly half of the subjects (n = 55) were assigned to the deceptive group and participated in taking $100 from a metal box located in a scenario room. The remaining subjects (n = 54) did not participate in the scenarios nor did they have knowledge of the mock theft.

Deceptive subjects were instructed to proceed to the scenario room and to remove the $100 bill from an open metal box located on a table in the scenario room. Each deceptive subject was told to hide the $100 bill on their person. Additionally, they were instructed to lie to the examiner about taking $100 from the metal box and having the money on their person. Next to the metal box was a 3" x 5" card with each deceptive subject's examiner room assignment. Non-deceptive subjects also entered the scenario room and picked up a 3" x 5" card with their examiner room assignment. However, the metal box containing the $100 bill as removed before non-deceptive subjects entered the scenario room. They had no knowledge of the theft, and were instructed to answer questions truthfully during the examination.

CVSA examiners conducted the pre-test interview as described in the NITV *Certified Examiners Course Manual* (1995). The relevant questions used were the same for all subjects (Appendix A). The control and irrelevant questions used were developed by each examiner, based on the rules of test question formulation taught in the NITV Certified Examiners Course. All test questions were reviewed with the subject before testing began. The lapel microphone was placed on the subject and the CVSA instrument was calibrated for the subject's voice level. The examination proceeded using the accepted CVSA format for

the Modified Zone of Comparison test and appropriate test questions.

The CVSA examiners conducted three examinations. The chart from the first examination was not evaluated, in accordance with NITV scoring procedures (*Certified Examiners Course Manual*, 1995). The second and third charts were numerically scored, and categorized as Deception Indicated (DI) or No Deception Indicated (NDI). In addition, all examinations were recorded on video/audio VHS tape for off-line analysis. When the examination was completed, the subject was escorted back to the briefing room for subject debriefing.

Scoring

Before data reduction and analysis, the original examiners independently evaluated each graphic recording. Based on their scoring they were asked to make a diagnosis of either DI or NDI. CVSA procedures do not allow for inconclusive determinations. The examiners' scores and decisions were not written on the charts. The decision for each subject was written by each examiner on a scoring sheet maintained by the examiner. All charts were marked only with the date of the examination and subject number. Charts blind scored by the three evaluators had all subject numbers removed, and were randomly coded.

Data analysis

The dependent measure for accuracy was the number of correct decisions made regarding deception and non-deception. Interrater agreement was determined by comparing the decision made for each subject among the evaluators, irrespective of the accuracy of the decision. Analyzes included a 2 x 2 chi-square analyzes of programming *vs.* decision, and a test of the significance of proportions of DI and NDI decisions when compared to change expectancy (0.50). An in-house program using common signal detection theory (SDT) procedures was used to assess instrument sensitivity.

Results

Evaluators made correct decisions on 163 of 327 charts (109 subjects x 3 evaluators), obtaining an overall accuracy of 49.8% ($z = -.05$, $p. = .96$), with a range of 45.9% to 54.1%. Their accuracy ranged from 54.5% to 63.6% for DI decisions, and 35.2% to 53.7% for NDI decisions. Administering examiners did slightly worse, achieving an overall accuracy of 48.6% ($z = .21$, $p = .84$), with an accuracy range of 33.3% to 55.6%. Their DI decision accuracy ranged from 38.5% to 66.7%, and their NDI decision accuracy range was 13.3% to 66.7%. No examiner obtained a combined (DI and NDI) accuracy rate significantly different from chance, nor were the results of chi-square analyses significant. Application of SDT to the data showed that overall instrument sensitivity was low. The noise and signal+noise distributions were completely overlapped, with the criterion line (beta) positioned near th means of the overlapped distributions, indicating nearly equal probability for DI or NDI decisions ($d'=0$, beta = 1.01). Interrater reliability (mean proportion of agreement) for decisions rendered was conducted by three blind score evaluators. These evaluators obtained a correct unanimous agreement rate of 26%, and a correct majority (2 of 3) agreement rate of 46%.

Discussion

As shown in a previous study (Cestaro, 1996a), the sensitivity of the CVSA is low when used in a low or no-stress situation, such as that encountered during a typical laboratory study. The CVSA manufacturer claims that stress

related to deception can be detected reliably by the instrument, and that stressful and non-stressful resposnes can be differentiated by trained operators. However, in this study, evaluators and administering examiners were not able to distinguish between deception and non-deception at rates better than chance levels of accuracy (50%). Prior to conducting the study, a power analysis of the proportion test for accuracy indicated that with N =100 (50 per group collapsed across programming [guilty, innocent] and an expected effect size of 0.20, power = .99 (p = .05). This means that, under the test conditions used in this study, there is a .99 probability that an effect of .20 greater than chance would have been detected, had one existed.

While every attempt was made to emulate the subject programming procedures reported in other studies, it is possible that the procedures used did not elicit physiologic responses during deception. While, in our opinion, it is unlikely that the low accuracy rates obtained are due to problems with the mock crime scenario, it is a possibility.

The test procedures incorporated in the study were the same as those used in field examinations, and all seven examiners (administering and evaluating) were trained and certified by the equipment manufacturer. All examiners had practical field experience in the pre-test, in-test, test analysis, and post-test phases of CVSA examination administration, and the administering examiners were permitted to conduct the examinations as learned in certification training conducted by the NITV. Examinations were monitored by a CVSA instructor from the NITV. The statistically significant decision concurrence rate, as shown by the results of the interrater reliability tests, seems to provide some level of confidence that the scoring methods employed among examiners were consistent. However, from a practical viewpoint, examiners obtained majority decision agreement on less than half of the subjects, and unanimous agreement on about one quarter of the subjects tested. The lack of instrument sensitivity to the measure(s) of interest impacted on the ability of examiners and evaluators to accurately and consistently discriminate between truthful and deceptive responses when assessing the subject's test charts.

In summary, although there is evidence to support the basic electrical theory of operation of the CVSA (Cestaro, 1996b), the instrument failed to function in a manner that would allow examiners to discriminate between truthful and deceptive responses from test subjects. Further research should examine the effects of increased levels of stress on subjects' responses to determine if there is a correlation between stress levels and instrument display characteristics. Although the CVSA instrument is purported to detect stress in human speech, there is still no unambiguous evidence to support that claim.

References

Ansley, N. & Garwood; M. (1984). The accuracy and utility of polygraph testing. *Polygraph, 13,* 3-131.

Certified Examiners Course Manual (1995) (Available from the National Institute for Truth Verification, West Palm Beach, FL.)

Cestaro, V.L. (1996a). A comparison between decision accuracy rates obtained using the polygraph instrument and the computer voice stress analyzer (CVSA) in the absence of jeopardy. *Polygraph, 25,* 117-127.

Cestaro, V.L. (1996b). A test of the computer voice stress analyzer (CVSA) theory of operation. *Polygraph, 25,* 101-116.

Florida Department of Law Enforcement. (1993, October). *Review of literature regarding voice stress analysis.* A report compiled by the Office of Inspector General FDLE, Tallahassee, FL.

Horvath, F. (1982). Detecting deception: The problem and reality of voice stress analysis. *Journal of Forensic Sciences, 27,* 340-351.

National Institute for Truth Verification. (1990). *Journal of Continuing Education (Vol. 10, No. 6).* West Palm Beach, FL: Author.

National Institute for Truth Verification.(1994). *Journal of Continuing Education (Vol. 12, No. 1).* West Palm Beach, FL: Author.

National Institute for Truth Verification. (1995). *Journal of Continuing Education (Vol. 12, No. 3).* West Palm Beach, FL: Author.

U.S. Congress, Office of Technology Assessment. (1983). *Scientific validity of polygraph testing* (OTA-TM-H-15). Washington, D.C.: U.S. Government Printing Office.

Appendix A

CVSA Relevant Test Questions (MZOC)

IR 1.

C 2.

IR 3.

R 4. Do you know who took that $100 bill from that metal box?

IR 5.

R 6. Did you take that $100 bill from that metal box?

IR 7.

C 8.

IR 9.

R 10. Do you suspect anyone of taking that $100 bill from the metal box?

IR 11.

R 12. Do you know who took that $100 from that metal box?

IR 13.

R 14. Did you take that $100 bill from that metal box?

IR 15.

* * * * * *

Errata

In the recent Janniro & Cestaro article (volume 27, number 1) the last sentence of the Data Analysis section should have included the following:

"Scoring reliability (in the form of interrater agreement) was assessed by a multiple rater kappa statistic (Fleiss, 1981)."

The last two sentences of the Results section should have been:

"Interrater reliability for all decisions rendered by the evaluators was high (kappa = .33, SE = .055, $p < .0001$). These evaluators obtained a correct unanimous agreement rate of 26%, and a correct majority (2 or 3) agreement rate of 46%."

We regret the error.

* * * * * *

The Validity Of The Modified General Question Test (MGQT)

Norman Ansley

Abstract

The polygraph technique called the Modified General Question Test (MGQT) has been widely used for several years. The present paper is a review of the relevant literature pertaining to this testing format, beginning with its historical development. The scientific literature was surveyed to provide an estimate of the MGQT's validity. Studies under consideration were divided by type: independent evaluation of confirmed field cases, examiner decisions for simulated MGQT formats, and independent evaluation of simulated MGQT tests. All studies found that detection of deception exceeded chance, some with criterion-related validity at 100%, but there was a substantial range of accuracies as a function of study conditions.

Key words: comparison question test, control question test, field study, laboratory study, MGQT, Modified General Question Technique, validity.

The Modified General Question Test draws its name from the Army C.I.D., which for a time in the 1950s used only a relevant-irrelevant test called the General Question Test (GQT) and the Peak of Tension Test (POT). Introduction of the Reid Control Question Test format into use by the Army C.I.D. first produced the title Reid Control Questioning Technique (Modified), or R.C.Q.T. Subsequently, the Army C.I.D. changed the title to the Modified General Question Test, or M.G.Q.T. Since then, the Reid and M.G.Q.T. tests have been changed in pretest, types of supplemental tests, and in question formulation. For example, the Army went to the Backster exclusive comparison question wording while the Reid format employs inclusive comparison question wording. There are other changes, enough to consider Reid and MGQT separate techniques.

Although MGQT is used for screening, there have been neither field nor simulated studies of its validity in screening. One field study of its utility in screening found it similar in results to the relevant-irrelevant test format (Weaver & Garwood, 1985). There is

The author is a Life Member of the American Polygraph Association and president of Forensic Research, Inc. Comments on this paper or requests for reprints should be sent to the author at 35 Cedar road, Severna Park, Maryland 21146.

one study of field validity (Putnam 1994), four studies of simulated MGQT for validity (Honts & Barland, 1990; Jones & Salter, 1989; Palmatier 1991, Podlesny & Truslow, 1993); and two studies of independent evaluations of simulated MGQT charts (Horvath 1988, Palmatier 1991). There is one study of confirmed field MGQT charts (Buckley & Senese, 1991).

