

Polygraph

VOLUME 28

1999

NUMBER 3

Contents

Reprint Effect of Posthypnotic Suggestions on the Accuracy of Preemployment Polygraph Testing Howard W. Timm	193
An Objective Method for Manually Scoring Polygraph Data Donald Krapohl & Barry McManus	209
A Response to Honts on the Issue of the Discussion of Questions Between Charts Stan Abrams	223
Fiscal Year 1998 Report to Congress on the Department of Defense Polygraph Program DoDPI	229
The Relative Utility of the Forensic Disciplines Gary D. Light & John R. Schwartz	240
Memory, Reason, Emotion, Anatomy and the Polygraph. The Mind-Body Link Offered by Neurobiology and Related Sciences William B. Anderson	259
History and Development of Polygraph Examinations in Bulgaria Svetoslav Zanev	263

Published Quarterly

© American Polygraph Association, 1999

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

Effect of Posthypnotic Suggestions on the Accuracy of Preemployment Polygraph Testing

Howard W. Timm¹

Abstract

The efficacy of both a posthypnotic polygraph countermeasure suggestion and a posthypnotic ideomotor lie detection suggestion were examined within an analog preemployment screening context. Forty-five subjects were randomly assigned to one of three equal-sized groups, controlling for their performance on the Harvard Group Scale of Hypnotic Susceptibility: Form A. Subjects assigned to one of the groups received the posthypnotic suggestion that the subject's index finger would rise whenever he or she lied during the polygraph test; subjects in another group were given a posthypnotic countermeasure designed to help them appear innocent whenever they lied; and those in the third group were not given either hypnotic intervention. None of the subjects given the ideomotor suggestion raised his or her finger when he or she lied during the polygraph test. The countermeasure suggestion also was ineffective, as was demonstrated by its failure to produce significantly more false negative responses.

Keywords: hypnosis, lie detection, polygraph, posthypnotic suggestions, psychiatry

Since the passage of the 1988 Employee Polygraph Protection Act, the use of polygraph preemployment screening is relatively rare in the United States. However, it is still used fairly frequently for certain types of "sensitive" positions (for example, law enforcement, intelligence, and certain security and drug-handling positions). In contrast to what one might expect, given the importance of screening out undesirable applicants within those spheres, comparatively little is known about the validity of this type of polygraph testing (OTA-TM-H-15). In addition to urging that further research be conducted to assess that procedure's validity, the U.S. Congress Office of Technology Assessment has recommended that studies also be conducted to determine how susceptible preemployment polygraph screening is to certain countermeasures.

This study examines the efficacy of both a posthypnotic polygraph countermeasure suggestion and a posthypnotic ideomotor lie detection suggestion within an analog preemployment screening context. Although there is a dearth of information concerning the efficacy of hypnosis as a polygraph countermeasure, few studies have been conducted. However, it may be important to note certain limitations associated with that body of research.

All the findings reported in the literature were based on small sample sizes, many of which were largely anecdotal, with a single subject. In several of these studies, it was apparent that the polygraph examiner knew the ground truth at the time he was conducting the test, which was further compounded by the use of subjective methods

Reprinted, with permission, from the Annual Book of ASTM Standards, copyright American Society for Testing and Materials, 100 Barr Harbor Drive, West Conshohocken, PA 19428.

This article is based on a paper presented at the Tenth International Congress of Hypnosis and Psychosomatic Medicine in Toronto, Canada, in 1985. The views expressed herein are those of the author and do not reflect the official policy or position of the U.S. Department of Defense or the U.S. Government. The research was conducted when the author was affiliated with the Center for the Study of Crime, Delinquency, and Corrections at Southern Illinois University at Carbondale, IL. Received for publication 1 Nov 1990; revised manuscript received 20 Feb 1991; accepted for publication 26 Feb 1991.

¹Program Manager, Defense Personnel Security Research and Education Center, Monterey, CA.

to analyze the polygraph charts. In all but one of them (Weinstein, Abrams & Gibbons, 1970), the only countermeasure used was hypnotically induced amnesia. While each of the studies reported using good hypnotic subjects, the hypnotic susceptibility of their subjects was not measured using a standardized hypnosis scale, nor were the effects of hypnosis countermeasures on a more general population of subjects examined. In addition, all of the testing situations involved either a mock crime or an information detection paradigm. Therefore, the results and conclusions reported in the following studies should be viewed taking into account those limitations.

One of the better hypnosis countermeasure studies was conducted by Germann (1961). His study involved five subjects who each attempted to deceive an examiner about names, places, and playing cards under three different conditions: (1) normal waking state, (2) hypnotized with suggested amnesia, and (3) waking state with posthypnotic suggestion for amnesia. In eight examinations, significant responses to the critical items (the questions on which the subjects lied) were found; in seven of the examinations, the results were inconclusive. Germann interpreted his results as supporting the hypotheses that hypnotically induced amnesia is not effective as a countermeasure; however, the relatively large number of inconclusives suggests that the hypnosis may, in fact, have decreased the detectability of the lies (Weinstein et al., 1970; Barland & Raskin, 1973).

Cumley (1959) reported a similar study involving two subjects who had been previously hypnotized on other occasions by the hypnotist involved in the experiment. The hypnotist gave the subjects a posthypnotic suggestion that they would be unable to remember the details of a mock crime. Cumley reported that guilt patterns were plainly present on the charts of both subjects; however, his report also indicated that the examiners knew the ground truth before the polygraph testing.

Tocchio (1963) reported the results of a single-subject hypnosis countermeasure experiment. In that study, a female committed a mock crime and then was given a

posthypnotic suggestion that "she would neither remember the offense nor the period of time in which it took place." Tocchio reported that the subject's polygraph charts indicated that she was guilty of the offense.

Bitterman and Marcuse (1945) conducted another single-subject study, which also indicated that hypnotically induced amnesia is not an effective countermeasure. A female subject was selected for her ability to enter "deep trance" based on the Friedlander-Sarbin scale (1938). The subject was told a word under hypnosis, then told she would not remember it, no matter how hard she tried. The subject was later given a polygraph test, during which one of two examiners attempted to detect the word presented. The procedure was repeated with the same subject a total of eight times (four exams were administered by each examiner). During two of the eight trials, the subject was given the word in the waking state. On four of the six trials involving the hypnotically induced amnesia, the examiners correctly detected the word; on the other two, they chose it as their second choice.

Weinstein et al (1970) conducted a study in which they examined the effect of hypnotically induced repression and guilt. They selected six college students on the basis of their ability to enter deep hypnotic states. The subjects were divided into two groups. The three members of the first group were told to enter an office and take one of three bills (\$1, \$5, or \$20). Then, they were hypnotized and told that they would not recall taking the money. The second group of students did not take any money; however, they were told under hypnosis that they had stolen one of the bills and that they would experience considerable guilt because of this. The examiner was completely misled by the three innocent students. In fact, he stated with certainty that each had taken the hypnotically suggested amount. The examiner was only partly convinced that the members of the guilty group had taken the money and identified the correct amount taken for only one of them.

Barland & Raskin (1973) mention three additional experiments, which also examined the efficacy of hypnotically induced amnesia as a countermeasure. In one of the experiments, conducted by William J. Bryan in

Miami in 1965, the examiners were able to interpret the charts correctly in spite of the subject's amnesia for the mock crime. However, in an experiment by William Mayer and Lt. Col. Joseph Ziglinski in Washington, DC, also with a single subject, hypnotically suggested amnesia was successful in misleading the examiner. Another informal experiment reported by Barland and Raskin in 1960 was conducted by the Lie Detector Committee of the U.S. Army Military Police School (Barland & Raskin, 1973), which apparently concluded that hypnotically suggested amnesia could be an effective countermeasure. However, Matte's (1980) description of an earlier draft of that report indicated that the procedure was rarely effective.

In addition to the possibility of using hypnosis as a polygraph countermeasure, certain people have advocated its use for detecting deception. For example, the use of ideomotor hypnotic questioning has been advocated in certain therapeutic situations involving clients who are resistant to providing information about their cases (Cheek, 1975; Cheek, 1976; Erickson & Rossi, 1975; Spithill, 1974; Udolf, 1981). Arons (1967) also advocates the use of the same technique as a form of "lie detector" in criminal cases and has trained law enforcement personnel to perform that technique.

Although the literature reflects the current applications of ideomotor questioning, it contains only one, relatively poorly designed, single-subject study (Tocchio, 1963), which has attempted to assess the validity of that technique. In that study, a secretary who was participating in a demonstration for television, complied with a posthypnotic suggestion that she scratch her ear whenever she lied about the details of a mock crime.

While there is little research on the extent to which hypnotized subjects will conform to suggestions that they make incriminating admissions or gestures, a few studies have examined subject compliance with antisocial or self-destructive suggestions in different contexts. In laboratory research, the following types of seemingly inappropriate behavior were reported after hypnotized subjects were directed to perform those acts:

throwing what they were told was acid in a person's face and picking up what appeared to be a poisonous snake (Rowland, 1939; Young, 1952); stealing, looking through a stranger's purse, verbalizing sexual fantasies in public (Brenman, 1942); stealing examination questions (Coe, Kobayashi & Howard, 1972); indecent exposure in public (Kline, 1958); and, for military personnel, giving out military secrets and physically attacking superior officers or friends (Watkins, 1947; Watkins, 1972).

In real life, the following acts also have been alleged to have been committed by individuals who were reported to have been given hypnotic suggestions to perform them: criminal acts followed by shooting oneself (Hammerschlag, 1957); committing a bank robbery in which innocent people were killed (Reiter, 1958); and heterosexual, pedophilia, and homosexual seductions (Kline, 1972; Perry, 1979; Tinterow, 1970). While it is distressing that individuals complied with the requests to perform the aforementioned acts, it is important to consider the possible role of several other factors that may have contributed to those actions: (1) a close relationship present or desired between the subject and the hypnotist; (2) use of hypnosis to deny personal responsibility for voluntary acts; (3) feeling obligated to help the researcher; (4) desiring to perform as a "good subject"; (5) believing no one would actually be harmed, since the activity was part of an experiment being conducted by presumably responsible researchers; and (6) believing that those acts committed were not really that objectionable to the subject (Coe et al, 1972; Barber, 1961; Barber, 1969; Conn, 1972; Conn, 1981; Levitt & Baker, 1983; Orne, 1959; Orne, 1961; Orne, 1972; Orne, 1965; Timm, 1983). Thus, while suggestions given under hypnosis may seem to be the principle cause of certain behaviors, the role of situational variables similar to those responsible for subject compliance in Milgram's studies on obedience must also be considered.

The present study is intended to examine the efficacy of both a posthypnotic polygraph countermeasure suggestion and a posthypnotic ideomotor lie detection suggestion within an analog preemployment screening context. As in almost all studies of detection of deception, the design employed

reflects a trade-off between the certainty with which ground truth is established and the study's external validity. The author hopes that the mix associated with the design chosen will be able to generate information which is of value in resolving the practical and theoretical questions addressed by this study.

Method

Subjects

The subjects consisted of 45 volunteers enrolled in selected criminal justice classes at a large Midwestern university. The courses from which the subjects were drawn were limited to those whose instructors agreed to permit their students to participate in the experiment for extra credit. To maintain consistency, the extra credit was standardized for all classes, at the fixed level of 3% of the total points for each class.

Prior to volunteering for the experiment, the subjects were informed of the purpose and design of the study. The subjects included 30 males and 15 females. Their ages ranged from 18 to 37 [mean (M) = 22.13; standard deviation (SD) = 3.109].