Field Validity of MGQT

Putnam (1994) reviewed all of his cases conducted at the Washoe County Sheriff's Office, Nevada from 1 January 1979 to 1 September 1982 to see which cases could be verified as an accurate or erroneous decision. The only ground for verification was confession of the examinee or confession of another person. There were 552 cases identified during that period. This experienced examiner used MGQT in 121 (22%) and Zone Comparison in 431 (78%) of the cases. Of those 552 cases he was able to verify the outcome of 285 cases (52%), and his data does not separate number of cases or accuracy by technique, MGQT or Zone Comparison. Inconclusives omitted, he made decisions in 65 verified truthful cases and 220 verified deceptive cases, a total of 285 cases. He was correct in 62 of 65 truthful (95%) and 219 of 220 deceptive (99%), totaling 281 of 285 (99%). [Extrapolating the proportions, there were about 14 truthful and 61 deceptive MGQT cases.]

Independent Evaluation of Field MGQT Charts (Table 1)

Buckley & Senese (1991) were interested in the effect of examinee race and gender on the accuracy of chart evaluation. They selected from the files of their office 40 sets of confirmed charts, 20 Caucasian and 20 African-American, with ten sets of truthful and ten sets of deceptive charts in each group. Nine experienced examiners, blind to all details, independently evaluated each set of charts (n. 360). They made 335 decisions, less 25 inconclusive decisions (7%), and were correct in 302 (90%). Of the NDI charts, they were correct in 143 of 163 (88%). Of the DI charts, they were correct in 159 of 172 (92%). There was no statistically significant difference in the accuracy of chart interpretation by race or gender. The charts were in MGQT format.

Simulations of MGQT (Table 2)

Honts & Barland (1990) conducted a laboratory study of the MGQT using 13 experienced examiners and field instruments. Subjects were 88 soldiers who were randomly assigned to truthful and deceptive roles, half in each group, with the deceptive lying about their participation in a mock theft. The decisions were based on a numerical analyses using the standard 7-point scale with +/- 5 cut-off for inconclusives. The examiners were correct in 18 of 29 decisions with truthful subjects (62%), with 15 inconclusives; correct in 38 of 42 decisions with deceptive subjects (90%), with 2 inconclusives. Overall correct in 56 of 71 decisions (79%), with 17 inconclusives.

Jones & Salter (1989) were interested in the validity of the Modified General Question Test (MGQT) and the Peak of Tension Test (POT) in detecting guilty persons from innocent persons in a mock crime experiment. There were nine volunteers ranging in age from 14 to 52 who were given peak of tension tests, followed immediately by Modified General Question Tests. Three of the nine were guilty of a theft: One stole a hundred dollar bill, one stole a Kruggerand coin, and one stole a 9mm. automatic pistol. The POT sequences of five items included these stolen items.

The examinations were conducted by an experienced examiner with a field instrument, who was blind to the roles of the subjects. Three charts were conducted in each test method, and scored with conventional methods. With the MGQT the examiner was correct in 5 of 5 truthful decisions (100%), with one inconclusive; correct in 3 of 3 deceptive decisions (100%), with no inconclusives. Overall, with the MGQT, he was correct in 8 of 8 decisions (100%), with one inconclusive. With the POT, the examiner was correct in 5 of 5 truthful decisions (100%), with one inconclusive; correct in 3 of 3 deceptive decisions (100%) with no inconclusives. Overall, with POT, he was correct in 8 of 8 decisions (100%), with one inconclusive. The same person was inconclusive in both test formats.

Podlesny & McGehee (1987) and Podlesny & Truslow (1993) [same report] evaluated the ability of the MGQT to distinguish between deception and truth, and within the deceptive subjects the ability to distinguish between three roles: confidants, accomplices, and thieves. All subjects, recruited from an urban population, were offered $100 extra if they passed the examination, regardless of role. Tests were given by a trained examiner on laboratory equipment recording the three basic physiological measures, and other data. There were 96 subjects, 24 innocent, 24 deceptives in the role of thieves, 24 deceptives as accomplices, and 24 deceptives as confidants of the thieves. The examiner, using numerical analysis, was correct in 16 of 22 truthful decisions (73%), with 2 inconclusives; correct in 61 of 66 deceptive decisions (92%), with 6 inconclusives; overall correct in 77 of 88 decisions (87%), with 8 inconclusives. The ability of the test to discriminate precise roles among the deceptive was above chance but poor. The authors considered inconclusive decisions to be errors because those decisions did not help discriminate roles, a consideration that is uncommon in polygraph statistical presentations. With chance at 33% for those deceptive, inconclusives included, the thieves were identified correctly in 15 of 24 decisions (63%), accomplices at 6 of 24 decisions (25%), and confidants at 12 of 24 decisions (50%), and for all decisions of deceptive roles, 33 of 72 (46%).

Independent Chart Evaluations of Simulated MGQT Tests (Table 3)

Horvath (1988) was interested in the difference in accuracy of MGQT test sequences with inclusive and with exclusive comparison questions (known as "control questions" in previous literature.) The original MGQT test sequence was the Reid Control Question Test which used inclusive probable lie comparison questions, questions that included the offense at issue. MGQT test formats developed later by organizations in the Department of Defense used exclusive probable lie comparison questions, questions that avoided the offense at issue by dates or locations. Horvath's simulated MGQT tests, 20 of each type, were evaluated independently by the experienced examiners. Excluding one inconclusive decision in each set (2.5%), the exclusive comparison question set evaluations were correct in 28 of 39 (72%). Of the exclusive NDI, correct in 12 of 19 (63%); of the exclusive DI, correct in 16 of 20 (80%). The inclusive comparison question set evaluations were correct in 34 of 39 (87%). Of the inclusive NDI, correct in 16 of 19 (84%); of the inclusive DI, correct in 18 of 20 (90%). (Horvath also conducted 20 sets of charts in which he replaced the comparison questions with irrelevant questions creating an unusual Relevant-Irrelevant test format. Applying MGQT numerical scoring the results were a disastrous

100% false positive rate, and 10% false negative rate. Using visual inspection of the unusual RI charts, the results were not better than chance.)

Palmatier (1991) was also interested in the difference in accuracy of MGQT set sequence with inclusive and exclusive comparison questions. He did not provide numbers for the DI and NDI status; but the examiner decisions for the exclusive comparison question tests were correct in 33 of 43 (77%), excluding 17 (29%) inconclusive decisions. The examiner decisions for the inclusive comparison question tests were correct in 37 of 49 (76%), excluding 11 (18%) inconclusive decisions (Table 2). The independent evaluation of the charts (Table 3) resulted in correct decisions for exclusive comparison question sets in 27 of 44 (61%), excluding 16 (26%) inconclusive decisions. For the inclusive sets, the evaluation was correct in 46 of 55 (84%), excluding 5 (8%) inconclusive decisions. Palmatier was also interested in the difference in accuracy of the MGQT and Zone formats and found them comparable when used in the same context. Also, there was no significant difference by gender of the examinees.

Table 1

Independent evaluation of Confirmed Field MGQT Charts

	NDI			DI			Total			
Buckley & Senese 1991	no.	no. correct	%	no.	no. correct	%	no.	no. correct	%	MGQT (inclusive/ Reid)
	---	---	---	---	---	---	---	---	---	
	163	143	88%	172	159	92%	335	302	90%	

Table 2

SIMULATED MGQT FORMATS, EXAMINERS DECISIONS

Author & Date	NDI no.	NDI no. correct	NDI %	DI no.	DI no. correct	DI %	TOTAL no.	TOTAL no. correct	TOTAL %	Technique
Honts & Barland, 1990	29	18	62%	42	38	90%	71	56	79%	MGQT
Jones & Salter, 1989	5	5	100%	3	3	100%	8	8	100%	MGQT
Palmatier, 1991	*	*	*	*	*	*	49	37	76%	MGQT(inclusive comparison questions)
	*	*	*	*	*	*	43	33	77%	MGQT (exclusive comparison questions)
Podlesny & Truslow, 1993	22	16	73%	66	61	92%	88	77	87%	MGQT
TOTALS	56	39	70%	111	102	92%	259	211	81%	MGQT

* not available

TABLE 3

INDEPENDENT CHART EVALUATION OF SIMULATED MGQT TESTS

	NDI			DI			TOTAL			Format
	no.	no. correct	%	no.	no. correct	%	no.	no. correct	%	
Horvath 1988	19	12	63%	20	16	80%	39	28	72%	MGQT (exclusive)
	19	16	84%	20	18	90%	39	34	87%	MGQT (inclusive)
Palmatier 1991	*			*			44	27	61%	MGQT (exclusive)
	*			*			55	46	84%	MGQT (inclusive)
Totals:	38	28	74%	40	34	85%	177	135	76%	

* Palmatier did not give numbers for NDI and DI.

APPENDIX

TEST FORMATS

R.C.Q.T. FORMAT
U.S. ARMY C.I.D.
(date unknown)

1. Irrelevant

2. Irrelevant

3. Secondary relevant (Did you ...)

4. Irrelevant

5. Primary relevant (Did you ...)

6. Comparison question

7. Irrelevant

8. Evidence connecting relevant (DYK where/what/how ...)

9. Relevant question (DYK or were you present ...)

9a. Alternate Relevant (DY participate in any way ...)

9b. Alternate Relevant (Are you deliberately withholding ...)

9c. Alternate Relevant (Concerning medication, are you withholding information about pills or medicine you have taken in the past ____ hours?*

10. Comparison question

* Taken by Army from an early Backster Zone Comparison test format.

MODIFIED GENERAL QUESTION TECHNIQUE (MGQT)

Federal Bureau of Investigation, 1985

1. Irrelevant
2. Irrelevant
3. Relevant (Did you participate ...)
4. Irrelevant
5. Relevant (Did you ...)
6. Comparison question
7. Irrelevant
8. Evidence connecting relevant (Is that you in the photograph?)
9. Relevant (Are you lying to me about anything ...)
10. Comparison question

Mixed series for third chart is: 4-1-9-6-2-3-10-5-6-8-10.

MODIFIED GENERAL QUESTION TECHNIQUE SUMMARY (MGQT TEST)

DoD Polygraph Institute, 1989

1. Irrelevant
2. Irrelevant
3. Relevant (plan, help, participate)
4. Irrelevant
5. Relevant (Did you ...)
6. Comparison question
7. Irrelevant
8. Evidence connecting relevant
9. Relevant (Do you know who, knowledge ...)
10. Comparison question

Mixed series for third chart: 4-1-5-6-3-10-9-6-8-10.

REID CONTROL QUESTION TEST (MGQT FORMAT)

Reid & Inbau 1977

1. Irrelevant
2. Irrelevant
3. Relevant
4. Irrelevant
5. Relevant
6. Comparison question
7. Irrelevant
8. Relevant
9. Relevant
10. Comparison question

Mixed Series: 4-1-9-5-6-2-3-10-7-5-10 (8 dropped, 10 asked twice)

References

Barland, G. H. (1990). *Analysis of DPI Studies.* Fort McClellan, Alabama: DoD Polygraph Institute.

Buckley, J. P. & Senese, L. C. (1991). The influence of race and gender on blind polygraph chart analyses. *Polygraph,* 20 (4) 237-258.

Honts, C. H. & Barland, G. H. (1990). *A laboratory study of the MGQT: An executive summary.* Fort McClellan, Alabama: DoD Polygraph Institute.

Horvath, F. (1988). The utility of control questions and the effects of two control question types in field polygraph techniques. *Journal of Police Science and Administration,* 16 (3) 198-209.

Jones, H. E. & Salter, S. (1989). Polygraph accuracy: An analog study. *Polygraph,* 18 (2), 69-74.

Minor, P. (1985). Modified General Question Technique (MGQT), Summary Sheet [FBI format]. Paper delivered at the 20th annual seminar of the American Polygraph Association, Reno, Nevada.

Palmatier, J. J. (1991). Analysis of two variations of control question polygraph testing utilizing exclusive and nonexclusive controls. Master of Science Thesis, Michigan State University.

Podlesny, J.A. & McGehee, C.M.* (1987). Investigation detection of deception: Role discrimination with a general question technique. SPR Abstract in *Psychophysiology,* 24 (5), 605-606.

Podlesny, J.A. & Truslow, C.M.* (1993). Validity of an expanded issue (modified general question) polygraph technique in a simulated distributed-crime-role context. *Journal of Applied Psychology,* 78(5), 788-797.

Putnam, R. L. (1994). Field accuracy of polygraph in the law enforcement environment. *Polygraph,* 23 (3), 260.

United States Army Military Police School (n.d.). Form: R.C.Q.T. - Reid Control Questioning Technique (Modified).

United States Army Military Police School (1976). Lesson Title: General Question Test (GQT) Construction.

United States Department of Defense Polygraph Institute (1989). Modified General Question Technique Summary.

Weaver, R. S. & Garwood, M. (1985). Comparison of Relevant-Irrelevant and Modified General Question Technique structures in a split counterintelligence-suitability phase polygraph examination. *Polygraph,* 14 (2), 97-107.

* Miss McGehee became Mrs. Truslow.

Polygraph Validity With Urinalysis Results As Ground Truth

Paul L. Mason

Abstract

Using live US Army polygraph cases conducted following positive urinalysis for marijuana and cocaine, this paper compares the polygraph decisions against the chemical drug tests. Assuming that the urinalysis was correct, there was a 1% false negative rate for the polygraph. It was not possible to disengage the effects of the testing examiners' knowledge of the urinalysis results, however an independent quality control process supported the decisions. The potential and limitations of generalization is discussed.

Key words: Exculpatory examination, PDD, polygraph, urinalysis, validity, zone comparison technique

One of the problems of ascertaining the accuracy of field polygraph examinations is establishing ground truth. This is usually accomplished by comparing the examiner's immediate post-test score with subsequent events, such as confession of the examinee or another person. The use of urinalysis examination results is one of those unusual situations where dependable evidence exists before the psychophysiological detection of deception (PDD) test is administered.

In the United States Department of Defense, a member of the armed forces who is to be disciplined for use of marihuana or cocaine detected during random drug testing may request an exculpatory PDD test. Thus, those who fail urinalysis testing and ask for a PDD test provide a pool of real examinees in which it is highly probable that they will be deceptive during a subsequent examination.

The author is a member of the American Polygraph Association, and a manager in a Federal polygraph program. The views expressed in this article do not reflect the official policy or position of the Department of Defense or the U.S. Government. The pilot findings were initially published by the U.S. Army Crime Records Center. Material from this report has also appeared in the April, 1993 issue of the *American Association of Police Polygraphists* newsletter under the title "Association between positive urinalysis drug test and exculpatory polygraph emanations." Correspondence with the author should be addressed to 6934E Ellingham Circle, Alexandria, Virginia 22315.

Urinalysis procedures in the Department of Defense use a combination of radioimmunoassay (RIA) and gas chromatography/mass spectrometry (GS/MS). The procedures used are considered medically sound and legally defensible if several steps are followed: a. witnessed urine collection, b. maintenance of external and internal chain-of-custody documents, c. records of procedures and worksheets regarding instrument operation, maintenance, and quality control, and records of personnel training and experience (*i.e.*, good laboratory practice), d. participation in accreditation and proficiency testing programs, e. reporting procedures and recall, and f. storage of all specimens, whether they gave negative or positive results (Hoyt, *et al.* 1987).

A disadvantage of using urinalysis test results as ground truth is that the examiners who conduct these exculpatory examinations know the examinee has tested positive for drugs and the probability of deception is very high.

Pilot Study

To test the concept and procedures the author conducted a pilot study of 111 examinees, all U. S. Army personnel who had tested positive for marihuana or cocaine or both. All had requested exculpatory examinations, and all were tested by the author. Of the 111, 110 were scored DI (Deception Indicated), one NDI (No Deception Indicated), and none INC (Inconclusive) or NO (No Opinion). Further investigation of the NDI case disclosed an error in urinalysis procedures (Mason, 1991)

Procedures

Examinees

There were 2,023 U.S. Army military persons who requested and were administered exculpatory PDD examinations because they had tested positive for marihuana or cocaine, or both, during random urinalysis examinations.

Examiners

All of the PDD examinations were conducted by one of the 40 U.S. Army Criminal Investigation Command (USACIDC) forensic psychophysiologists. All of the examiners were graduates of the Department of Defense Polygraph Institute Basic Course, or the Army predecessor course, and all certified under DoD and Army regulations.

Quality Control

All the exculpatory PDD examination charts and reports were sent to the U.S. Army Crime Records Center for quality control review. That consisted of a blind review and scoring of the charts by a senior examiner. Scoring systems used by the reviewer was the same as that used by the examiner, with rules appropriate to the testing format.

Instruments

All the examinations were conducted on standard analog field instruments which simultaneously record cardiovascular (heart rate and blood volume), electrodermal (resistance or conductance), and respiratory (thoracic and abdominal) activity.

Test Formats

Test formats were those approved by USACIDC, and most of the examinations were conducted with either the Zone Comparison Test or

the U.S. Army Modified General Question Test. Probable lie, exclusive time-barred comparison questions were used for comparison. All examinees were advised of their legal rights and all signed a polygraph examination consent form prior to the examination. If the examinee was represented by a defense attorney, and the results were DI, a post-test interrogation was not conducted. All test questions were reviewed with the examinee prior to the conduct of the intest phase of the examination. Relevant questions related to the use of illegal drugs indicated by the urinalysis report.

Results

The results reported here are those of the examiners, and based on their scoring of the charts. Of the 2,023 examined, 1,896 were scored as Deception Indicated (DI), 24 as No Deception Indicated (NDI) 7, and 96 as No Opinion/Inconclusive (NO). The distribution is 94% DI, 1% NDI, and 5% NO/INC. See table 1. Excluding the No Opinion/Inconclusive decisions, the distribution is 99% DI and 1% NDI. See Table 2.

Table 1.

Distribution of Decisions

	DI	NDI	NO/INC	TOTAL
NUMBER:	1,896	24	103	2,023
PERCENTAGE:	94%	1%	5%	100%

Table 2

Distribution of Decisions Without NO and INC

	DI	NDI	TOTAL
NUMBER:	1,896	24	1,920
PERCENTAGE:	99%	1%	100%

There was no information gathered on quality control review, confirmation by confession, or the results of investigation of NDI cases.

If we consider the urinalysis results as ground truth, we have a false negative rate of one percent, and a true positive rate of 99 percent.

Discussion

The advantage of this research methodology is that an imperfect but highly reliable test, independently establishes the use of marihuana or cocaine usage which is denied in the exculpatory PDD testing. If we consider all the NDI decisions to be errors, we have an estimated field validity rate for deceptive cases that is very high. The research method here has a significant drawback: the examiners were aware at the time of testing of the probability of deception, as all of the examinees in the program have failed urinalysis examinations. Critics will suggest that this knowledge may influence the conduct or scoring of the PDD examination. This criticism is mitigated by the quality control procedures and the accompanying blind chart analysis. Also, these examinations are only a part of an examiner's case load. While not included in this study, an independent analysis of these charts by a team of experienced examiners would be a useful supplement. Unlike the quality control reviewers, the research reviewers would not know the issues, questions, case facts, and other data. They would know only the test format. The high validity rate for these exculpatory PDD tests must be tempered by the fact that urinalysis testing, used for ground truth, is not perfect. However, this methodology offers an alternative to the follow-up studies that have been used to establish field validity.

REFERENCES

Hoyt, D. W., Finnigan, R. E., Nee, T., Shultz, T. F. & Butler, T. J. (1987). Drug testing in the workplace - are methods legally defensible. Journal of the American Medical Association 258, (4) 505-509.

Mason, P. L. (1991) Association between positive urinalysis drug testing and exculpatory polygraph examinations. U.S. Army Crime Records Center.

* * * * * *

Call For Papers

The journal *Polygraph* is issuing a call for papers on the chart evaluation methods. Papers can be research, historical, instructional, experimental, or theoretical in nature. In addition to scoring rules, it is incumbent upon all authors to identify important characteristics of the data to be scored, *i.e.*, test formats, physiological features, multiple-issue or single-issue tests, number of charts, etc. Manuscripts regarding automated scoring system will not be included in this call, but are always welcome and will be subject to the standard submission process. Send manuscripts to: Editor, American Polygraph Association, P.O. Box 4085, Anniston, AL 36204. For style guidance, see Instructions to Authors, found in this edition of *Polygraph*, and APA's *Terminology Reference* (1997).

* * * * * *

Psychophysiological Detection of Deception Considerations in Background Investigations

Joseph E. Phipps

Abstract

Background investigations have changed little over the years. If one wants to know about a person's character, habits, or his past behavior, one has to go out and interview persons that know them best. Over the years technology and computer data banks have allowed us to collect much more information with less time and cost. But over time technological improvements have developed other tools that must be considered if we are to continue to do a better job. The use of polygraph in the detection of deception is one of these tools. This paper assesses the background investigative process employed by the federal government and review studies that indicate the use of polygraph techniques may improve the process.

Keywords: Background investigations, detection of deception, polygraph, screening.

Discussion

There are many regulations within the federal government that stipulate and set standards for a person to meet in order to qualify for a position or access to classified information. One of these standards for sensitive positions and access is the completion of a Background Investigation (BI). These regulations stipulate what types of record checks will be conducted, who will be interviewed and what types of derogatory information would disqualify an individual. What is in short supply in the literature is research addressing just how accurate the background investigation is in uncovering adjudicable information.

The author is a CW4 in the U.S. Army, and an instructor with the Department of Defense Polygraph Institute. All views expressed are the author's, and do not represent the position or policies of the Department of Defense Polygraph Institute or the U.S. Government. Requests for reprints should be addressed to the author at DoDPI, Building 3195, Ft. McClellan, AL 36205.