Apparatus

A Stoelting field polygraph (Model 122656) was used to record both the respiration and skin resistance response (SRR) of the subject. Respiration was recorded using a pneumatic tube positioned around the subject's thoracic area. The SRR was recorded from two stainless steel electrodes attached to the volar surface of the first and third fingers of the subject's right hand. All SRR recordings were made with the instrument in the automatic centering mode.

The instrument used to score respiration objectively was a Tektronix digitizer (Model 4662) interfaced with a Tektronix microcomputer (Model 4051) programmed to measure the curvilinear distances between points on a sheet of paper. The subjects' left hands were videotaped using a Panasonic color video camera (Model WV-3110), connected to a Panasonic portable video cassette recorder (Model NV-8410), using ½-inch TDK brand super-avilyn high-output, high-resolution video recording tape.

Initial Testing

The first phase of this study involved the subjects completing the Harvard Group Scale of Hypnotic Susceptibility: Form A (HGSHS:A), selecting their subject number, and completing a questionnaire which addressed their prior criminal involvement. Fifteen misconduct questions were included on the questionnaire (for example, "Have you ever used LSD?") Next to each question the subjects were asked to identify both their degree of prior involvement in the matter addressed and their degree of concern that their being asked that question on a pre-employment polygraph test might jeopardize their ability to secure employment with the criminal justice agencies and departments to which they were planning to apply.

Subjects were asked to indicate their degree of concern and involvement using two different five-point rating scales. The degree of concern options ranged from "not concerned at all about responding to that question on a polygraph test" (Point 1) to "I am extremely concerned about responding to that question on a polygraph test" (Point 5). The subjects' degree of involvement options consisted of the following: (1) never did it and never knew any friends, relatives, or acquaintances who did it; (2) never did it personally, but others I know have done it; (3) only did it once; (4) have done it, but only a few times; and (5) have frequently engaged in that activity.

The first phase of the experiment was conducted in a large (80-seat) auditorium on campus. Volunteers were given the opportunity to select one of three consecutive nights to complete that phase. They were required to be at the auditorium at the designated starting time and forewarned that, if they were late, they would not be permitted to enter the room or to participate in the study that evening.

Five minutes after the designated time, the auditorium door was locked to prevent late arrivals from entering the room and disturbing the other participants. The subjects were informed of the tasks to be completed that evening, told of the steps that would be taken to preserve their confidentiality throughout the experiment, and asked to fill out the forms they would be given later as honestly and completely as possible. Subjects were then

told to disperse throughout the classroom in such as fashion that none of the subjects would be able to see what the others had written. Then, a box containing slips of paper with different subject numbers on them, ranging from 1 to 60, was passed around the room, and the subjects were told to select any one they desired. Next, each of the subjects was given an HGSHS:A test booklet and a copy of the questionnaire. Each was told to write his or her subject number on them, to put the slip of paper containing the subject number in his or her wallet, and to be certain not to lose it. The subjects were warned not to write their names on any of the documents, except for a card indicating the class to which they wanted the extra credit applied, and they were warned not to write their subject numbers on those cards.

The HGSHS:A was administered first. To help standardize that procedure, an audio tape recording of that test was used, which conformed to the specifications contained in the HGSHS:A manual prepared by Shor and Orne (1962). The same tape was played during each of the three evening sessions. After the subjects finished the HGSGS:A, they completed their misconduct questionnaires, then turned in those materials by placing them in a larger envelope, which was located on a desk away from where the investigator was standing.

Hypnotic Treatment Conditions

During the second phase of the experiment, 15 subjects were randomly assigned to each of three groups, controlling for their levels of hypnotic susceptibility. They were either assigned to the countermeasure condition, the ideomotor condition, or to the control group. A one-way analysis of variance (ANOVA) comparing the HGSHS:A means for the three groups indicated that none of those means significantly differed from the others [$F(2, 42) = 0.25$; $P = 0.778$].

Subjects from all three conditions were scheduled to meet with the investigator in his office. After they had been greeted, the subjects were informed of the polygraph procedure that would be followed during the next stage, the questions they would be asked, and the scoring procedure that would be employed. Any questions they had regarding

those procedures were answered, and the subjects made their appointments for taking the polygraph test. Subjects assigned to the control group were then thanked for meeting with the investigator and dismissed.

Those assigned to both the countermeasure condition and the ideomotor condition were then informed that they would be hypnotized and were given a brief explanation about hypnosis and the procedure that would be used. After answering the subjects' questions concerning those matters, the investigator administered an eye-fixation hypnotic induction, followed by a deepening suggestion and a hand-closure suggestion. [For a description of each of these procedures see Weitzenhoffer (1957)]

Subjects assigned to the countermeasure group were then given the following additional suggestions:

You have seen how by using your imagination, you could make your hands be drawn together. By using your imagination you will also be able to prepare yourself to pass the polygraph test. You will be prepared, so you will pass the test. Let's imagine together what will happen the day you take the polygraph test.

(Polygraph procedure reviewed)

You will hear a voice that sounds like mine on the tape. It might seem like it takes some time for the first question to be asked. Then you will hear the voice on the tape say, "Are you presently attending _____ University?" When you hear that question you will say yes, but when you do, you will feel nervous, concerned and guilty that you did say yes. You will feel your breathing become shallower, your hands perspire, and a tense feeling in your stomach. In fact, each time you say yes during the polygraph test you will feel that same way. You will feel nervous, concerned and guilty about saying yes; you will feel your breathing become shallower, your hands perspire, and a tense feeling in your stomach. That feeling will stop as soon as you hear the next question.

That feeling will stop as soon as you hear the next question. Whenever you say no, regardless of whether or not your statement is true, you will feel calm and relaxed, just as you do now. You will breathe nice and deeply and feel cool and relaxed, just as you feel now.

I am going to ask you a few questions now that will be on the polygraph test and I want you to respond like you will when you take the test and then to tell me how you feel.

(Four questions reviewed, the first question, a relevant question, the control question, then another relevant question. After each question the subject is queried about feelings, desired responses are reinforced by saying, "good, that is just the way you will feel when you actually take the polygraph test.")

I am going to count to three now. When I do, you will wake up. Even though you will be completely awake, you will remember what we said, and when you close your eyes to take the polygraph test, you will have the same feelings. You will feel tense, nervous, and guilty each time you say yes. This will last until you hear the next question. When you say yes, you will have a tense feeling in your stomach, you will breathe shallowly, and your hands will perspire. However, whenever you say no, you will always feel relaxed and calm. I'm going to count to three now; when I do you will feel wide awake and very good. I'm counting to three now ...1...2...3.

Subjects assigned to the ideomotor condition were given the following suggestions:

Just as your hands have been drawn together you will feel your index finger on your left hand rise up each time you tell a lie during the polygraph test. It will seem that the harder you try to keep it down, the more it is drawn upward. It will seem that the harder you try to keep it down, the more it is drawn upward. I am going to ask you a question now, and I want you to lie. I am going to ask you a question now, and

I want you to tell me a lie. When you do lie, your index finger on your left hand, your left hand, will be drawn up. Here is the question; I want you to respond no and watch what happens. Is today (*the correct day of the week*)?

(If the subject's finger did not begin to rise after 10 s, he or she was told "It's starting to creep up, the harder you try to keep it from rising the more it wants to rise. The harder you fight to keep it from rising the more it wants to rise.")

I want you to lie also to this question and watch how automatically the finger rises this time. Is it the month of (*correct month*)? Good. The harder you try to keep it down each time you lie, the stronger it is drawn upward. When you take the polygraph test, each time you tell a lie your index finger on your left hand will automatically rise. When you tell the truth, it will stay down. When you lie, it will go up. When you tell the truth, it will stay down.

I'm going to wake you up now. I'm going to count to three and clap my hands. When I do, you will wake up feeling wide awake and refreshed. I am going to count to three now and clap my hands. When I do, you will be wide awake and refreshed; however, when you take the polygraph test, your index finger on your left hand will rise each time you lie. I'm counting to three now ...1...2...3.

Polygraph Testing

Within three days after receiving their respective treatments, the subjects in all three groups reported individually to the polygraph testing room. The subjects met with the polygraph examiner, who was unaware of the treatment conditions or the subjects' responses on the misconduct questionnaire.

After the instructor had asked to see the subject's identification number, the subject was given a description of the equipment and the procedure that was to be used. None of the subjects was asked his or her name, and only the subject number was used for identification purposes. The subjects were shown the questions and reminded that the

examiner's task was to "weed out unsuitable candidates applying for a position at a conservative criminal justice department, which feels that any applicant who either admits to or has a deceptive response to any of the infraction questions under investigation is unsuitable for the position." They were also reminded that their role was to pass the preemployment screening procedures and to be offered the position, regardless of their extent of previous misconduct. Subjects were also reminded that they should say yes only to those questions identified by the examiner (that is, the four irrelevant questions and one control question).

The subjects were then shown the videotaping equipment and informed that it was there to record whether certain subjects will subconsciously raise the index finger on their left hands each time they lie. After the polygraph attachments were connected to the subject, the camera zoomed in on the subject's left hand and focused. The examiner stated out loud the subject's identification number after starting the video recorder.

Next the polygraph test was administered, which consisted of the following questions:

1. Are you presently attending _____ University?
2. Have you ever used heroin?
3. Have you ever smoked marijuana?
4. During the last three years, have you stolen merchandise from a retail store?
5. Have you ever taken LSD?
6. Are you enrolled in at least one Administration of Justice class this term?
7. In the last three years, have you driven while you were clearly drunk or under the influence of drugs?
8. Do you smoke marijuana on a fairly regular basis?
9. Have you ever sold drugs?
10. In the last three years, have you ever participated in a fistfight?
11. Do you presently reside somewhere in _____?
12. Have you ever purchased an item you thought was stolen?

13. In the last three years, have you been fired from a job?
14. Have you ever had your driver's license revoked or suspended?
15. Have you ever used cocaine?
16. Are you currently in _____ Hall?
17. Have you ever illegally used "speed" or another type of amphetamine?
18. Have you ever attempted to elude a police officer, either while on foot or in an automobile?
19. In the last three years, have you stolen something valued at more than \$10 from an employer?
20. Are you worried that I might ask you some additional questions that you were not told about?

After Questions 1 through 20 were administered, the order was reversed and Questions 19 through 1 were repeated, followed by Question 20. To increase the standardization of the questioning procedure, the questions were tape recorded and asked at 20-s intervals. Subjects were told to keep their eyes closed during the testing.

After the polygraph test was over, the video recorder was stopped and the polygraph attachments were removed. The subjects were thanked and informed that they would be told later in the term how they did. No subjects were permitted to see their charts or to find out how well they did, since their feedback to other volunteers might have contaminated the study.

At the end of the term, the subjects were given the opportunity to find out which questions they appeared to have their greatest responses to, based upon a cursory examination of their charts. For each subject, a sheet containing his or her results and a letter of thanks for participating was placed in a sealed envelope with the subject's number written on the outside. Those envelopes were left with a receptionist, who was instructed to give them their feedback envelopes after they had shown her their subject number slips.

Before giving the subjects their envelopes, however, she gave each one another letter requesting that, prior to receiving his or her feedback, he or she complete the same

misconduct questionnaire taken earlier and either place it in the investigator's mail box or send it to him using campus mail. They were informed that the reason for requesting them to complete the questionnaire a second time was that they were now certain that their responses would be confidential and that they had had more time to think about the questions.