Richards J. Heuer, Jr. wrote a series of articles for the U.S. Government providing background information for personnel security. *Polygraph* (1995) published one of his articles on Drug Use and Abuse and in it Heuer talks about the analysis of 7,232 BIs adjudicated by 14 different federal agencies. Drug use was the most important issue in 59% of the issue cases at agencies that use the lifestyle polygraph examination as part of the clearance process. Heuer stated "In addition to showing the significance of drug issues, this documents the value of the lifestyle polygraph in surfacing drug use that escapes identification by other investigative tools."

During the course of a BI, very little credible adjudicable information is developed during the interview phase of an investigation except for the subject interview. The Defense Personnel Security Research Center conducted a study (Carney, 1996) in which a team of personnel security adjudicators reviewed 1,028 cases; 290 from the Department of Defense, 159 from the Central Intelligence Agency, 282 from the National Reconnaissance Office, and 297 from the Office of Personnel Management. The study stated in its findings that adjudicable information from sources other than the subject was developed in only 19% of the cases. In only 2.4% of the cases the applicant was either denied clearance or resigned because of security issues. Out of the 1,028 cases, there were a total of 45 individuals that were denied access. In 20 of the denied cases the significant source of the adjudicable information was the polygraph.

During the same study a detailed analysis was given to the number of issues that were developed from each type of interview and record check. The most productive interview was that of the ex-spouse. In 12.3% of the cases (20 out of 163) some adjudicable information was developed. The next highest was former employer interviews at 5.6% (56 out of 1,005). While the subject is unquestionably the most productive source, the information the subject provided was self serving. When comparing the severity of the information provided by the subject with the severity of the information provided by the source (i.e. ex-spouse, former coworkers, employers, friends, etc.) the two agreed in only 6.7 % of the denied cases. In the remaining 93.3% of the cases the subject consistently under reported the severity of the issue in the case. The use of the polygraph tells a different story. For the denied cases there was an 88.2% agreement between severity of information developed by the polygraph with the information developed from other sources. As an example, a person will readily admit to experimentation with marijuana, but will be reluctant to divulge recreational or habitual use. The reason is clearly self-interest. The typical interviewee will only provide enough to convince the interviewer that there is no problem, so to qualify for the job or access. While investigators can go to employers, teachers, neighbors, and friends, the best source of factual information is the subject, and obtaining this information depends on the subject's willingness to speak candidly. With the polygraph as part of the screening process, the subject's ability to understate his or her actual behavior is diminished, as the individual knows he cannot pass unless he tells the whole truth.

The most productive record checks were medical records. Of the medical records reviewed, 5.6% (10 out of 178) produced some useful information. Individuals outside of law enforcement may find this surprising. Many people believe the National Agency Check (NAC) which incorporates a review of FBI files

would produce more issue information. The fact is that only an estimated one-third of all crime is reported to the police, and of the one-third reported, an arrest is made in less than one-fourth of those cases. Compound this with prosecutorial discretion where on average only a little over 50% of those arrested are actually prosecuted. Studies also reflect of those convicted, disposition of their case after arrest is estimated to be recorded only about 60% of the time. In sum, when a crime is committed, the perpetrator only has about a 3% chance of having his name recorded in the FBI files.

The Department of Defense published a document in 1984, titled The Accuracy and Utility of Polygraph Testing. One study looked at polygraphs given to 318 applicants for employment with the National Security Agency in FY 81. These applicants, who were in the military or had recent military service, were cleared for Sensitive Compartmented Information access and served in cryptologic positions.

Table 1 reflects that from the 318 applicants, 343 separate admissions were made concerning adverse adjudicable information. Some of the issues included various degrees of terrorist activities, disclosure of classified information, drug and alcohol use, involvement in serious crime, and deliberate falsification of security forms. This derogatory information was developed after all had undergone Special Background Investigations and some had update investigations and the individuals were in a position of trust.

Table 1 lists only that adjudicable information obtained during the polygraph examination which was not in the security file. The table excludes all information obtained during BIs. This information strongly suggests that individuals with serious background flaws are not deterred by the BI.

Also published in the Department of Defense document was a study that compared adjudicable information obtained from polygraph interviews with adjudicable information obtained from the special background investigation. The investigators were to look at the investigations separately and review the developed information. Some of the "dual track" cases had to be discarded from the study because information developed during the polygraph interview was totally disqualifying, and the Defense Investigative Service was advised to stop the investigation. Other cases were deleted from the study because information from the polygraph interview was given to the Defense Investigative Service as leads for further investigation. In the remainder of the cases polygraph interviews and background investigations were conducted independently, without exchange of information. From a total of 248 randomly selected dual track cases, 194 cases were selected as meeting the criterion of complete independence, required for the study. The results of the study revealed that 78% of all information used for evaluation was obtained from polygraph reports. The remaining 22% of information used for evaluation came from the background investigation. In 113 cases (58%), the polygraph report was of particular value to the clearance adjudicator because it was the only investigative method to develop information for evaluation. By comparison, the BI developed the only information in two cases (1%). This study supports the hypothesis that the application of polygraph techniques produce more adjudicable information than standard BI.

Table 1

Admissions During Polygraph Interviews

Fiscal Year 1981, from 318 Military Applicants

Topics	Number of Person Making Admissions	Percentage
Communist, Fascist or Terrorist Activity	1	0.3%
Divulgence of Classified Information:		
To family	13	4.1%
To others	10	3.1%
Soviet Bloc Travel	15	4.7%
Soviet Bloc Contacts	6	1.9%
Drugs:		
Marijuana	141	44.3%
Marijuana over 100 times	15	5.0%
Uppers, Downs, etc.	31	9.7%
Cocaine and Heroin	12	3.8%
Crimes:		
Misdemeanors	38	11.9%
Felonies	12	3.8%
Alcohol Problems:		
Past	7	2.2%
Present	1	0.3%
Deliberate Falsification of Security Forms	7	2.2%
Serious Credit Problems	14	4.4%
Psychological Problems:		
Treatment	14	4.4%
Suicide Attempts	5	1.6%

Polygraph Validity

Polygraph testing is used by the US government for pre-employment and security screening on individuals requiring special access to select programs, materials or agencies. Norman Ansley (1996) wrote the following concerning the accuracy of the polygraph:

"In the past 75 years over 250 studies have been conducted on the accuracy of polygraph testing....the preponderance of available information indicates that when a properly trained examiner utilizes an established testing procedure, the accuracy of the decision made by the polygraph examiner is generally in the range of 85-95% for specific issue testing. (P. 138,139)"

Since many different types of techniques and testing parameters were used to determine validity and accuracy it is difficult to draw from the data a precise figure for the accuracy of polygraph testing in all settings. Pre-employment or security screening polygraph testing is different from specific issue testing as the questions are more numerous and cover a broader area of topics. Norman Ansley (1996) had the following to say about the accuracy of pre-employment testing:

"To date there has been only a limited number of research projects on the accuracy of polygraph testing in the pre-employment context, primarily because of the difficulty in establishing ground truth. However, since the same physiological measures are recorded and the same basic psychological principles may apply in both the specific issue and pre-employment examinations, there is no reason to believe that there is a substantial decrease in the accuracy rate for the pre-employment circumstance. The few studies that have been conducted on pre-employment testing support this contention. While the polygraph technique is not infallible, research clearly indicates that when administered by a competent examiner, the polygraph test is one of the most accurate means available to determine truth and deception. (p. 139)"

Without question, examiner competence and the way they are trained will have a direct relationship on the validity of the test they conduct. This is why the polygraph community, while we have made great strides in recent years, must continue to standardize our training and techniques.

The Department of Defense maintains very stringent standards for polygraph examiners. They have also recently instituted a Quality Assurance Program to assist DoD and federal agencies in standardizing and following established procedures. All DoD and federal polygraph examiners have at least a 4 year college degree and complete a course of instruction at the DOD Polygraph Institute, which has curriculum on forensic psychophysiology meeting the requirements of a master's degree-level of study.

Cost Considerations

Looking at the cost factor and comparing a BI to the polygraph interview requires working with estimated numbers. The cost is dependent on the type of leads required by the scope of the investigation. The cost of a Single Scope Background Investigation within the Department of Defense is currently in excess of $1,700.

The cost of a polygraph interview is dependent upon whether the subject comes to the examiner, or if the examiner goes to the subject. Much of the cost attributed to the polygraph is examiner salary, travel and equipment. For the present comparison analysis an examiner can

conduct 300 screening polygraphs a year at a rate of two per day. This factors in the time required to write the polygraph report, 40 hours of required training per year, vacation time, and time consumed by no-shows. When calculating the examiner's salary along with one quality control and one administrative support person for every ten examiners, training, equipment and administrative costs, screening examinations could be conducted for as little as $300. For the sake of this paper and knowing that some cases will require travel by the examiner to the subject, an average cost of $350 will be used. Adding the cost of credit and national agency checks will bring the total for a polygraph interview investigation to $380.

A cost comparison of these investigations reveals a substantial difference: approximately $1,700 for a standard BI and $380 for a polygraph interview investigation. A significant cost savings of approximately $1300 per case. One agency alone closed 57,919 initial cases last year and could theoretically have realized a cost savings of over $75,000,000.

Time Considerations

Another important factor is that a polygraph interview investigation would take considerably less time to conduct than a normal BI. The polygraph portion could be scheduled and completed within weeks and adjudicators would just have to wait for the outcome of the credit and national agency checks.

In one study, employees were hired on a probationary status pending the outcome of their BIs. They were also administered initial polygraph examinations, and 65 employees were terminated immediately based on information derived from the polygraph. This 1953 study discussed in the Department of Defense publication, The Accuracy and Utility of Polygraph Testing stated:

"Calculated on the basis of a nine-month clearance period at the GS-4 level, this group of individuals would have cost the government $1,814,322 in salary alone while awaiting the completion of a background investigation. This figure is low, since the average salary of individuals awaiting clearances in the agency was above the GS-4 level and the average background time was more than nine months." (p. 19)

We must remember that this was a 1953 study; the conduct of a BI today is somewhat shorter. Conversely, salaries have also increased dramatically in the 44 years since this study. Some statistics show that initial BIs are now taking between 120 and 180 days.

Deterrence of the Polygraph

Christopher Boyce, better known as the Falcon in The Falcon and the Snowman, (INSCOM, 1991) is serving 68 years in prison for espionage. He stated before Congress that in regard to the polygraph he "greatly feared it." Had he known about the polygraph he would not have considered an act of espionage. He further stated that polygraph should be considered one of the best deterrents to those toying with the thought of espionage. Recent Espionage Cases, produced by the Department of Defense Security Institute, is a short study of 88 individuals arrested on espionage charges from 1975 to 1989. Many of these individuals talked of their fear of the polygraph in detecting their activities. James Hall III, an Army Warrant Officer who was convicted of selling Top Secret information to East Germany and the Soviet Union stated in an interview that had he been scheduled for a polygraph, he probably would have

turned himself in because "I know I would have failed the test." He went on to say that "One of the good approaches to detecting people engaged in espionage is the polygraph." These two examples support a hypotheses that the use of the polygraph provides a deterrent effect not observed with the standard BI.