Objective Scoring Procedures

The charts were analyzed by objectively scoring respiration, SRR amplitude, and SRR maximum height. With the field polygraph used, a rising SRR pattern on the polygraph chart indicated less electrical resistance, suggesting an emotional or cognitive reaction. In order to score both the respiration and the SRR responses, it was necessary to correct for tangent errors, which resulted from the use of fixed-length pivoting polygraph pens. This was accomplished by making a tracing of the semicircle path of travel of the polygraph pen when the chart was stationary. This tracing was then placed over the polygraph chart and aligned with each question marker tick at the top of the chart. A line was then drawn intersecting the points on the SRR and respiration patterns where the constructed tangent error templates crossed them.

Respiration patterns were scored by measuring the curvilinear length of the pattern recorded by the polygraph respiration pen, beginning when each question was asked and ending 15s later. The patterns associated with the 40 questions asked during the polygraph test were traced using a Tektronix digitizer.

The SRR amplitude was scored by measuring the vertical rise of the largest wave occurring between the onset of the stimulus question and 15s later. The length of the vertical rise was measured from its lowest point before the wave began a positive slope to the highest point it reached within the 15s period. When no positive SRR rise on the chart occurred during the 15s period, a value of 0.1 mm was recorded, otherwise the actual values in mm were recorded.

The method used to analyze the charts is referred to as the adjacent question comparison method. The dependent measurement

values for the question immediately preceding and immediately following each relevant question were added together. That total was divided by two and compared with the value associated with the response in the middle. If the mean for the adjacent SRR amplitude values was lower than or equal to 0.5 times the value of the SRR response in the middle, the middle question was given an SRR amplitude score of 3. If it was greater than 0.5, but less than or equal to 0.7, it was given a score of 2; if it was greater than 0.7, but less than or equal to 1, it was given the score of 1. All values over 1 were scored at 0. Thus the larger an SRR amplitude response was in comparison with those immediately surrounding it, the higher it was scored. Respiration values were compared in a similar manner; however, different cutoff points were selected, and lower values, instead of higher values, were considered indicative of deception (Timm, 1982; Timm, 1982a). If the mean of the adjacent respiration values was greater than or equal to 1.15 times the value of the respiration response in the middle, the middle question was given a respiration score of 3. If it was less than 1.15, but greater than or equal to 1.05, it was given a score of 2; if it was less than 1.05, but greater than or equal to 1, it was given the score of 1. All values under 1 were scored as 0. Therefore, the smaller a respiration response was in comparison with those immediately surrounding it, the higher it was scored.

As previously noted, each question was asked twice during the polygraph test. The two SRR amplitude values for each question were added together, as were the two respiration values. If the sum of those two values was four or greater, it was classified as indicating possible guilt for that dependent measure under the low exclusion criterion; values of only three or greater were classified as suggesting deception using the high exclusion criterion. Thus, more applicants would be rejected in a real testing situation using the high exclusion criterion, than would be rejected under the low exclusion criterion.

In addition to analyzing the polygraph charts, the videotapes were also viewed and scored. An assistant, who worked independent of both the investigator and the polygraph examiner, and who was unaware of both the

subjects' treatment groups and their responses on the questionnaire, viewed the tapes. That assistant was instructed to note any noticeable rise of the subjects' index fingers, as well as the questions on which those rises occurred.

Results

The assistant who viewed the tapes reported that none of the subjects' index fingers rose enough for him to detect any elevation during the testing. Therefore, the ideomotor treatment appears to have been unsuccessful with respect to producing the suggested effect during the questioning.

The number of admissions made by subjects on the questionnaires ranged from 0

to 10. The percentage of subjects responding affirmatively to each offense question contained on that instrument is presented in Table 1. The question producing the highest number of affirmative responses (82.2%) was the question asking whether they had ever smoke marijuana; while the question yielding the lowest number of affirmative responses (0%) was the question asking whether they had ever used heroin. The mean number of admissions made by the subjects in each group was 4.80, 4.93, and 4.40 for Groups 1, 2, and 3, respectively [$F(42, 2) = 0.125$; $P = 0.88$]. Therefore, it appears that the number of admissions made by the subjects was consistent across all three groups and that it should not be a major factor affecting the outcome of other analyses.

Table 1.
Percentage of subjects responding affirmatively to the offense questions.

% Responding Affirmatively	Offense Question
00.0	Have you ever used heroin?
82.2	Have you ever smoked marijuana?
20.0	During the last three years, have you stolen merchandise from a retail store?
13.3	Have you ever taken LSD?
60.0	In the last three years, have you driven while you were clearly drunk or under the influence of drugs?
55.6	Do you smoke marijuana on a fairly regular basis?
28.9	Have you ever sold drugs?
33.3	In the last three years, have you participated in a fist fight?
28.9	Have you ever purchased an item you thought was stolen?
11.1	In the last three years, have you been fired from a job?
8.9	Have you ever had your driver's license revoked or suspended?
33.3	Have you ever used cocaine?
51.1	Have you ever illegally used "speed" or another type of amphetamine?
26.7	Have you ever attempted to elude a police officer, either while on foot or in an automobile?
18.2	In the last three years, have you stolen something valued at more than \$10 from an employer?

The number of questions correctly and incorrectly classified using the objective scoring method and two different exclusion levels is presented in Tables 2 and 3. None of the treatment conditions resulted in a significant effect on the mean number of false positive, false negative, true positive, or true negative responses attained by their respective subjects at the $P = 0.05$ level.

To help ascertain whether treatment effects were being masked by subjects with low hypnotic susceptibilities, four additional analyses were run using a field-oriented procedure excluding subjects who scored five or under on the HGSHS:A. None of those calculations yielded statistically significant differences either. However, the mean number of true positive responses using the low exclusion criterion was 0.33 for those in the countermeasure condition ($n = 9$), 0.56 for those in the control group ($n = 9$), and 1.0 for those in the ideomotor condition ($n = 7$) [$F(2,22) = 1.9$; $P = 0.17$]. Although an order of

those means and their magnitude conform to what one might expect given the treatments, both the probability value attained and the number of tests conducted support the notion that those perceptible differences may have occurred simply by chance. The exclusion criteria affected the types of errors produced in the expected manner. The high exclusion criteria raised the level of true positives and false negatives at the expense of true negatives and false positives, while the opposite relationship held for the low exclusion criteria.

The type of error that occurred was also affected, as one might expect, by the number of admissions that were made. The mean number of responses correctly and incorrectly classified is presented in Tables 4 and 5. As the number of admissions increased, true positives and false negatives increased at the expense of true negatives and false positives, which is also perfectly logical.

Table 2.
Mean number of questions correctly and incorrectly classified using the adjacent question comparison method and low exclusion criteria.

Group	True Positive	True Negative	False Positive	False Negative
SRR Amplitude*				
1	1.0	9.0	1.2	3.8
2	1.1	9.1	0.9	3.8
3	0.5	9.0	1.6	3.9
$F(2, 42)$	1.5	0.00	1.8	0.00
Probability	0.22	1.0	0.17	1.0
Respiration*				
1	0.5	9.3	0.9	4.3
2	0.8	9.9	1.1	4.1
3	0.9	9.6	1.0	3.5
$F(2, 42)$	1.1	0.18	0.04	0.38
Probability	0.34	0.84	0.96	0.68

*SRR amplitude and respiration response scores of 4 or greater are classified as indicating possible guilt.

Table 3.
Mean number of questions correctly and incorrectly classified using the adjacent question comparison method and high exclusion criteria.

Group	True Positive	True Negative	False Positive	False Negative
SRR Amplitude*				
1	2.0	7.3	2.9	2.8
2	2.2	6.6	3.4	2.7
3	1.6	7.2	3.4	2.7
<i>F</i> (2, 42)	0.51	0.30	0.34	0.01
Probability	0.60	0.74	0.72	1.0
Respiration*				
1	1.5	8.1	2.1	3.3
2	1.6	7.7	2.3	3.3
3	1.3	8.7	1.9	3.1
<i>F</i> (2, 42)	0.16	0.42	0.18	0.05
Probability	0.85	0.66	0.84	0.95

*SRR amplitude and respiration response scores of 3 or greater are classified as indicating possible guilt.

Another matter of interest was whether there were differences in the level of concern associated between false positives and true negatives and between true positives and false negatives, which might help account for the nature of the subjects' responses. To examine this possibility, a series of *t* tests was conducted using the field-oriented scoring system and the high exclusion criterion. As expected, the mean concern level of the true positive responses [mean (*M*) = 2.96] was higher than the mean associated with the false negatives (*M* = 2.65) for SRR amplitude; however, it did not reach statistical significance at the *P* = 0.05 level (*t* = 1.59; *P* = 0.06). Given the fact that the number of cases was lowered to 30 because of missing cases (subjects who either had no true positives or no false negatives) and that the relationship conformed to what one might expect, this hypothesis bears further consideration. However, it should also be noted that there were no differences in the level of concern with respect to the SRR amplitude false positives

and true negatives (*M*_{fp} = 1.22; *M*_{tn} = 1.23; *N* = 39; *P* = 0.94).

Discussion

The failure of the posthypnotic suggestion given to subjects in the ideomotor condition to result in any of their fingers being elevated during the polygraph testing was somewhat of a surprise, given that virtually all of them raised their fingers during the treatment phase. This appears to support the notion that demand characteristics (Orne 1959; Orne 1961) play an important role in this type of situation, as perhaps do the potential consequences of their actions. It is also possible that more subjects in the ideomotor condition would have been detected if a strain gage device had been attached to the subjects' index fingers; however, it would probably be more effective for detecting deception if used to determine when the subjects were pressing down in order to be certain that their fingers did not rise.

Table 4.
Mean number of SRR amplitude responses correctly and incorrectly classified by the number of offenses committed.

No. of Offenses	<i>n</i>	True Positive	True Negative	False Positive	False Negative
High Exclusion^a					
0	4	0.0	12.3	2.8	0.0
1	3	0.7	11.3	2.7	0.3
2	6	0.3	11.5	1.5	1.7
3	3	1.3	9.3	2.7	1.7
4	6	1.5	9.0	2.0	2.5
5	6	1.3	8.8	1.2	3.7
6	3	1.0	7.3	1.7	5.0
7	5	1.8	6.0	1.8	5.2
8	4	1.7	5.5	1.5	6.2
9	1	5.0	4.0	2.0	4.0
10	4	1.5	4.2	0.8	8.5
Low Exclusion^b					
0	4	0.0	13.8	1.3	0.0
1	3	0.3	12.7	1.3	0.7
2	6	0.0	12.2	0.8	2.0
3	3	0.3	10.7	1.3	2.7
4	6	1.5	10.0	1.0	2.5
5	6	1.0	9.5	0.5	4.0
6	3	1.0	8.3	0.7	5.0
7	5	1.2	7.2	0.6	5.8
8	4	0.8	6.0	1.0	7.2
9	1	3.0	5.0	1.0	6.0
10	4	0.5	4.5	0.5	9.5

^aSums of mean response comparison and adjacent question comparison values of 6 or greater are scored as indicating possible guilt.

^bSums of mean response comparison and adjacent question comparison values of 7 or greater are scored as indicating possible guilt.

The results of the polygraph testing can be interpreted many different ways and undoubtedly will be shaped to some extent by one's ideological perspective. On the negative side, numerous false positives and negatives occurred. On the positive side, most of the responses were correctly classified. Also on a positive note, the scoring system could be adjusted to take into consideration the possible harm that might be cast upon a given organization and the number of good candidates applying for a position. Most people

would probably support using different preemployment selection criteria for screening soldiers for a position in a nuclear missile silo, where false negatives might destroy humanity, as opposed to screening inventors, where false positives might result in the loss of many valuable contributions.

Numerous factors have undoubtedly affected the internal and external validity of this study. The use of criminal justice majors as subjects appears to have made the study

Table 5.
Mean number of respiration responses correctly and incorrectly classified by the number of offenses committed.