The deterrence value is the most difficult part of an investigation to measure. The investigative process should not only detect past and present activity, but should deter future activity of a derogatory nature knowing the activity will come out in a subsequent investigation. This could be one of the more beneficial aspects of the polygraph technique.

Summary

The literature clearly shows that a standard BI is defective in ascertaining a large amount of adjudicable information. Evidence to demonstrate the deterrence power of the BI is lacking. While the utility of the polygraph, in conjunction with the BI, is our best defense to ensure the security of our national secrets, an argument can be made that the polygraph alone is less time consuming, less costly, less manpower intensive, produces more adjudicable information, and produces a deterrence not present with a normal BI. As government continues to shrink, and budgets are cutback, it is incumbent upon government security programs to maximize the use of their remaining investigative tools. While money could be saved by conducting more interview telephonically or through the mail, but the quality of the investigation would certainly be compromised. The polygraph as a tool has great potential, and when utilized correctly, can replace other investigative methodologies. The data justifies further research on a large body of cases. While the literature tends to show advantages of polygraph utility over the standard BI, a new government dual-track study is needed to take in the advances of modern technology, changes in the background investigatory procedures, and the computerized polygraph.

References

Ansley, N. (1996). The Accuracy of Polygraph Examinations, Polygraph, 25(2) 138-139.

Carney, R. M. (1996). SSBI Source Yield: An Examination of Sources Contacted During the SSBI. Defense Personnel Security Research Center.

Department of Defense, (1984), The Accuracy and Utility of Polygraph Testing, US Government Printing Office, 18 - 22.

Department of Defense, (1990), Recent Espionage Cases, DoD Security Institute, 2-26.

Heuer, Richards J. Jr. (1995), Drug Use and Abuse: Background Information for Security Personnel, Polygraph, 24(3), 154-155.

US Army, Headquarters, Intelligence and Security Command, (1991), The Counterintelligence Scope Polygraph Examination Program, PAN A2101-91-002

An Analysis Of The Psychodynamics Of The Directed Lie Control Question In The Control Question Technique

James Allan Matte, Ph.D.

Abstract

The Directed-Lie Control Question (DLCQ) has been recently introduced as a replacement for the traditional Probable-Lie Control Question (PLCQ) in control question tests. Its proponents assert that the DLCQ improves the standardization of comparison question formulation, increases the accuracy of psychophysiological veracity examinations using the polygraph, and eliminates the manipulation and alleged intrusiveness required by the Probable-Lie Control Question. This study reviews and evaluates existing studies of the DLCQ and further analyzes the psychodynamics of the DLCQ versus the PLCQ to determine its construct and criterion validity.

Keywords: Comparison question test, control question test, detection of deception, directed lie, PDD, polygraph, probable lie, validity

Background:

The control question technique includes a variety of control question tests (Reid, Marcy, Arther, Backster, Matte, DoDPI, Utah, Integrated) (Matte, 1996). All of them employ some type of control question, either non-current exclusive, current exclusive or non-exclusive (Matte, 1996) designed to elicit a probable-lie from both the Innocent and Guilty examinee, which is used for comparison with the relevant test question(s) contained in the same test.

The Zone Comparison Test (ZCT) developed by Cleve Backster is probably the most researched and utilized control question test in the criminal justice system, and analog and field validation studies have demonstrated that the ZCT possesses a very high degree of accuracy and reliability. (Arellano, 1990; Elaad & Schahar, 1985; Matte & Reuss, 1989; Putnam, 1983; Raskin, Barland, & Podlesny, 1978; Widacki, 1982).

The author wishes to express his sincere appreciation for the critical reviews of this article by Cleve Backster and Stan Abrams, Ph.D., and to Donald J. Krapohl for his insightful suggestions for revision.

Dr. Matte has published numerous articles in *Polygraph*, and he is the author of Forensic Psychophysiology Using the Polygraph: Scientific Truth-Verification-Lie Detection. The author can be contacted for reprints and information at J.A.M. Publications, 43 Brookside Drive, Williamsville, NY 14221-6915, or e-mail at JAMpublications@mattepolygraph.com.

An Analysis Of The Psychodynamics Of The Directed Lie Control Question In The Control Question Technique

James Allan Matte, Ph.D.

Abstract

The Directed-Lie Control Question (DLCQ) has been recently introduced as a replacement for the traditional Probable-Lie Control Question (PLCQ) in control question tests. Its proponents assert that the DLCQ improves the standardization of comparison question formulation, increases the accuracy of psychophysiological veracity examinations using the polygraph, and eliminates the manipulation and alleged intrusiveness required by the Probable-Lie Control Question. This study reviews and evaluates existing studies of the DLCQ and further analyzes the psychodynamics of the DLCQ versus the PLCQ to determine its construct and criterion validity.

Keywords: Comparison question test, control question test, detection of deception, directed lie, PDD, polygraph, probable lie, validity

Background:

The control question technique includes a variety of control question tests (Reid, Marcy, Arther, Backster, Matte, DoDPI, Utah, Integrated) (Matte, 1996). All of them employ some type of control question, either non-current exclusive, current exclusive or non-exclusive (Matte, 1996) designed to elicit a probable-lie from both the Innocent and Guilty examinee, which is used for comparison with the relevant test question(s) contained in the same test.

The Zone Comparison Test (ZCT) developed by Cleve Backster is probably the most researched and utilized control question test in the criminal justice system, and analog and field validation studies have demonstrated that the ZCT possesses a very high degree of accuracy and reliability. (Arellano, 1990; Elaad & Schahar, 1985; Matte & Reuss, 1989; Putnam, 1983; Raskin, Barland, & Podlesny, 1978; Widacki, 1982).

The author wishes to express his sincere appreciation for the critical reviews of this article by Cleve Backster and Stan Abrams, Ph.D., and to Donald J. Krapohl for his insightful suggestions for revision.

Dr. Matte has published numerous articles in *Polygraph*, and he is the author of Forensic Psychophysiology Using the Polygraph: Scientific Truth-Verification-Lie Detection. The author can be contacted for reprints and information at J.A.M. Publications, 43 Brookside Drive, Williamsville, NY 14221-6915, or e-mail at JAMpublications@mattepolygraph.com.

However, the methodology used to develop and effectively introduce the probable-lie control question (PLCQ) requires the psychological manipulation of the examinee in order to elicit a probable-lie that will function as designed. This is a complicated procedure which requires a complete understanding of the psychological factors involved, and much experience in its application. It is therefore not surprising that some forensic psychophysiologists have readily accepted the introduction of a new application to the probable-lie control question, which removes the necessity for the careful development and psychological manipulation of the examinee, with the utilization of an old concept, the Directed Lie (Fuse, 1982; Golden, 1969; Reali, 1973) to the probable-lie control question. Briefly stated, the examinee is directed to lie to a probable-lie control question which the examinee perceives or is told will provide his/her unique physiological pattern of his/her lie to be used for comparison to the neighboring relevant question(s) included in the same test. Proponents of the Directed Lie Control Question (DLCQ) assert that DLCQ's are easy to construct, and the same DLCQ's can be used for all tests. They eliminate the psychological manipulation required by Probable-Lie questions, and they involve no invasion of privacy. (Horowitz, Kircher, Honts, & Raskin, 1997).

Research:

Several research studies have been conducted on the Directed Lie, involving different test formats. In 1981, Dr. Gordon H. Barland reported his results of a Validity and Reliability Study of Counterintelligence Screening Tests, wherein he used 56 U.S. Army employees in a mock screening situation to determine the accuracy of PV examinations with the Counterintelligence Screening Test (CIST) using directed lie questions. This CIST is derived from the federal version of the Zone Comparison Test. The test contained 13 questions: five relevant, four directed lies, two symptomatic, one sacrifice-relevant and one irrelevant. Five different models of field polygraph instruments were used which recorded respiration, skin resistance and relative blood pressure. This study demonstrated that the Directed-Lie control question identified the truthful significantly better than chance but failed to identify the deceptive subjects on the relevant questions much better than chance levels. (Barland, 1981)

Millard E. Addison, Special Assistant for Scientific Investigations at the Department of the Navy Naval Investigative Service conducted a study during the period 1977-1978 on Silvestro F. Reali's Lie/Truth (Positive Control) technique which employs directed lies for comparison with their neighboring relevant questions. Addison had his ten forensic psychophysiologists use the Positive Control technique on various types of criminal and counterintelligence cases after an MGQT, Zone Comparison Test or Relevant-Irrelevant test to see if the two tracked. They "found the two tracked on approximately 95 percent of the NDI (No Deception Indicated) calls, were inconclusive in 4 percent, and DI in less then 1 percent; however, on DI (Deception Indicated) calls, we found only a 60 percent tracking ratio with 40 percent NDI or inconclusive" (Addison, 1982).

In 1987, an analog study on the "Validity of the Positive Control Physiological Detection of Deception Technique" was conducted by Lawrence N. Driscoll, Charles R. Honts, Ph.D., and David Jones, J. D., Ph.D. The

results of this experiment "indicate the positive control test to be an inferior detection of deception technique as compared to the control question test. This finding is indicated by the dramatically increased percentage of inconclusive outcomes for the positive control (45 percent) as compared to the control question test (10 percent), and in an increased false negative rate for the positive control test (22 percent) as compared to the control question test (0 percent)...the positive control test was not demonstrated to be a valid discriminator of truthtellers and deceivers." (Driscoll, Honts, & Jones, 1987).

It is recognized that the Positive Control Test (PCT) does not employ the same test format as the traditional control question test. However the Positive Control Test does employ Directed Lies for comparison against its neighboring relevant questions, and the psychodynamics of the Directed Lie can be seen from the results of its inordinate percentage of false negatives which imply that the same elements depicted in Table 1, were present in aforesaid studies of the Positive Control Test. Furthermore, when a guilty examinee is instructed to tell a lie during the conduct of a PCT, the guilty examinee suffers a conflict in that he/she is compelled to give an affirmative answer(admit) to a crime he/she committed. In addition, the guilty examinee is instructed to tell a lie but instead verbalizes a truthful answer which he/she portrays as a lie hence is deceiving the examiner. M.E. Addison, Special Assistant for Scientific Investigations, U.S. Navy in reporting his research findings of the PCT to this author stated that "when a guilty examinee responds affirmatively to an offense he/she committed, it may have a greater emotional effect than responding negatively to the same question." All five elements depicted in Table 1 responsible for potential false negatives are present in the Positive Control Technique as they are in the Directed Lie Control Question Technique. We must understand that it is not the dissimilarities between the two techniques that are important, but the similarities which encompass the same elements listed in Table 1, that reflect the causative factors for the false negatives in the DLCQ.

Use of the Directed-Lie control question was reported by L. S. Fuse in 1982, wherein he indicated that the DLCQ had evolved over the previous sixteen years and was found to be most effective in multiple-issue tests. However, Fuse cautioned that there had to be the correct amount of emphasis on the directed lie because excessive emphasis would dampen the response to the relevant question, and insufficient emphasis could cause a false positive response. (Fuse, 1982)

In 1988, Honts and Raskin conducted *A Field Study of the Validity of the Directed Lie: Control Question*, wherein they reported on 25 confirmed criminal tests in which the Directed Lie control question (DLCQ) procedures were used. Confirmation (ground truth) was acquired through admissions, physical evidence that conclusively exonerated the examinee, or there was a retraction of the allegation. Regarding the latter, the alleged victim retracted the allegation, denying that the offense had taken place. Two control questions and one DLCQ were used for comparison with three relevant questions in each one of these tests. Using this method, the researchers reported that "The original examiners in these cases reached conclusive decisions on 24 of the 25 cases and 92 percent of those decisions were correct; there was one false-negative error." The results of blind scoring revealed a 90 percent accuracy when only traditional

Table 1

COMPARISON OF PLCQ AND DLCQ POTENTIAL RESPONSE ELEMENTS

Response Elements	Focus of Psychological Set and Potential Response PLCQ		DLCQ	
Subjects:	**Innocent**	**Guilty**	**Innocent**	**Guilty**
1. Fear of Detection	Yes	No	No	No
2. Fear of Error	No	No	Yes	No
3. Hope of Error	No	No	No	Yes
4. Fear of Physiological Comparison	No	No	No	Yes
5. Perceived as Threat to Outcome of Test	No	No	Yes	Yes
6. Perceived Threat of Past Offense Reflecting on Capacity to Commit Current Offense	Yes	No	No	No
7. Shame-Embarrassment	Yes	No	No	No
8. Perceived Relationship with Relevant Question	No	No	Yes	Yes
9. Invitation to Counter-measures	No	No	No	Yes
Total Affirmatives	3	0	3	5
Potential Errors		False Negative		False Negative

control questions were used, with both errors being false positives. When using both traditional controls and one DLCQ, an accuracy of 95.6 percent was obtained, with one false negative as the only error. Dr. Abrams (1991) expressed his concern about the criteria used to establish ground truth in 11 of the 25 subjects used in Honts, *et al.* study, inasmuch as those eleven subjects were suspects in child sexual abuse cases and one of the criteria used for verification was the retraction of an allegation. Dr. Abrams, a Clinical Psychologist and Forensic Psychophysiologist stated that "It is not at all unusual for a child victim of sexual abuse to retract his or her accusation, but that does not necessarily mean that the abuse did not occur." Abrams quoted Toth & Whalen (1987) "Whatever a child says about sexual abuse, she is likely to reverse it. Beneath the anger of impulsive disclosure remains the ambivalence of guilt and the martyred obligation to preserve the family. In this chaotic aftermath of disclosure, the child discovers that the bedrock fears and threats underlying the secrecy are true. Her father abandons her and calls her a liar. Her mother does not believe her and decompensates into hysteria or rage." (Abrams, 1991).

In 1991, Dr. Stan Abrams studied the directed-lie control question approach in ten verified cases consisting of six confirmed deceptive subjects and four confirmed truthful subjects. In all ten real-life cases, verification was determined by confessions. The instructions given to the subjects were taken verbatim from an audio-taped examination conducted by Dr. Raskin. The wording was exactly the same and a very strong effort was made to maintain the same inflection. All of the polygraph charts were numerically scored using the traditional seven-position scale (+3, 0, -3). However Dr. Abrams admits that "since this was an experimental procedure, the writer did not feel that he could risk jeopardizing the polygraph findings. Therefore instead of the DLCQ being utilized once in each test in the series, as ordinarily would be the case, it was employed only in the last test of the series and it followed the final relevant question. In this way it would not impinge on the test results in any way that could invalidate the test if it were to be used as evidence in court." The results revealed that all four truthful subjects' scores increased from an average score of +2.75 to +5.25. However, five of the six deceptive subjects' scores increased from an average score of -1.6 (deception) to an average score of +3.6 (truthful) and the sixth deceptive subject's score increased from zero (inconclusive) to +4 (truthful). Thus, five of the six confirmed deceptive subjects produced minus (deception) scores and one produced a score of zero (inconclusive) when the relevant questions were compared to the normally used control questions, but when compared to the DLCQ, all six confirmed deceptive subjects produced plus (truthful) scores, hence false negatives for all six confirmed deceptive subjects. Apparently, the same factors that caused the unacceptable number of false negatives when using the DLCQ in studies conducted by Driscoll, et al. (1987), Addison (1982), Barland (1981) were operative in Abrams' study.

An analog study was conducted by Horowitz, Kircher, Honts, and Raskin (1997) in which "The Role of Comparison Questions in Physiological Detection of Deception" was examined. In this study, comparison questions in physiological detection of deception examinations were studied with 60 guilty and 60 innocent participants in a mock crime experiment." Different types of

comparison question were used in four conditions: relevant-irrelevant (RI) participants answered only relevant and neutral questions; trivial directed lie (TDL) participants were instructed to lie to three of the six neutral questions; personal directed lie (PDL) participants were instructed to lie to personally relevant questions; and probable lie (PL) participants received traditional probable lie comparison questions." Respiration, cardiovascular, vasomotor, and electrodermal activity were recorded but the procedure deviated from field methods in several respects. "Respiration was recorded from two 25-cm Hg strain gauges placed around the upper thorax and the abdomen. SC was recorded from two 10-mm Ag/AgCl electrodes filled with 0.05 molar NaCl in a Unibase medium attached with adhesive collars to the middle and ring fingers of the left hand. Electrocardiograms were acquired from Lead II." A Beckman Type R Dynograph was used to acquire the physiological recordings. The authors reported that "manipulation of the comparison questions produced different patterns of physiological responses for Innocent but not for Guilty participants. The Relevant-Irrelevant (RI) test produced an unacceptable rate of false positive decisions."

Horowitz, *et* al. reported that "Excluding inconclusives, the R-I test produced 100% correct decisions for guilty participants but only 22% correct decisions for innocent participants. The highest overall accuracy was obtained for participants in the PDL condition (85.7% correct decisions)." Horowitz further reported that "Because the R-I test has been criticized for producing a high number of false positive outcomes, the number of errors produced by each of the four test structures for only the innocent participants was compared.[1]* The R-I test produced 11 false positive results, whereas the other tests produced only 2 each." However a review of the data by this author in Table 3 of cited study shows that for the Guilty experimental group, the TDL group had 73% correct decisions, the PL group had 73% correct decisions, and the PDL group had 84% correct decisions. For the Innocent experimental group, the TDL group had 84% correct decisions, the PL group had 86% correct decisions, and the PDL group had 87% correct decisions. However, for the Innocent experimental group, the Personal Directed Lie (PDL) generated negative scores for respiration. "In fact, the PDL yielded respiration scores that were significantly more negative than those obtained with the R-I test. Only the PL comparison question generated respiration length scores in the predicted direction. Inspection of the numerical scores for respiration confirmed these results."

Horowitz noted that the "respiration responses by Innocent participants to DL questions (both PDL and TDL) were opposite to that predicted by prior research, whereas respiration responses by PL participants were as strongly in the predicted direction." Horowitz stated that "respiration may be the least reliable physiological measure when scored numerically, and respiration length had the largest drop in validity when the computer scoring model was cross-validated," citing Kircher & Raskin, 1988. They suggested that "When DL questions are

---

[1] Other research studies (Barland 1981; Addison 1982; Driscoll, *et.al.* 1987; Abrams 1991) mentioned in this analysis reveal that the Directed Lie control question errs mostly in favor of the guilty subject (False Negatives).

used, perhaps respiration responses should not be used or should be weighted the least of the physiological measures." They also indicated "that personal significance increases the potency and utility of the DL questions. The PDL questions produced stronger cardiovascular and skin conductance responses by innocent participants than did the TDL or PL questions." The authors temper their recommendation in the use of the directed lie comparison question with "the caveat that their study is a laboratory analog of a field situation that is difficult to model."

Aside from the fact that the psychodynamics of the aforesaid (Horowitz, et al 1997) analog study are quite different than field studies which incorporate the strong emotions of fear of detection by the guilty, fear of error by the innocent, and potential anger, which are all lacking in analog studies, this study rationalizes the dismal performance of the Directed Lie with the Innocent participants when evaluated through their respiration patterns, by stating that respiration may be the least reliable physiological measure when scored numerically. However, it should be noted that in many field studies, respiration recorded by the pneumograph was shown to have equal diagnostic value and in some field studies respiration had greater diagnostic value than its neighboring parameters (Buckley & Senese, 1991; Elaad, 1985; Elaad & Kleiner, 1990; Matte & Reuss, 1992; Nakayama & Yamamura, 1990; Slowick & Buckley, 1975). An experimental scoring technique proposed and tested by Jayne (1990) also supported the pneumograph as providing the most diagnostic information. The electrodermal (GSR) is sometimes the least effective parameter (Jayne, 1990; Matte & Reuss, 1992). Furthermore, a study by Elaad, Bonwitt, Eisenberg, Meytes in 1982 revealed that respiration was the only one of the three parameters (pneumo, GSR, cardio) not affected by beta blockers. Interestingly, Elaad, *et al.* concluded that "respiration seemed to improve the overall detection rate especially because skin resistance responses have the quality of rapid habituation." Thus respiration appears less vulnerable to habituation.

Analysis:

The Directed Lie Control Question (DLCQ) may be perceived by the guilty examinee as a relevant exemplar of deception question, inasmuch as the guilty examinee is led to believe or perceives that the DLCQ will provide the forensic psychophysiologist (FP) with a physiological fingerprint of his/her lie pattern which will be compared with his/her response to the neighboring relevant test question. Even the subtle presentation of the DLCQ must raise the guilty examinee's suspicion that the acquisition of his/her known lie pattern is for the purpose of comparison with his/her physiological pattern to the relevant test questions. Hence the DLCQ may well be perceived by the guilty examinee as a deception exemplar question. The guilty examinee thus is likely to believe that without the DLCQ, identification of a lie to the relevant test question is significantly reduced or improbable. The focus of the guilty examinee will normally be on the test question that offers him/her the greatest threat to his/her immediate security. The DLCQ offers a new and immediate competing threat to the guilty examinee that may be equal or even greater than the threat offered by the relevant question because of its perceived capacity to identify his/her lie to the relevant question. The DLCQ further invites the guilty examinee to employ countermeasures to defeat the perceived purpose of the DLCQ which in turn elicits greater mental

exercise and selective attention (psychological set) from the guilty examinee. In short, the DLCQ offers the guilty examinee a new and immediate competing threat which may be equal or greater than that offered by the expected relevant test questions.

Table 1 depicts a construct validity analysis of the Probable-Lie Control Question versus the Directed-Lie Control Question in the control question technique. The Table shows that the DLCQ contains five elements (Nr. 3, 4, 5, 8, 9) which may cause a false negative result, whereas the PLCQ contains none. Both the DLCQ and the PLCQ each contain three elements (PLCQ 1, 6, 7; DLCQ 2, 5, 8) which may elicit a response from the Innocent subject which indicates an approximate equal capacity to identify the Innocent examinee.