No. of Offenses	<i>n</i>	True Positive	True Negative	False Positive	False Negative
High Exclusion^a					
0	4	0.0	13.8	1.3	0.0
1	3	0.3	10.3	3.7	0.7
2	6	0.2	11.5	1.5	1.8
3	3	0.7	10.3	1.7	2.3
4	6	1.7	9.5	1.5	2.3
5	6	1.3	7.8	2.2	3.7
6	3	2.3	7.3	1.7	3.7
7	5	1.8	6.0	1.8	5.8
8	4	1.3	6.0	1.0	6.7
9	1	1.0	6.0	0.0	8.0
10	4	2.8	4.5	0.5	7.2
Low Exclusion^b					
0	4	0.0	14.7	0.3	0.0
1	3	0.0	10.7	3.3	1.0
2	6	0.2	12.0	1.0	1.8
3	3	0.7	11.0	1.0	2.3
4	6	0.8	10.0	1.0	3.2
5	6	0.8	8.5	1.5	4.2
6	3	1.7	8.3	0.7	4.3
7	5	0.8	7.0	0.8	6.2
8	4	0.5	6.7	0.3	7.5
9	1	0.0	6.0	0.0	9.0
10	4	1.2	4.8	0.3	8.8

^aSums of mean response comparison and adjacent question comparison values of 6 or greater are scored as indicating possible guilt.

^bSums of mean response comparison and adjacent question comparison values of 7 or greater are scored as indicating possible guilt.

more realistic, because many of them will be required in the future to take preemployment polygraph tests for the positions they desire. Differences between this study and standard field practice with respect to the questioning procedure, choice of questions, scoring procedure, consequences of detections, and not using the questioning situation to solicit

admissions detract from the study's external validity. However, the author believes that the study provides the initial state for the continued refinement of the scoring procedures described and appears to have shed some light upon several important practical and theoretical issues.

References

- Arons, H. (1967). *Hypnosis in Criminal Investigation*. Power Publishing: South Orange, NJ.
- Barber, T.X. (1961). Antisocial and Criminal Acts Induced by Hypnosis: A Review of Experimental and Clinical Findings. *Archives of General Psychiatry*, 5, 301-312.
- Barber, T.X. (1969). *Hypnosis: A Scientific Approach*. Van Nostrand Reinhold, New York.
- Barland, G.H. & Raskin, D.C. (1973). Detection of Deception. *Electrodermal Activity in Psychological Research*, W.F. Prokasy & D.C. Raskin, Eds. Academic Press: New York.
- Bitterman, M.E. & Marcuse, F.L. (1945). Autonomic Response in Posthypnotic Amnesia. *Journal of Experimental Psychology*, 35, 248-252.
- Brenman, M. (1942). Experiments in the Hypnotic Production of Antisocial and Self-Injurious Behavior. *Psychiatry*, 5, 49-61.
- Cheek, D.B (1975). Maladjustment Patterns Apparently Related to Imprinting at Birth. *American Journal of Clinical Hypnosis*, 18, 75-82.
- Cheek, D.B. (1976). Short-Term Hypnotherapy for Frigidity Using Exploration of Early Life Attitudes. *American Journal of Clinical Hypnosis*, 19, 20-27.
- Coe, W.C., Kobayaski, K. & Howard, M.L. (1972). An Approach Toward Isolating Factors That Influence Antisocial Conduct in Hypnosis. *International Journal of Clinical and Experimental Hypnosis*, 20, 118-131.
- Conn, J.H. (1972). Is Hypnosis Really Dangerous? *International Journal of Clinical and Experimental Hypnosis*, 20, 61-79.
- Conn, J.H. (1981). The Myth of Coercion Through Hypnosis: A Brief Communication. *International Journal of Clinical and Experimental Hypnosis*, 29, 95-100.
- Cumley, W.E (1959). Hypnosis and the Polygraph. *Police*, 4, 39.
- Erickson, M.H. & Rossi, E.L. (1975). Varieties of Double Bind. *American Journal of Clinical Hypnosis*, 17, 143-157.
- Friendlander, J.W. & Sarbin, T.R. (1938). The Depth of Hypnosis. *Journal of Abnormal Social Psychology*, 33, 453-475.
- Germann, D.C. (1961). Hypnosis as Related to the Scientific Detection of Deception by Polygraph Examination: A Pilot Study. *International Journal of Clinical and Experimental Hypnosis*, 9, 309-311.
- Hammerschlag, H.E. (1957). *Hypnotism and Crime*. Wilshire Book Co.: Hollywood, CA.
- Kline, M.V. (1958). The Dynamics of Hypnotically Induced Antisocial Behavior. *Journal of Psychology*, 45, 239-245.
- Kline, M.V. (1972). The Production of Antisocial Behavior Through Hypnosis: New Clinical Data. *International Journal of Clinical and Experimental Hypnosis*, 20, 80-94.

- Levitt, E.E. & Baker, E.L. (1983). The Hypnotic Relationship—Another Look at Coercion, Compliance and Resistance: A Brief Communication. *International Journal of Clinical and Experimental Hypnosis*, 31, 125-131.
- Matte, J.A. (1980). *Art and Science of the Polygraph Technique*. Charles C. Thomas: Springfield, IL.
- Orne, M.T. (1959). The Nature of Hypnosis: Artifact and Essence. *Journal of Abnormal Social Psychology*, 58, 277-299.
- Orne, M.T. (1961). The Potential Uses of Hypnosis in Interrogation. *The Manipulation of Human Behavior*, A.D. Biderman and H. Zimmer, Eds., Wiley: New York, 169-215.
- Orne, M.T. & Evans, F.J. (1965). Social Control in the Psychological Experiment: Antisocial Behavior and Hypnosis. *Journal of Personality and Social Psychology*, 1, 189-200.
- Orne, M.T. (1972). Can a Hypnotized Subject be Compelled to Carry Out Otherwise Unacceptable Behavior? *International Journal of Clinical and Experimental Hypnosis*, 20, 101-117.
- Perry, C. (1979). Hypnotic Coercion and Compliance to It: A Review of Evidence Presented in a Legal Case. *International Journal of Clinical and Experimental Hypnosis*, 27, 187-218.
- Reiter, P.J. (1958). *Antisocial or Criminal Acts and Hypnosis: A Case Study*. Charles C. Thomas: Springfield, IL.
- Rowland, L.W. (1939). Will Hypnotized Persons Try to Harm Themselves or Others? *Journal of Abnormal and Social Psychology*, 34, 114-117.
- Shor, R.E. & Orne E. C. (1962). *Harvard Group Scale of Hypnotic Susceptibility*. Consulting Psychologists Press: Palo Alto, CA.
- Spithill, A.C. (1974). Treatment of a Monosymptomatic Tie by Hypnosis: A Case Study. *American Journal of Clinical Hypnosis*, 17, 88-93.
- Timm, H.W. (1982). Analyzing Deception from Respiration Patterns. *Journal of Police Science and Administration*, 10, 47-51.
- Timm, H.W. (1982a). Effect of Altered Outcome Expectancies Stemming from Placebo and Feedback Treatments on the Validity of the Guilty Knowledge Technique. *Journal of Applied Psychology*, 67, 397-400.
- Timm, H.W. (1983). The Factors Theoretically Affecting the Impact of Forensic Hypnosis Techniques on Eyewitness Recall. *Journal of Police Science and Administration*, 11, 442-450.
- Tinterow, M.M. (1970). *Foundations of Hypnosis: From Mesmer to Freud*. Charles C. Thomas: Springfield, IL.
- Tocchio, O.J. (1963). Lie Detection Under Hypnosis. *Police*, 8, 9-11.
- Udolf, R. (1981). *Handbook of Hypnosis for Professionals*. Van Nostrand Reinhold: New York.
- United States Congress, Office of Technology Assessment (1983). Scientific Validity of Polygraph Testing: A Research Review and Evaluation—A Technical Memorandum (OTA-TM-H-15). U.S. Government Printing Office, Washington, DC.

- Watkins, J.G. (1947). Antisocial Compulsions Induced Under Hypnotic Trance. *Journal of Abnormal and Social Psychology*, 42, 256-259.
- Watkins, J.G. (1972). Antisocial Behavior Under Hypnosis: Possible or Impossible? *International Journal of Clinical and Experimental Hypnosis*, 20, 95-100.
- Weitzenhoffer, A.M. (1957). *General Techniques of Hypnotism*. Grune and Stratton, New York.
- Weinstein, E., Abrams, S. & Gibbons, D. (1970). The Validity of the Polygraph with Hypnotically Induced Repression and Guilt. *American Journal of Psychiatry*, 126, 1159-62.
- Young, P.C. (1952). Antisocial Uses of Hypnosis. *Experimental Hypnosis*, L.M. LeCron, Ed. MacMillan Publishing Co.: New York.

An Objective Method for Manually Scoring Polygraph Data

Donald Krapohl & Barry McManus

Abstract

In an effort to improve the accuracy and interrater agreement of the traditional seven-position scoring system, objective scoring rules were devised using three key physiological tracing features previously identified by Kircher and Raskin (1988). The method assigns ordinal values to ratios that were created by dividing measurements of response features from the relevant questions by those of comparison questions. The physiological data for this project had been collected from live criminal cases in which ground truth had been established independently. For the training set of 300 cases, the three features were extracted from the data by software. When the scoring system was then tested against the 300 training cases, it demonstrated an accuracy of 82.3% when No Opinion (NO) decisions were counted as errors, and 93.9% when the NO outcomes were excluded. This outcome suggested that the method had potential, and cross validations were conducted. The first of two cross validations with 60 holdout cases found an accuracy of 83.3% including NO outcomes, and 89.3% without NOs. A second cross validation, using experienced polygraph examiners to manually perform the measurements on another set of 100 cases, rendered an accuracy of 82.3% when NO outcomes were included, and 93.3% correct without NOs. The cases used in the second cross validation had also been previously scored manually with the traditional seven-position scoring system by experienced examiners as part of another study. Those scorings produced fewer correct decisions than did the objective scoring method, and the poorer interrater agreement from that method limited the potential validity.

Keywords: cardiovascular activity, decision rules, electrodermal activity, objective scoring, reliability, respiration, respiration line length, scoring, seven-position scoring, validity

It is axiomatic that standardization of data assessment procedures improves interrater agreement. In the last 60 years several semi-objective scoring systems have been introduced to the field of psycho-physiological detection of deception (PDD) with the goal of standardizing the manner in which the physiological tracings are interpreted. (Winter, 1936; Backster, 1963; Honts & Driscoll, 1988; Gordon & Conchetti, 1988). A common shortcoming of all manual scoring systems to date is that the assignment of scores relies, to varying degrees, on subjective estimates of reaction intensity. Few existing manual scoring systems use any objective measurements for the assignment of scores, and none rely entirely on such measurements.

Because subjectivity has not been eliminated from the scoring systems, scorers in the field are vulnerable to the influence of extra-polygraphic information, drift, experience, source of training, or other factors (Honts & Perry, 1992; Honts & Amato, 1999). These factors may or may not introduce error into the decision process of a particular scorer, however, they would be expected to degrade agreement across examiners who had been influenced by the factors differently. Such degradation in reliability constrains validity, in addition to working against the acceptance of the technique in the scientific community. Complicating the existing field conditions more, even in the few instances that truly objective rules are taught and practiced, the

This paper was reviewed by Dr. Virgil Williams.