The negative effect of the evidentiary element of the DLCQ is exacerbated by the methodology of the DLCQ which particularly emphasizes the review of the DLCQ between the collection of each polygraph chart data (Horowitz, *et al.*1997), thus inordinately increasing the strength of the DLCQ resulting in undue influence and alteration of the guilty examinee's psychological set. This routine manipulation of the examinee's psychological set, without specific evidence of a control question malfunction, raises the potential for a false negative result to a higher probability.

The traditional reviewed probable-lie control question is designed to be structurally weaker than its neighboring relevant questions so that it will not offer an equal threat to the guilty examinee thus causing an inconclusive result. This is accomplished by having the control question embrace a similar category of offense but separated from the relevant question with a time bar that significantly removes the control question from the current time frame occupied by the relevant question. Thus while the relevant test question offers an immediate threat to the well-being of the guilty person, the earlier-in-life control question offers a threat of lesser consequence because any transgression hidden by the control question occurred much earlier in the examinee's life, and not an issue of the examination.

There are a variety of fears, and the comparison between relevant and control questions should encompass the same type of fear, that is, fear of detection in a lie, ideally regarding a similar offense, one deliberately weaker than the other. The DLCQ does not offer a fear of detection in a lie since the lie is directed. But the DLCQ is believed to provide incriminating evidence identifying the guilty examinee's lie pattern to the relevant questions, thus represents a significant threat to the guilty examinee's ability to conceal his/her lie to the relevant questions, hence rendering the DLCQ a relevant deception exemplar question of an equal or greater threat to the guilty examinee than the relevant test question against which it is being compared.

The DLCQ can also be compared to the deception exemplar attribute of the Key question in a Control-Stimulation Test (CST). The Control-Stimulation Test is designed to reassure the innocent examinee of the accuracy of the test and of the competency of the forensic psychophysiologist. It also serves to stimulate the guilty examinee. The CST further serves as a Control test to establish the examinee's capability and manner of response to a known lie under controlled conditions. (Hickman, 1978; Lovvorn, 1978). It can further serve to identify a Spot Responder or prove

the falsity of such a claim by a deceptive examinee. (Decker, 1996). The Control-Stimulation Test's Key question is usually perceived by the examinee as a Deception Exemplar relevant question inasmuch as it is introduced to the examinee as a means of acquiring a physiological sample of the examinee's lie pattern, as described above, which is another reason for administering the CST as the first test in the series of tests to be administered to the subject. The examinee will from the start know that the CST was the device used to acquire that data, and will thus not associate that role to the exclusive control questions used in the zone comparison test.

Conclusion:

The Probable-Lie Control Question has been shown to be a valid and reliable control question in control question tests both in laboratory and field studies. The Directed Lie Control Question however appears to lack construct validity in that it fails to demonstrate adequate capacity to function as a control question of less intensity than the relevant question for the guilty examinee. It further lacks criterion validity in that there appears to be inadequate data demonstrating a relationship between test outcomes and a criterion of ground truth supporting the effectiveness of the DLCQ for all examinees in control question tests. There exists a wholly insufficient body of data to support its use as a valid and reliable control question in control question tests.

References:

Abrams, S. (1991). The directed lie control question. *Polygraph*, 20(1), 26-31.

Abrams S. (1976). The control question: A technique for effective introduction. *Polygraph*, 5(4), 290-292.

Addison, M. E. (1982, May 9, 11). Letters from Dept. of the Navy's Special Assistant for Scientific Investigations to J. A. Matte.

Ansley, N. (1990). The validity and reliability of polygraph decisions in real cases. *Polygraph*, 19(3), 169-181.

Arellano, L. R. (1990). The polygraph examination of Spanish speaking subjects. *Polygraph*. 19(2), 155-156. [abstract]

Backster, C. (1990). Backster zone comparison technique chart analysis rules. Handout at lecture of the 25th annual seminar/workshop of the American Polygraph Association, 14 August 1990, Louisville Kentucky.

Barland, G. H. (1981, May 12). A validity and reliability study of counterintelligence screening tests. Unpublished manuscript, Security Support Battalion, 902nd Military Intelligence Group, Ft. George G. Meade, MD.

Barland, G. H. & Raskin, D. C. (1973). Detection of deception. In W. F. Prokasy & D. C. Raskin (Eds.), Electrodermal activity in psychological research (418-477), New York: Academic Press.

Buckley, J. P., and Senese, L. C. (1991). The influence of race and gender on blind polygraph chart analyses. *Polygraph*, 20(4), 247-258.

Capps, M. H., and Ansley, N. (1992). Anomalies: The contributions of the cardio, pneumo, and electrodermal measures towards a valid conclusion. *Polygraph*, 21(4), 321-340.

Decker, R. E. (1996, January 16). Telephone conversation with J. A. Matte.

Department of Defense Polygraph Institute Research Division Staff. (1995, June). A comparison of psychophysiological detection of deception accuracy rates obtained using the counterintelligence scope polygraph and the test for espionage and sabotage question formats. *Polygraph*. 26(2). 79-106.

Driscoll, L. N., Honts, C. R., Jones, D. (1987). The validity of the positive control physiological detection of deception technique. *Polygraph*. 16(3), 218-225.

Elaad, E. (1985). Validity of the control question test in criminal cases. Unpublished manuscript. Scientific Interrogation Unit, Criminal Investigation Division, Israel National Police Headquarters, Jerusalem, Israel.

Elaad, E., Shahar. E. (1985). Polygraph field validity. *Polygraph*, 14(3), 217-223.

Elaad, E., Ben-Shakhar, G. (1997). Effects of item repetitions and variations on the efficiency of the Guilty Knowledge Test. *Psychophysiology*, 34(5), 587-596.

Elaad, E., Bonwitt, G., Eisenberg, O., Meytes, I. (1982). Effects of beta blocking drugs on the polygraph detection rate: A pilot study. *Polygraph*, 11(3), 225-233.

Elaad, E., and Kleiner, M. (1990). Effects of polygraph chart interpreter experience on psychophysiological detection of deception. *Journal of Police Science and Administration*, 17(2), 115-123.

Fuse, L. S. (1982). Directed lie control testing (DLCQ) technique. Unpublished manuscript.

Golden, R. I. (1969, August). The 'Yes-No' technique. Paper presented at the American Polygraph Association Seminar in Houston, Texas.

Golden, R. I., and Hunter, F. L. (1970, Nov-Dec). The measurement of upper and lower respiration in lie detection. *American Polygraph Association Newsletter*, 4(6), 16-19.

Hickman, R. C. (1978). Usefulness and theory of the stimulus test. *Polygraph*, 7(3), 182-185.

Honts, C. R, and Raskin, D. C. (1988). A field study of the validity of the directed lie control question. *Journal of Police Science and Administration*, 16, 56-61.

Horowitz, S. W., Kircher, J. C., Honts, C. T, Raskin, C. D. (1997). The role of comparison questions in physiological detection of deception. *Psychophysiology*, 34(1), 108-115.

Horvath, F. (1991). The utility of control questions and the effects of the control question tests in field polygraph techniques. *Polygraph*, 20(1), 7-25.

Jayne, B. C. (1990). Contributions of physiological recordings in the polygraph technique. *Polygraph*, 19(2), 105-117.

Lovvorn, D. J. (1978). A modified controlled stimulation test technique. *Polygraph*, 7(3), 188-194.

Matte, J. A. (1996). Forensic Psychophysiology Using The Polygraph: Scientific Truth-Verification-Lie Detection. Williamsville, New York: J.A.M. Publications.

Matte, J. A. (1993). The review, presentation and assurance of intended interpretation of test questions is critical to the outcome of polygraph tests. *Polygraph*, 22(4), 299-312.

Matte, J. A., and Reuss, R. M. (1992). A study of the relative effectiveness of physiological data in field polygraph examinations. *Polygraph*, 21(1), 1-22.

Matte, J. A., and Reuss, R. M. (1989). A field validation study of the Quadri-Zone Comparison Technique. *Polygraph*, 18(4), 187-202.

Matte, J. A. and Reuss, R. M. (1989). Validation Study on the Quadri-Zone Comparison Technique. Research Abstract, LD 01452, Vol. 1502, 1989, University Microfilm International.

Nakayama, M. and Yamamura, T. (1990). Changes of respiration pattern to the critical question on Guilty Knowledge Technique. *Polygraph*, 19(3), 188-198.

Putnam, R. L. (1983). Field accuracy of polygraph in the law enforcement environment. Unpublished manuscript, American Polygraph Association archives.

Raskin, D. C., Barland, G. H., Podlesny, J. A. (1978). Validity and reliability of detection of deception. Washington, D. C.: National Institute of Law Enforcement and Criminal Justice.

Raskin, D. C., Kircher, J. C., Honts, C. R., Horowitz, S. W. (1988). Validity of control question polygraph tests in criminal investigation. *Psychophysiology*, 25(4), 474. [abstract]

Reali, S. F. (1978). Reali's Positive Control Technique. *Polygraph*, 7(4), 281-285.

Ryan, R. (1989). The accuracy of respiration, GSR, and cardiovascular polygraph responses utilizing numerical evaluation. Unpublished master's thesis, Reid College.

Slowik, S. and Buckley, J. (1975). Relative accuracy of polygraph examiner diagnosis of respiration, blood pressure, and GSR recordings. *Journal of Police Science and Administration*, 3(3), 305-310.

Toth, P. A., and Whalen, M. P. (Eds.) (1987). Investigation and Prosecution of Child Abuse. National Center for the Prosecution of Child Abuse, 1033 N. Fairfax Street, Suite 200, Alexandria, VA.

Widacki, J. (1982). Analiza prezestanek diagnozowania w. badanich poligraficznych (The analysis of diagnostic premises in polygraph examinations). Uniwersytetu Slaskeigo, Katowice. [Text in Polish; Abstract in English)

* * * * * *

Psychophysiological Detection of Deception Accuracy Rates Obtained Using the Test for Espionage and Sabotage

Department of Defense Polygraph Institute
Research Division Staff

Abstract

Previous research conducted by the Department of Defense Polygraph Institute (DoDPI) indicated that the decisions of examiners who administered the Test for Espionage and Sabotage (TES), were significantly more accurate at identifying programmed guilty examinees than were the decisions of examiners who administered either of two Counterintelligence Scope Polygraph (CSP) formats. The new format differs from previous security screening formats in that: (a) the number of issues being tested is reduced; (b) the number of repetitions of the questions used to calculate question scores is restricted to three; (c) between test stimulation is eliminated; (d) the order of questions within the question sequence cannot be altered; (e) each relevant question is compared to the same comparison questions; (f) the pretest is brief, more standardized and follows a logical sequence of information presentation; and (g) the Directed Lie Comparison (DLC) questions eliminate many of the problems associated with Probable Lie Comparison (PLC) questions. The procedures utilized during this study were identical to those in the previous study, but only the TES format was utilized. The replication was done in order to further validate the accuracy of the examiners' decision in identifying programmed guilty and innocent examinees, when the TES format was administered. The data collected in this study were evaluated using the new criteria developed from the previous study. Ten certified examiners from the Office of the Secretary of the Air Force conducted 88 examinations. The examiners had been trained to administer the TES and had been utilizing the TES when conducting security examinations. Ninety-eight percent of the innocent examinees and 83/3% of the programmed guilty examinees were correctly identified.