Both authors are polygraph examiners in federal service. The authors appreciate the assistance of Dr. Gordon Barland and Dr. Andrew Dollins of the Department of Defense Polygraph Institute, and Dr. Sethuraman of Florida State University for their statistical guidance. We are also grateful for the technical information provided by Dr. John Kircher, and the assistance of the two federal examiners who performed the blind scorings. Opinions expressed in this paper are those of the authors, and do not represent the position of the US Government.

literature does not supply evidence that any of the prevailing scoring rules are empirically based.

Related to the problem of assigning scores to responses is the issue of the decision rules. Scoring rules are how numbers are assigned to reaction comparisons, while decision rules are those that specify how those numbers are used to produce polygraph decisions of deception and nondeception. Many polygraph schools teach that the sum of all individual scores that are equal to +6 or greater warrants a decision of No Deception Indicated (NDI), -6 or smaller as DI (Deception Indicated), and all others as No Opinion (i.e., DoDPI, 1992). Critics have argued that the common decision thresholds are arbitrary (Furedy & Heslegrave, 1988) since there is no suggestion in the literature that they are optimum for all or even some subset of polygraph cases. Moreover, some polygraph practitioners employ the spot score rule (Light, 1999), which renders polygraph decisions that are based on the score of a single question irrespective of the total score for the examination.

All cutting scores establish some level of accuracy, that is, the act of setting decision thresholds dictates the error rate. User acceptance of any given set of thresholds assumes a concomitant risk. Though +/-6 thresholds are widely accepted, they may not be optimal in all settings since the consequences of error are unique to the case. If a system of scoring were in place which permitted perfect agreement among scorers, it would be possible to regulate the level of risk by selecting the appropriate threshold that corresponded with the probability of error. Barland (1985) proposed that scores could be used to build a distribution for the purpose of attaching probabilities of error to individual scores. Working with laboratory data, Barland constructed a table in which he determined the likelihood of error across cutting scores. There were two factors that limited the generalizability of the table, however. First, as Barland notes, the data were all from laboratory cases. There is a general consensus that the incentives and conditions between field and laboratory polygraph examinations are quite dissimilar, limiting to some degree what can be said about one from the other.

Secondly, the scorings were from a semi-objective scoring system. The lack of perfect objectivity in the method introduced a source of variability into scores, and diminished to some unknown degree the predictive value of the probability figures.

In the main, the more a total score is divergent from zero, the more likely it is to accurately indicate the veridical status of the examinee's statement. In other words, one can have greater confidence in a polygraph decision if the score on which it was based is further from zero than a decision where the score approaches zero. However, the lack of truly objective scoring rules in the field complicates the problem of establishing fixed cutting scores, since individual differences in scoring procedures blur the predictive power of the cutting scores. Put simply, a +6 overall score produced for one examiner is not always a +6 scoring for a second examiner on the same case, which limits the value of fixed cutting scores in traditional scoring.

These conditions call for the development of a manual scoring system that eliminates human estimates of relative reaction magnitudes entirely. The foundation for creating such an objective manual scoring system can be found in the work of Kircher & Raskin (1988). In the evaluation of polygraph data, Kircher et al. found the highest prediction value for guilt status to be three key physiological features; respiration line length (RLL) (Timm, 1982), electrodermal response (EDR) amplitude, and blood volume (BV) amplitude. It might, therefore, be possible to substitute these three features for the various criteria examiners currently use in manual scoring, creating a scoring system based entirely on simple measurements.

The Kircher features have specific measurement windows. RLL, which is the length of the respiration waveform, is measured for 10 seconds immediately after stimulus onset. In the assessment of EDR amplitude, a response must begin after 0.5 seconds from stimulus onset up to 5 seconds after the subject's verbal response. The measurement is from the level at stimulus onset to the maximum within 20 seconds. BV is the mean of the pulse wave, measured at stimulus onset until the presentation of the

next question. Artifacts tracings are not measured, nor are those from cases where an examinee has confessed to countermeasures. Each of the measurements produced a single raw value, permitting simple and objective scoring rules.

Simplicity of method should not be the only criterion for success, however. Standardization, accuracy, and defensibility are also key elements to the optimal scoring system. To maximize the usefulness of the scoring system, it must satisfy the following four criteria.

1. The system must be based entirely on objective measurements.
2. Scoring and decision rules must be individually empirically justified.
3. The method must, on average, produce more correct decisions than the conventional seven-position scoring method, and have a demonstrated inter-rater agreement greater than seven-position scoring.
4. The system must be practical enough to be accomplished in the field by nonscientist polygraph examiners.

For the present effort, data were used to develop a model of response patterns for the three key Kircher features, and to translate the pattern into the familiar seven-position scoring system used in the field. The seven-position scores resulting from that process were applied against the training data to determine whether the system had promise under optimum conditions. Two cross validation studies were then performed. One used an automated feature measuring system and the other employed field examiners measuring the features manually. The objective of this endeavor was to deliver a reliable and validated scoring procedure based on what is known about psychophysiological responding during polygraph testing.

Method

Training Polygraph Cases

Digitized polygraph data in a database of confirmed cases at the Department of

Defense Polygraph Institute were subjected to stratified random sampling. A total of 300 cases drawn, with the requirement that there be an equal number of truthful and deceptive cases. Confirmation of the cases came from the confession of the subject, confession of someone besides the subject (thereby exculpating the subject), or from other irrefutable evidence. Inclusion of cases in the database did not require that the original polygraph examiners make the correct decision regarding the subjects' veracity, but only that the cases be definitively resolved regarding the examinees' guilt or innocence. Selection criteria for this study did not include whether the examination was of a suspect, witness, or victim, nor were factors such as gender, age, education or race considered. Each selected polygraph case was a single-issue examination, and was conducted employing the DoDPI Zone Comparison Technique (ZCT) (DoDPI, 1992). These examinations consisted of three relevant, three probable-lie comparison (exclusionary), one sacrifice relevant, one irrelevant, and two symptomatic questions. Three charts from each case were used, and if the testing examiner conducted more than three charts, only the first three completed charts were used. The cases were all from field criminal investigations, and had been conducted by various US federal, state and local law enforcement agencies. All cases in the DoDPI confirmed cases database at the time of this study were recorded on the Axciton computer polygraph (Axciton Systems, Houston, TX).

Scoring Rules

The Kircher features were measured on all of the cases using software developed for DoDPI by the Johns Hopkins University Applied Physics Laboratory (Extract, 1998). The Extract software makes absolute measurements of the Kircher features, and drops those values into a commercially available spreadsheet. For the purposes of the present study, only those measurements associated with the relevant and comparison questions were considered. All measurement values were converted to ratios, with the value of response to the relevant question divided by that of the comparison question (R/C), creating a single number representing the subject's relative response intensity to the relevant question.

Consistent with DoDPI scoring rules (DoDPI, 1992) the response to the first relevant question is compared to that of the adjacent comparison question that evoked the stronger response. The stronger response in the current study was defined as the shorter respiration line length, and greater EDR and blood volume amplitudes. The values from the second and third relevant questions were divided by those of the adjacent preceding comparison question. In cases of missing values, the value of the nearest question of the same type was substituted for the absent values. Missing values resulted from artifactual responses, loss of signal, or negative amplitudes, and they constituted less than 2% of the original data.

The resulting ratios for all 300 cases were used to construct a grand distribution of ratios by channel, and each distribution was divided into seven portions such that each contained the same number of ratios. The seven divisions corresponded with the

positions of the 7-position scale; +3, +2, +1, 0, -1, -2, -3. Since all ratios were constructed as R/C, ratios from the EDR and blood volume data greater than 1.0 indicated larger reactions to relevant questions than to comparison questions, while ratios smaller than 1.0 indicated the opposite pattern. However, since shorter RLL is associated with greater arousal, the pattern of ratios from the respiration channel was the reverse of the EDR and blood volume; smaller ratios indicated greater reactions to the relevant questions. If one of the two respiration channels produced a 7-position score different from the other respiration channel, but of the same sign, the respiration channel providing the score further from 0 was preferred. If the respiration scores were on opposite sides of 0, a 0 was assigned. Each physiological channel provided 2700 ratios for the construction of a distribution (300 cases X 3 charts X 3 questions). The seven divisions are presented in Table 1.

**Table 1. Scoring rules based on 300 cases.
RLL data from the two respiration channels were pooled.**

Ratios (R/C)

7-Position Value	RLL	EDR	BV
3	>1.25	<.44	<.54
2	1.11 - 1.25	.44 - .67	.54 - .71
1	1.04 - 1.10	.68 - .92	.72 - .88
0	.97 - 1.03	.93 - 1.20	.89 - 1.05
-1	.88 - .96	1.21 - 1.60	1.06 - 1.29
-2	.79 - .89	1.61 - 2.44	1.30 - 1.67
-3	<.79	>2.44	>1.67

RLL = Respiration Line Length
EDR = Electrodermal Response
BV = Blood Volume

Carefully conducted studies (Kircher & Raskin, 1988; Olsen, Harris, Capps & Ansley, 1997) convincingly show that the electrodermal amplitude is the greatest contributor to PDD accuracy. In the present study, the scores from the electrodermal responses were doubled, to take advantage of this well

documented phenomenon. Therefore, the weightings of the polygraph channels were: respiration 25%, blood volume 25%, and electrodermal response 50%. Appendix A is a score sheet that demonstrates how the weighting is accomplished in practice.

The decision rules were: total scores for a case that are equal to or greater than +6 were called No Deception Indicated (NDI), those equal to or lower than -6 were called Deception Indicated (DI), and all others were No Opinion (NO). The spot score rule (rendering decisions based on cutting scores for individual questions) was not used.

Cross Validation I

Polygraph Cases

Sixty new digitized cases were selected from the DoDPI confirmed case database, satisfying the same requirements as the training data in terms of test format, proportion of truthful cases, and number of charts. None of these cases were used in the training set.

Scoring

The APL Extract program was used to measure the Kircher features, and these features were then converted to ratios, the ratios to scores, the scores summed, and decisions rendered, all in an identical manner as the training steps.

Cross Validation II

Polygraph Cases

Paper strip charts were made of 100 digitized field cases that had been part of another DoDPI study (DoDPI96-P-0001). They were all DoDPI Zone formats with either two or three relevant questions, two or three probable-lie comparison questions, and three charts. They were single-issue examinations,

and ground truth was established independently. Sixty-five of the cases were confirmed deceptive, and the remaining 35 were confirmed truthful.

Measurements

Two experienced polygraph examiners with a federal agency independently measured the Kircher features for the 100 cases, and entered the measurements into an electronic spreadsheet. RLL was measured by use of a digital map measuring meter (Brookstone, model 213066), and the EDR and BV amplitudes were measured in millimeters with a commercially available ruled plastic overlay designed for polygraph data analysis.

These 100 ZCT cases had also been manually scored by three experienced DoDPI examiners using DoDPI scoring method (DoDPI, 1998), and those results had been reported previously (Blackwell, 1999). The scorings of those cases were recalculated here to remove the effect of the spot score rule, and thus allow direct comparison with the present scoring system.

Results

Training set

Using the scoring rules in Table 1, and weighting the EDR data as described earlier, the training set had an average hit rate of 82.3%, error of 5.3%, and a NO of 12.3%. Excluding the NO outcomes, the scoring system on the training set had 93.9% correct decisions, and 6.1% error. Table 2 breaks down the data by decision.

Table 2. Correct, incorrect and No Opinion calls for 300 confirmed truthful and deceptive PDD cases using the objective scoring method and +/- 6 cutting scores.