Keywords: Counterespionage, DLC, detection of deception, directed lie comparison question, espionage, polygraph, PDD, screening, TES

Acknowledgments: Sheila D. Reed, Ph.D., served as principle investigator throughout planning, data collection, and drafting of this manuscript. We appreciate the support of the Office of the Secretary of the Air Force (OSAF) polygraph program, and specifically Bruce Thompson and Jim Morrison. Also thanks to Edith Andreasen, Richard Baird, Ray Brafford, Debbie Habel, Michael Rhodes, Donald Schupp, Ed Stoval, James Vaughan, Michael Walker, Harrison Wright, Earl Taylor, Sam Braddock, Gordon Barland, Jeff St. Cyr, Linda Knickerbocker, and Joan Harrison-Woodard. This research was supported by funds from the Department of Defense Polygraph Institute's project DoDPI93-P-0045. The views expressed in this article do not reflect the official policy or position of the Department of Defense or the U.S. Government. Submit reprint requests to Andrew B. Dollins, Ph.D., Department of Defense Polygraph Institute, Building 3195, Fort McClellan, Alabama 36205-5113.

The accuracy of decisions for determining deception in a mock screening situation, using three psychophysiological detection of deception (PDD) formats has been compared [Department of Defense Polygraph Institute (DoDPI) Research Division Staff, 1997]. Two of the formats were counterintelligence scope polygraph (CSP) formats; one in which probable lie comparison (PLC) questions are asked and the other in which directed lie comparison (DLC) questions are asked. The third format was the test for espionage and sabotage (TES) in which: (a) the number of issues being tested is reduced; (b) the number of repetitions of the questions used to calculate question scores is restricted to three; (c) between test stimulation is eliminated; (d) the order of questions within the question sequence is constant; (e) each relevant question is compared to the same comparison questions; (f) the pretest is brief, more standardized and follows a logical sequence of information presentation; and (g) DLC questions are asked in place of the standard PLC questions.

The decisions of the examiners who administered the TES format were significantly more accurate (83.3%) at identifying the programmed guilty (PG) examinees than were the decisions of the examiners who administered either the CSP-PLC (55.6%) or the CSP-DLC (58.6%) format. There were no significant differences among the accuracies of the examiners' decisions at identifying the innocent examinees.

This study replicated the procedures utilized in the previous study, but only the TES format was administered. In addition, the data were evaluated using a scoring method that was developed using the data collected during the previous study (DoDPI Research Division Staff, 1997).

Methods

Examinees

Eighty-eight examinees were recruited by a local employment agency under contract to the Department of Defense Polygraph Institute and were paid $30.00 for their participation. Individuals who met the following criteria were excluded from participation: (a) less than 19 or more than 60 years of age, (b) not in good health, (c) pregnant, or (d) did not have the equivalent of a high school diploma. Thirty-four male (M = 27.2, SD = 9.7) and 54 female (M = 27.8, SD = 10.2) examinees were scheduled for testing. Thirty-three examinees were PG.

Examiners

Ten certified examiners (9 males and 1 female) were selected by and from the Office of the Secretary of the Air Force (OSAF)to conduct the examinations. Selection of the examiners was determined by the agencies. Although examiner selection was not random (selection criteria generally involve availability and experience), the examiners were considered representative of the CSP examiner population. The examiners had been trained in the administration of the TES and, for one month, had been conducting government examinations using the format. Examiners conducted two practice examinations before conducting an examination for the project. Each examiner completed two 4-hour examinations (morning and afternoon) on four days and one 4-hour examination on one day for a total of nine examinations each. The examiners were not given any information regarding the base rates. They did not receive feedback regarding the accuracy of their decisions until the end of the study, and they were blind as to whether the examinee was PG.

Apparatus

The examiners used standard analog field polygraphs manufactured by either Lafayette or Stoelting. Standard respiratory, electrodermal, and cardiovascular responses were recorded. The electrodermal component was operated in the manual mode. The examinations were conducted individually in large (6.2m x 6.2m) rooms in a building located on Fort McClellan. The scenarios used to program examinees guilty were enacted in another building located approximately two miles from the examination building. There were no video recording devices nor one-way mirrors in the examination rooms. The examinations were audio taped.

Scenarios

The PG examinees enacted one of four mock scenarios. Each scenario was representative of one of the four relevant questions. The espionage scenario required one examinee to steal a classified document from an office and to give the document to a second examinee. The second examinee received the document and placed it inside a vehicle located in the parking lot. Examinees who enacted the sabotage scenario, stole either a classified document or a classified computer disk. The examinee either put the document through a paper shredder or with a pair of scissors, cut the disk into pieces. An examinee who enacted the unauthorized contact scenario was asked to meet with a German agent who was sitting in a car in the parking lot. The agent requested that the examinee obtain some classified information to be given to the agent at a later time. During the enactment of the unauthorized disclosure scenario, the scenario setter was called out of his office midway through briefing the examinee regarding some classified computer information. A third person, who appeared to be fixing a window screen, entered the office and engaged the examinee in conversation regarding what the examinee had been told. All PG examinees received $100.00 as payment for their participation in the "crime." In addition, all PG examinees wrote a statement indicating that "for the purposes of this project" they had engaged in espionage, sabotage, unauthorized contact, or unauthorized disclosure, depending on which scenario they enacted.

Scoring and Decision Criteria

Scoring procedures developed during a previous study (DoDPI Research Division Staff, 1997) were used to evaluate the data. If the original decision was conclusive--significant responding (SR) or no significant responding (NSR)--then the decision was final. If a conclusive decision could not be made then the physiological responses to the first presentation of each relevant question were reevaluated by comparing them to the physiological responses only to the first presentation of the second comparison question. If, after the rescore, a conclusive decision was not possible, then the test was considered inconclusive (INC).

Procedures

During each session, ten examinees were given information regarding the research project, their participation, and the PDD examination. If they agreed to participate, they signed a consent form indicating that they were voluntarily participating in the research project. The examinees were taken in groups of two either to another building to be programmed guilty, or to the testing site. The PG examinees received information regarding the purpose of the scenario they were to participate in and signed an additional consent form indicating that they agreed to participate in the scenario. After

they enacted one of the scenarios, they were transported to the testing site. The transportation of the examinees to the testing site was timed so the examiners were not able to discern which examinees were innocent and which were programmed guilty.

The examinations were conducted according to guidelines provided to the examiners. Each examiner provided a numeric score and a decision (SR, INC, NSR) based on the numeric score, for each test. An NSR decision concluded the subtest. If the decision was INC, the examiner briefly discussed the questions with the examinee to determine if the examinee understood the questions. Then, the test was administered again. If, based on the data from the second test, the examiner's decision was INC, then the decision for that subtest was INC. When the examiner rendered an SR decision, the examiner confronted the examinee with the results.

Programmed guilty examinees were instructed to confess their guilt if they were confronted by the examiner, but not to reveal any details of their activities. Once a PG examinee confessed, the examination was concluded. However, an innocent examinee who responded significantly to the relevant questions--a false positive (FP) decision--was questioned by the examiner to determine if there was a legitimate real-world explanation for the examinee's physiological responses to the relevant questions. The examiner recorded any information provided by the examinee and concluded the examination. Two examiners, otherwise not involved with the study, independently evaluated the information obtained from the examinees who received FP decisions. If the two examiners agreed that the information was significant enough to justify the examinee's physiological responding--a false positive decision with justification (FPWJ)--then that examinee's data were not included in the original data analyses.

If the decision for the first subtest was either NSR or INC, the examiner conducted the second subtest. If, however, the decision for the first subtest was SR, then the second subtest was not conducted. All of the examinees tested during a session were debriefed simultaneously. Examinees who participated in mock scenarios returned the $100.00.

Data reduction and analyses

The data from 82 examinees were included in the analyses. The remaining six examinees were excluded for the following reasons: One PG examinee confessed to the examiner prior to the examination, one PG examinee was unable to understand the instructions of the scenario setter, two examinations were incomplete, and two FPWJ examinees were excluded.

If the scoring based on the physiological responding during an initial test resulted in an INC decision and a second test was conducted, unless otherwise indicated, only the result of the second test was included in the analyses. The percentages of innocent examinees and PG examinees correctly identified were calculated. Chi-square tests were conducted to determine if the numbers of correct decisions in identifying innocent and PG examinees were significantly different from chance. The significance criterion was set at .05.

Results

Excluding the one inconclusive decision, 98% of the innocent examinees and 83.3% of the PG examinees were correctly identified. The number of correct decisions, inconclusive decisions and errors made by the examiners are presented in Table 1. Both the number of innocent examinees correctly identified [X^2 (1, N = 51) = 47.08, p < .001] and the number of PG examinees correctly identified [X^2 (1, N = 30) = 13.33, p < .001] were significantly greater than chance. When the two FPWJ examinees are included in the analysis of the accuracy of decisions identifying innocent examinees, 94.3% of the innocent examinees were correctly identified. Including the FPWJs, the number of innocent examinees correctly identified was significantly greater than chance [X^2 (1, N = 53) = 41.68, p < .001]. There were a total of five (10%) innocent examinees and five (17%) PG examinees for whom the initial test results were inconclusive. After retesting, the results remained inconclusive for only one innocent examinee (1.9%).

Table 1

Number of Correct Decisions, Inconclusive (INC) Decisions, and Errors Made by the Examiners in Identifying Programmed Guilty and Innocent Examinees

	Decisions		
Role	Correct	INC	Errors
Innocent	50*	1	1
Guilty	25*	0	5

Note: Analyses tested whether each distribution was significantly different from change.
* p < .001.

Discussion

Excluding the one inconclusive decision, 98% of the innocent examinees and 83.3% of the PG examinees were correctly identified. These results mirror the findings of the previous study, with respect to the accuracy of decisions obtained using the TES format. Although many questions remain regarding the generalizability of the TES format to field situations, the TES format appears to have greater validity than the format currently used by the federal government.

Further testing is required to answer some of the questions raised by the current and the previous studies (DoDPI Research Division Staff, 1997): (a) does the caveat "during this project" affect the accuracy of the decisions identifying innocent and PG examinees, (b) does the effect of the question caveat impact on the generalizability of the format to field situations, (c) does the reduced number of relevant issues addressed during the test contribute to the increase in the accuracy of identifying PG examinees, and (d) will the results generalize when different issues are addressed and different relevant questions are utilized.

Reference

Department of Defense Polygraph Institute Research Division Staff (1997). A Comparison of Psychophysiological Detection of Deception Accuracy Rates Obtained Using the Counterintelligence Scope Polygraph (CSP) and the Test for Espionage and Sabotage(TES) Question Formats. *Polygraph*, 26(2), 79-106.

* * * * * *