	Nondeceptive Cases (n=150)			Deceptive Cases (n=150)		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
With No Opinion	82.7%	6.0%	11.3%	82.0%	4.7%	13.3%
Without No Opinion	93.2%	6.8%	--	94.6%	5.4%	--

Cross Validation I

Table 3 is the outcome of the cross validation using a holdout set of 60 confirmed cases. Overall accuracy was 83.3% with NO cases included, and 89.3% without them. In

the with-NO data, hit rates for non-deceptive cases were better than with deceptive cases, and the proportions of accurate decisions were significantly different ($z=2.06$, $p<.05$), though this did not hold true for the without-NO condition ($z=1.05$, ns).

Table 3. Percentage of correct, incorrect and No Opinion calls for confirmed truthful and deceptive PDD cases using the proposed objective scoring method and +/- 6 cutting scores.

	Nondeceptive Cases (n=30)			Deceptive Cases (n=30)		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
With No Opinion	93.3%	6.7%	0%	73.3%	13.3%	13.3%
Without No Opinion	93.3%	6.7%	--	84.6%	15.4%	--

Cross Validation II

Table 4 shows the accuracy of the objective scoring method based on the measurements performed by two experienced examiners, and the objective scoring system. Including NO outcomes, the average accuracy for the two objective scorers on nondeceptive cases was 81.5%, and 93.5% when those

outcomes were excluded. For deceptive cases, the objective scorers averaged 83.1% accuracy with NO outcomes, and 93.1% without them. Hit rates for deceptive and nondeceptive cases were not significantly different ($z=0.21$, ns). The proportion of agreement between the two objective scorers was 0.86, which was significantly greater than chance ($z = 7.63$, $p<.01$).

Table 4. Percentage of correct, incorrect, and No Opinion outcomes for 100 cases using the objective scoring system.

	Nondeceptive Cases (n=35)			Deceptive Cases (n=64)		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
Objective Scorer 1	82.9%	5.7%	11.4%	83.1%	6.2%	10.8%
Objective Scorer 2	80.0%	5.7%	14.3%	83.1%	6.2%	10.8%
Average	81.5%	5.7%	12.9%	83.1%	6.2%	10.8%
Average w/o NOs	93.5%	6.5%	--	93.1%	6.9%	--

The cases used in this study had been previously scored blindly by three federal examiners as part of another study (DoDPI96-P-0001, Blackwell, 1999). Table 5 shows a reevaluation of those scoring data, to exclude the effects of the spot score rule. Average

accuracy across the three Blackwell scorers for non-deceptive cases was 58.1% when NO outcomes are included, and 91.1% without them. Average accuracy for deceptive cases was 72.3% including NO outcomes, and 98.0% when those outcomes were excluded. The

proportions of correct decisions for deceptive and nondeceptive cases were not significantly different from one another when the NO cases were included ($z=1.44$, ns), nor when they were excluded ($z=1.60$, ns). Average proportion of

agreement for decisions between all pairs of these examiners with these cases was .72, which was significantly better than chance ($z=5.52$, $p<.05$).

Table 5. Percentage of correct, incorrect, and No Opinion results from three scorers of physiological data from 100 field polygraph examinations employing traditional scoring rules without the spot score rule (n=100 decisions per scorer). (Reanalyzed data from Blackwell, 1999.)

	Nondeceptive Cases (n=35)			Deceptive Cases (n=64)		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
Scorer 1	77.1%	5.7%	17.1%	67.7%	1.5%	30.8%
Scorer 2	48.6%	2.9%	48.6%	78.5%	1.5%	20.0%
Scorer 3	48.6%	8.6%	42.9%	70.8%	1.5%	27.6%
Average	58.1%	5.7%	36.2%	72.3%	1.5%	26.1%
Average w/o NOs	91.1%	8.9%	--	98.0%	2.0%	--

The recalculated scoring data from the Blackwell study were compared to those of the objective scoring system since the only difference in the conditions between the two was the method of scoring. Table 6 shows the proportion of agreement among all pairs of scorers, and each scorer against ground truth. Agreement between the objective scorers was

greater than the average agreement between pairs of traditional scorers (.86 vs .72, $z=2.43$, $p<.05$). The proportion of agreement between the objective scorers' decisions and ground truth was significantly greater than the average proportion of agreement of traditional scorers and ground truth (.82 vs .67, $z=2.47$, $p<.05$).

Table 6. Proportion of agreement between pairs of objective scorers and traditional scorers with 100 polygraph cases.

	OS 2	TS 1	TS 2	TS 3	Ground Truth
OS 1	0.86	0.75	0.71	0.65	0.83
OS 2		0.69	0.67	0.64	0.82
TS 1			0.70	0.72	0.71
TS 2				0.73	0.68
TS 3					0.63

OS 1 = Objective Scorer 1
OS 2 = Objective Scorer 2

TS 1 = Traditional Scorer 1
TS 2 = Traditional Scorer 2
TS 3 = Traditional Scorer 3

Table 7 displays the average score assigned to each channel for nondeceptive and deceptive cases (electrodermal scores were unweighted for this comparison). While the EDR receives the most diagnostic scores for both nondeceptive and deceptive cases, the objective scoring system assigned more diagnostic scores to the RLL and BV for both nondeceptive and deceptive cases than did the average traditional scorer. The average scores across channels for the objective scoring system are more closely clustered than those

of the traditional scorers. It would appear from these data that traditional scoring had a de facto method of weighting the electrodermal channel, that is, by having a reduction in scores assigned to the other two channels. Viewed from another perspective, the data would indicate that the objective scoring system captures more scores in the respiration and blood volume channel than does the traditional scoring system, suggesting some optimization for those channels.

Table 7. Mean scores by channel for nondeceptive and deceptive cases for five scorers and two scoring systems. For the nondeceptive cases, each scorer had 315 RLL, EDR and BV scores (35 cases X 3 charts X 3 questions). For deceptive cases, each scorer had 585 RLL, EDR and BV scores (65 cases X 3 charts X 3 questions).

	Mean Scores for Nondeceptive Cases				Mean Scores for Deceptive Cases			
	RLL	EDR*	BV	Total	RLL	EDR*	BV	Total
OS 1	3.37	4.97	3.11	11.45	-4.91	-6.72	-4.97	-16.6
OS 2	3.51	4.57	3.14	11.22	-4.08	-6.62	-4.6	-15.3
TS 1	2.34	3.34	4.06	9.74	-0.08	-7.33	-2.83	-10.24
TS 2	0.66	3.00	2.11	5.77	-1.97	-6.31	-3.20	-11.48
TS 3	1.34	3.77	1.97	7.08	-0.85	-7.29	-1.29	-9.43
Average OS	3.44	4.77	3.13	11.34	-4.5	-6.67	-4.79	-15.96
Average TS	1.45	3.37	2.71	7.53	-0.97	-6.98	-2.44	-10.39

* Note: The EDR scores were unweighted for this comparison.

RLL = Respiration Line Length

EDR = Electrodermal Response

BV = Blood Volume

OS = Objective Scorer

TS = Traditional Scorer

Critics of the objective scoring system could argue that this system produces higher average scores because of its tendency to give more 2s and 3s than the traditional scoring system, and that same propensity may also make it more inclined to give scores in the wrong direction. To complement the data of average scores, it is also important to examine how often total scores for each channel were on the correct side of zero. Table 8 is a side-by-side comparison, using the 100 Blackwell cases again, of the percentages of scores in the

correct direction. Only one significant difference in proportions between the systems was detected at the .05 level: the proportion of scores for the respiration channel assigned by the objective scoring system was greater than those assigned by the traditional scoring system ($z=2.08$, $p<.05$). From these data, it would appear that the objective scoring system is no more inclined to assign scores in the incorrect direction than does the traditional scoring system.

Table 8. Average percentage of total scores in the correct direction for nondeceptive and deceptive cases by channel for two scoring systems.

	Nondeceptive Cases (n=35)				Deceptive Cases (n=65)			
	RLL	EDR	BV	Average	RLL	EDR	BV	Average
Average OS	65.7%	81.4%	54.2%	67.1%	71.5%*	80.8%	72.3%	74.9%
Average TS	62.9%	69.4%	69.4%	67.2%	53.9%*	83.5%	70.3%	69.2%

* Percentage is significantly different from corresponding percentage in other scoring system. ($p < .05$).

RLL = Respiration Line Length
 EDR = Electrodermal Response
 BV = Blood Volume
 OS = Objective Scorer
 TS = Traditional Scorer

The substitution of subjective estimates of relative reaction intensity with objective measurements sets the stage for perfect agreement among those who perform the objective scoring system, permitting the calculation of error for each cutting score. In that vein, the present developmental data, the first cross validation data, and the second cross validation data (one of the two manual scorings was chosen by coin toss) were pooled for the purpose of estimating error rates from the distribution of scores for the deceptive ($n=245$) and nondeceptive ($n=215$) cases, following the Barland methodology (1985). Table 9 shows the results. According to these data, if an examiner, utilizing the objective scoring system and the ZCT technique with the DoDPI protocol, scored an examination as a -8, Table 9 would indicate that a mere 6% of nondeceptive cases using the objective scoring system have scores that low or lower. In other words, all things considered, an NDI call with that score is a long shot for being correct. Similarly, a DI call based on a +20 score would have a 1% chance of being correct. If the user wishes to avoid errors, wide cutting scores can be applied. Conversely, if decisiveness (avoidance of NOs) is important in a specific application, a larger proportion of decisive calls will result from narrowing the NO band with narrower cutting scores.

Of course, overall error is not simply a function of score, but also of base rate. Lower or higher base rates of deception alter the

likelihood of error on any given single case. Table 9 assumes an equal proportion of deceptive and nondeceptive examinees in the population. Moreover, the listed error estimates are specific to three-question, three-chart, single-issue ZCT examinations scored with the objective scoring system outlined in this paper, with no consideration for spot scores. They may not generalize to other configurations. Examiners should use Table 9 data with an awareness of those limitations.

Discussion

It was previously reported that among field cases there is evidence of an asymmetry in the response patterns of deceptive and nondeceptive examinees to relevant and comparison questions (Krapohl, 1999; Raskin, Kircher, Honts, & Horowitz, 1988). Deceptive examinees appear to respond stronger on average to relevant questions than nondeceptive examinees respond on average to comparison questions. Traditional scoring and decision rules are almost exclusively symmetrical; that is, there is an implicit expectation that deceptive response patterns (R to C) are reverse images of nondeceptive response patterns (C to R). Since this symmetry has not been borne out in the data, it would appear that traditional scoring may not be as sensitive to truthfulness as it is to deceptiveness. The objective scoring system tested here was devised to compensate for the asymmetry by setting the scoring rules

Table 9. Probability estimates for scores of deceptive and nondeceptive cases when using the objective scoring system.

Score	Probability of a truthful subject having this score, or lower	Probability of a deceptive subject having this score, or higher
-40	0.01	
-38	0.01	
-36	0.01	
-34	0.01	
-32	0.01	
-30	0.01	
-28	0.01	
-26	0.01	
-24	0.01	
-22	0.01	
-20	0.02	
-18	0.02	
-16	0.02	
-14	0.03	
-12	0.04	
-10	0.05	
-8	0.06	0.20
-6	0.07	0.18
-4	0.09	0.15
-2	0.11	0.13
0	0.13	0.11
2	0.15	0.09
4	0.18	0.08
6	0.21	0.06
8	0.24	0.05
10		0.04
12		0.03
14		0.03
16		0.02
18		0.02
20		0.01
22		0.01
24		0.01
26		0.01
28		0.01
30		0.01
32		0.01
34		0.01
36		0.01
38		0.01
40		0.01

to match the asymmetry. The effect of this compensatory action can be seen most clearly with the cases scored by both the objective and traditional scoring systems. As seen in Table 4, correct, incorrect and NO outcomes for deceptive cases had highly similar proportions to those of the nondeceptive cases with the objective scoring system. Table 5 shows the same cases scored traditionally, with the predicted shift in efficiency toward identification of the deceptive, with a corresponding decrement in detecting the nondeceptive. Field examiners who prefer any of the conventional seven-position scoring systems in their various forms should be cognizant that the scoring and decision rules of that system may not be sensitive to the trend of asymmetry in physiological responding. Identification and validation of any fixes to this shortcoming will require further investigation.

A practical disadvantage of the objective scoring system is the amount of time required to perform it, from 30 to 45 minutes per case. Traditional scoring of polygraph charts can usually be completed in a few minutes, whereas the objective scoring requires careful measurements, followed by a series of algebraic computations. Even if the mathematics were performed automatically in an electronic spreadsheet, the exacting measurements require much more time than do the assignment of values in the traditional scoring. This additional time requirement can be a significant resource issue in some settings. However, in the evidentiary arena, where polygraph decisions are arguably most important, and interrater agreement can be a nettlesome issue, the objective scoring system provides a common means for the valid interpretation of the physiological data, regardless of the philosophy of the participating examiner experts. It also affords those examiners an opportunity to estimate the probability of error, a feature not found in any other manual polygraph scoring method. Error estimates are important in evidentiary settings, and a lack of this capability has been a major criticism from the judicial bench.

Of the four requirements of the scoring system outlined in the introduction of this paper, three were fully satisfied. First, the present scoring system was based exclusively

on objective measurements of the physiological data. Second, all of the scoring rules used empirical data and the findings of previous research, so that all were individually and scientifically justified. Third, the method rendered more correct decisions than the traditional scoring, and enjoyed greater interrater agreement with the present data. The last requirement, that the method be practical enough for field use, was only partially supported. Despite very high agreement, there was a degree of difference in the measurements produced by the two field examiners applying the objective scoring system. Those differences may be attributable to measurement error, examiner fatigue, recording error, or other human factors. Some of the errors may be avoided in the field when a single case is the focus of the scoring, versus the scoring of 100 cases with its attendant 7200 measurements and data entries as was done in this study. Nevertheless, the effectiveness of this objective scoring method may always be constrained to some degree by the scorer's ability to record accurate measurements.

The objective scoring system outlined in this paper is not offered as a substitute for traditional scoring in everyday field use. Indeed, the modest error rates demonstrated in the traditional scoring system with the experienced polygraph examiners are comparable to those of the objective scoring system, notwithstanding the higher NO rate of the former. Given the relative ease of the traditional scoring method, there are good reasons to recommend its use in most applications. The objective scoring method was designed specifically for the singular use in evidentiary settings, where the additional time demands of the system are offset by its better precision. It offers a common method for all scorers, regardless of training, to arrive at precisely the same score. It is the authors' hope that it will be brought to bear on the problem of the courtroom "battle of the polygraph experts," so that the polygraph discipline can help clear, rather than cloud, the issue of witness and suspect veracity.

Final Thoughts on Scoring

The phenomenon of differential attentional set exploited in conventional PDD is demonstrably reliable and robust. This may

be one reason why all published studies of PDD have found a significant effect for discrimination between deceptive and nondeceptive subjects, even when examiners use very different scoring methods [for an appreciation of these dissimilarities, see *Polygraph* (1999) 28(1)]. Differences in scoring methodology have largely been ignored as an empirical question, since the incidence of opposite decisions is somewhat rare despite some variability in scores. The low rate of opposite outcomes, however, is probably more a testimony to the magnitude of the effect of psychological set than the efficacy of a given scoring technique. Almost every scoring method yet devised works pretty well. Arbitrary and idiosyncratic rules seem to be tolerated by the PDD data, and this has created a fertile field for rule makers. There has been no sense of urgency to test whether a particular scoring or decision rule within a scoring technique is valid. Rule makers simply point out that the system works, shifting the attention from the possibility that there may be unnecessary, inefficient, or even bad rules in the system. With increasing scrutiny by scientists from other disciplines, there are clouds on the horizon for the profession's blissful nescience. The polygraph

discipline would profit from a departure from the theory-heavy data-light approach that tends to plague discussions and instruction of manual scoring in PDD, and to subject scoring procedures to empiricism where rules can be tested, and confirmed or disconfirmed in the scientific tradition. Our friends in the scientific community warn us that this is the direction we must go.

The discipline of polygraphy has grown to the stage of maturity where it should be cautious of scoring procedures unless there are published data to support them, regardless of the prevalence of the practice in the field. Absent data, these rules should be considered intelligent, logical, unsupported personal opinions, regardless of how confidently they are taught. They are subject to individual experience, periodic adjustments, blind spots, and hindsight explanations. They are problematic for the polygraph profession. One is reminded of the candid words of the brilliant scientist Lev Landau who, when commenting on his field of theoretical physics, said, "We are often in error, but never in doubt." In other words, confidence is no substitute for evidence. Let's look at the data.

References

- Backster, C. (1963). The Backster chart and reliability rating method. *Law and Order*, 11, 63-64.
- Barland G.H. (1985). A method for estimating the accuracy of individual control question tests. *Proceedings of Ident-85*, 142-147.
- Blackwell, N.J. (1999). PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examinations from actual criminal investigations. DoDPI96-P-0001. Reprinted in *Polygraph*, 28(2), 149-175.
- Capps, M.H. & Ansley, N (1992). Numerical scoring of polygraph charts: What examiners really do. *Polygraph*, 21(4), 264-320.
- Department of Defense Polygraph Institute. (1992). Zone comparison test. Ft. McClellan, Alabama.
- Furedy, J.J., & Heslegrave, R.J. (1988). Validity of the lie detector: A psychophysiological perspective. *Criminal Justice and Behavior*, 15(2), 219-246.
- Gordon, N.J. & Cochetti, P.M. (1987). The horizontal scoring system. *Polygraph*, 16(2) 116-125.

- Honts, C. R. & Driscoll, L.N. (1988). A field validity study of rank order scoring system (ROSS) in multiple issue control question tests. *Polygraph*, 17(1), 1-16.
- Honts, C.R. & Perry, M.V. (1992). Polygraph admissibility: Changes and challenges. *Law and Human Behavior*, 16(3), 357-379.
- Honts, C.R. & Amato, S. (1999). The automated polygraph examination. Technical Report to the Central Intelligence Agency. Contract No: 110224-1998-MO.
- Kircher, J.C. & Raskin, D.C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology*, 73, 291-302.
- Krapohl, D. (1999). Short report: Proposed method for scoring electrodermal responses. *Polygraph*, 28(1), 82-84.
- Light, G.D. (1999). Numerical evaluation of the Army Zone Comparison Test. *Polygraph*, 28(1), 37-45.
- Olsen, D. E., Harris, J.C., Capps, M.H., & Ansley, N. (1997). Computerized polygraph scoring system. *Journal of Forensic Sciences*, 42(1), 61-70.
- Raskin, D.C., Kircher, J.C., Honts, C.R., & Horowitz, S.W. (1988). A study of the validity of polygraph examinations in criminal investigations: Final report to the National Institute of Justice. Grant No. 85-IG-CX-0040.
- Timm, H. W. (1982). Effect of altered outcome expectancies stemming from placebo and feedback treatments on the validity of the guilty knowledge technique. *Journal of Applied Psychology*, 67, 391-400.
- Winter, J.E. (1936). A comparison of the cardio-pneumo-psychograph and association methods in the detection of lying in cases of theft among college students. *Journal of Applied Psychology*, 20, 243-248.

Appendix – Objective Scoring Worksheet

DK 9/99

Chart 1	MEASUREMENTS						RATIOS R/C			SCORES (circle) see table		
	CQ1	RQ1	CQ2	RQ2	CQ3	RQ3	Spot 1	Spot 2	Spot 3	Spot 1	Spot 2	Spot 3
RLL 1										1 resp score only, the stronger.		
RLL 2										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
EDA										- + 0 2 4 6	- + 0 2 4 6	- + 0 2 4 6
BV										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
Total												

Chart 2	MEASUREMENTS						RATIOS R/C			SCORES (circle) see table		
	CQ1	RQ1	CQ2	RQ2	CQ3	RQ3	Spot 1	Spot 2	Spot 3	Spot 1	Spot 2	Spot 3
RLL 1										1 resp score only, the stronger.		
RLL 2										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
EDA										- + 0 2 4 6	- + 0 2 4 6	- + 0 2 4 6
BV										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
Total												

Chart 3	MEASUREMENTS						RATIOS R/C			SCORES (circle) see table		
	CQ1	RQ1	CQ2	RQ2	CQ3	RQ3	Spot 1	Spot 2	Spot 3	Spot 1	Spot 2	Spot 3
RLL 1										1 resp score only, the stronger.		
RLL 2										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
EDA										- + 0 2 4 6	- + 0 2 4 6	- + 0 2 4 6
BV										- + 0 1 2 3	- + 0 1 2 3	- + 0 1 2 3
Total												

Sub totals (all charts)												
										Grand Total		

Scoring Table							
RLL	-3	-2	-1	0	1	2	3
	0.00 to 0.79	0.80 to 0.89	0.90 to 0.96	0.97 to 1.03	1.04 to 1.10	1.11 to 1.25	1.26 to 999
EDA	-6	-4	-2	0	2	4	6
	999 to 2.45	2.44 to 1.61	1.60 to 1.21	1.20 to 0.93	0.92 to 0.68	0.67 to 0.44	0.43 to 0.00
BV	-3	-2	-1	0	1	2	3
	999 to 1.67	1.66 to 1.30	1.29 to 1.06	1.05 to 0.89	0.88 to 0.72	0.71 to 0.54	0.53 to 0.00

Case # _____

Examiner _____

Date _____

A Response to Honts on the Issue of the Discussion of Questions Between Charts

Stan Abrams

Abstract

This paper is in response to Honts' statements published in the last edition of *Polygraph* that Abrams has inappropriately testified against the admissibility of the Utah version of the directed lie test, resulting in this testimony being rejected as evidence. Abrams has provided evidence supporting his view that this technique neither should be admitted into court nor employed as a polygraph technique because there is insufficient and conflicting research findings, there is a discussion of questions between charts, and there is an excessive emphasis placed on the directed lie which has resulted in false negative findings.

Key words: directed lie, reply

This paper is in response to Honts' article "The Discussion of Questions Between List Repetitions (Charts) Is Associated with Increased Test Accuracy" (1999). In his paper, Honts indicated that this writer has testified against the directed lie test in U.S. v. Gilliard because of the discussion of questions between charts resulting in polygraph testimony not being admitted into evidence in that case and in others where the court or jury findings favored the opposing side for which they testified. However, the problem of discussion between charts is only one of the reasons that the writer believes that this particular procedure should not be admitted into evidence, or even viewed as an acceptable polygraph technique. The major reasons are that the research is minimal and conflicting, and that the directed lie test (DLT) used by Honts places too much emphasis on the directed lie question (DLQ) resulting in false negative findings, i.e., finding a deceptive person truthful.

Government polygraphists developed the DLT mainly for use as a screening examination. In their approach, the DLQ replaced the probable lie question (PLQ). Raskin modified this approach to be used in a multiple issue specific test. However, in doing so, he made two major changes in the

procedure. There is discussion about the questions between charts, but more importantly, an emphasis is placed on the hybrid directed lie question (HDLQ), as this approach was labeled in U.S. v. Gilliard. The purpose of this was to differentiate the government's DLT from Raskin and Honts' hybrid directed lie test (HDLT). It is assumed by Abrams that the original purpose of the emphasis on the HDLQ was to increase the truthful subject's reactivity in response to the HDLQ, thereby, reducing the likelihood of false positive responses, i.e., for a truthful person to be inaccurately labeled as deceptive. It is the writer's contention that in the process, too much emphasis was placed on the HDLQ, thereby, resulting in an increase in false negatives. Therefore, it is important to separate the DLT, which is sometimes employed by government examiners, from the HDLT that is used by Honts and Raskin because of these two significant differences that exist. The government DLT does not permit any discussion between charts and certainly does not place any undue emphasis on the DLQ. Another reason for separating the two procedures is that in some of the cases in which this writer has testified, both the DLT and the HDLT were grouped together and it was implied that the HDLT approach was used by many governmental agencies. In fact, this

Dr. Abrams is a regular contributor to *Polygraph*. Any reprint requests can be forwarded to him at 1618 S.W. First Ave. Suite P-2, Portland, OR 97201

is not the case, but rather, government examiners in some instances do employ the DLT but not the HDLT.

The first information published on the government's DLT was reported by Fuse in 1982. He indicated that he found this approach to be very effective in multiple issue screening tests, but he specifically warned that if too much emphasis were placed on the DLQ, it could result in a false negative finding. Of course, the reverse could occur if stress were placed on the relevant question. Gelb also indicated the risk of this occurring in the use of his pre-incident indicator procedure (1994). He stated that if excessive emphasis were placed on the pre-incident, that too could cause a false negative finding to occur.

Interestingly enough, and consistent with these findings, Honts and Perry wrote "...that an unethical and dishonest polygraph examiner could manipulate the subject and the examination in such a way as to produce a desired result. For example, an examiner who was motivated to produce a deceptive result might ask over-general or provocative relevant questions, and spend a great deal of time on their review and presentation. Subsequently, this unethical examiner could ask very narrow, specific or inappropriate control questions and spend a great deal of time on their review and presentation. An examiner predisposed to produce a truthful result could take the opposite approach, overemphasizing the control questions and minimizing the relevants" (1992).

In discussing their research on countermeasures, Honts, Raskin, and Kircher (1994) indicated that mental activity, such as counting backwards from a hundred by sevens, was an effective countermeasure that caused a physiologic response much like that of the reaction that occurs in association with lying. There is a parallel here with the HDLT subject who is instructed to lie to the question "Did you ever tell even one lie in your life?" and then told to "think of a time that you actually did lie." One must assume that the latter is also mental activity similar to that used as a countermeasure. Therefore based on their own research, they have increased the likelihood of obtaining a false negative result. Unlike their research on countermeasures,

however, they are not only instructed to "Think of a time when you lied" immediately before the testing starts, but between each chart as well. Consider the impact of this repetition where the examiner repeated the test six times. Based on their own research findings, mental activity increases the likelihood of obtaining a false negative result.

In essence, all of this means that there is a delicate balance that exists between the comparison and relevant questions and many variables can tip this balance in either of those two directions. Too much discussion of one or the other during the pretest, a difference in inflection or loudness when the questions are being asked, any discussion between charts that stresses either the relevant or comparison questions, or any mental activity on one question versus another can weigh the balance in the direction of that particular emphasis.

In contrast to Honts' statement that this writer was in disagreement with the discussion of questions between charts, the objection, much more specifically was to the unequal discussion of questions between charts. It was found that in the HDLT the HDLQ was emphasized more than the relevant question in every test of this nature that was reviewed for possible testimony. Moreover, there was little standardization of this approach in that the instructions varied not only with different subjects, but also with different examiners.

Another disadvantage of this procedure was that the presentation made to the subject between the charts at times motivated the subject to respond and become involved in a discussion. Almost always, this was related to the HDLQ, which also had to result in an even greater tendency to react to those questions as compared to the relevants. Interestingly, in two cases, the subjects even laughed during the HDLQs, apparently because the lie that they denied was so outlandish.

The purpose of Honts' study on the impact of the discussion of the polygraph questions between charts, would appear to be based, at least in part on the writer's testimony against the HDLT in U.S. v. Gilliard in 1996. In Honts' discussion of this case, he emphasized that Abrams was unable to cite

any literature on the effect of discussing questions between charts. He neglected to mention Abrams' study that was presented before the court in the Gilliard case, which at the time, was one of the only three studies of the HDLT that existed.

Abrams employed an actual tape of Raskin's presentation of the HDLT and repeated this verbatim with subjects being examined in real crime situations. Only cases in which the accuracy was confirmed were included. It was found that when it was compared to the results with the PLQ, there was a definite trend for both the truthful and deceptive subjects to move more in the direction of truthfulness. Some of these changes, particularly with deceptive subjects, were quite strong, indicating that placing emphasis on the HDLQ through a discussion between charts could certainly influence even deceptive subjects to appear more truthful and in some instances to be found truthful. The reader is referred to this article to obtain a more complete description of how the HDLQ is presented to a subject (Abrams, 1991).

One of the other two studies presented in the Gilliard testimony was the laboratory research of the HDLT conducted by Horowitz. He used a mock crime design and compared the relevant/irrelevant (RI) approach, the probable lie test (PLT), and two forms of the HDLQ, one using personal issues and the other employing neutral issues. For the truthful subjects, the accuracy rate excluding inconclusives was reported to be 87% for the personal HDLT and 84% for the neutral HDLT. For the PLT 86% accuracy was reported and 22% for the RI approach. With deceptive subjects, for the personal and neutral HDLTs, 84% and 73% accuracy respectively were reported. The PLT accuracy was found to be 73% and the RI 100%. At the end of the article, the author warned of the risk of generalizing from laboratory research to actual testing in the field. This is a legitimate concern since probable lie question validity research in the field usually demonstrates about 95% accuracy for deceptive subjects in contrast to this study, which reported only 73%. The result for the deceptive subjects is so at variance with the usual field research that these findings have to be viewed with some doubt.

The final study that had been conducted in this area was by Honts and Raskin who did a field study of 25 subjects. Verification, however, consisted of not only admissions, but also physical evidence that reportedly completely exonerated the subject, and retractions by the victims. Since 6 of the 13 truthful subjects were accused of the sexual abuse of children, and knowing how frequently children who were actually abused recant, this would place a considerable degree of doubt on their so called confirmed truthful subjects. The authors compared the relevant questions with two PLQs and one HDLQ. Blind scoring produced 95.6% total accuracy, but they reported that only 90% accuracy was reached when only the PLQs were employed. Since it is highly likely that some of their confirmed truthful subjects were actually deceptive, the results of this study must be considered flawed. In addition, attempting to generalize from a study employing two PLQs and one HDLQ is weak since it does not match any of the field studies that are now being conducted by Honts and Raskin. It is certainly not the same as the HDLT used by Honts in the Gilliard case, which used one PLQ and two HDLQs.

Based on the above three studies, Abram's major reason for testifying against the HDLT being admitted into evidence has been because there is both minimal and conflicting research that exists on the HDLT. In fact, that was the reason given by the Court for rejecting polygraph admissibility in U.S. v. Gilliard.

Further evidence of this is seen even more recently in a 1998 article in *Polygraph*, in "A Guide to DoDPI Research Interests." Dollins stated that, "It has not, however, been verified that a specific issue examination using directed lie comparison questions is as effective or accurate as one using probable lie comparison questions." It is assumed that in this case, DLT refers to both the government version and the hybrid approach.

One also should be aware that in a poll of the APA accredited polygraph schools conducted by this writer and presented in U.S. v. Gilliard there was only one school that actually taught the use of the HDLT.

In U.S. v. Gilliard, the defendant previously had been found guilty of defrauding Medicare, and in the case raised by Honts, he had been accused again at a completely different time and situation, but of the same offense. According to Honts, Abrams' testimony caused the rejection of his polygraph testimony, but he neglected to mention the fact that Gilliard was convicted of 102 counts of fraud. Although a conviction might not be the best verification of guilt, it would certainly suggest that there was other compelling evidence in this case.

Another reason for the writer testifying against the HDLT approach has been that he had never found truthful findings when the charts were independently evaluated. In the Gilliard case, for example, a high level FBI examiner and Abrams individually scored the charts, and neither polygraphist found the charts supported a call of truthfulness.

Honts also indicated in his paper that he had considerable concern about Abrams testifying against polygraphy. It must be recognized, however, that in each case that he cited, and in others that were not mentioned, the tests were only conducted by either Honts or Raskin employing the HDLT. If there were any evidence to the questionable validity of this approach, it can be found in the fact that, with the exception of one time, in no case in which the writer testified or assisted in the case was the HDLT admitted in evidence. In that case where it was admitted, New Mexico v. Mead, Raskin found the defendant truthful when he denied sexually abusing two children with one score being +21 and the other +10. Both tests had a probability of accuracy of .90. The defendant pled guilty after Raskin was cross examined. Because Raskin has argued that people who plead are not necessarily guilty, it should be pointed out that the court would not accept the plea until the defendant actually described in detail what sexual acts he had committed on the two children. In several other cases where polygraph had already been admitted into evidence, the court or jury found in opposition to the HDLT.

The final and perhaps most compelling reason for testifying against those cases employing the HDLT is the outcome of many of these cases. If the Gilliard case were the only

one in which strong truthful polygraph scores were obtained (+27 with a .956 probability of truthfulness) and the court rejected admitting his polygraph findings, one could assume that the decision of the court, and then later the jury, were in error or that the examiner came to an occasional incorrect conclusion. However, there have been too many cases in which subjects have been found truthful with high polygraph scores, but confessions occurred, defendants pled, or were found guilty in court. Because some of these cases were of a high profile nature, it resulted in considerable damage to polygraphy. In the infamous Hofmann double homicide case in Utah, the HDLT was employed by Honts, and Hofmann was found truthful. The call was confirmed by Raskin. Unfortunately for polygraphy, these findings were reported on television just at about the same time that a confession was being obtained. Raskin claimed that the error was due to the subject using hypnosis as a countermeasure. There were other high profile cases that received considerable media attention and also were detrimental to polygraphy. In the Seattle Chinatown massacre, thirteen people were killed, but the examiner reported that Mak was truthful. Mak received the death penalty. MacDonald, the Green Beret physician who is in Federal prison for killing his wife and two children also was viewed as being honest. In the notorious Woodward case, where the English nanny who admitted to the police that she shook the baby and dropped her on her head, she too was seen as being truthful. These are only a few of those cases in which those using the HDLT have provided grist for the media, which allows them to sensationalize the inaccuracy of polygraphy.

Honts stated that this writer not only hurts polygraphy, but society as well by testifying against the HDLT. Consider New Mexico v. House, where the defendant who was reported as truthful in stating that his collision when he drove in the wrong direction on a freeway was due to a migraine headache rather than the fact that he was found to be legally intoxicated. The defendant killed a mother and her three children. He was found guilty by the court. In New Mexico v. Wilson, an elementary teacher was found truthful with a .95 probability of the results being accurate when she denied molesting any of twelve

children in an elementary school. Fortunately for the children she abused and those that she would have molested later, the police obtained an admission. There have been other cases of truthful findings that have later been shown to be inaccurate. These include, Anderson v. Samrock and the Bernalillo Sheriff's Office, Griffith v. Melgaard, U.S. v. Freedman, New Mexico v. Raebuck, Idaho v. Kildare, Wyoming v. Reno, and New Mexico v. Martins to name some of them. For a list of more details and additional cases, the reader should consult the testimony in U.S. v. Clayton and Dalley or Steve Griffith v. Muscle Improvement Inc. One should also recognize that the worst part of this is how many other cases exist in which the experts have convinced a jury that the defendant was innocent and this was not in fact the case.

In Honts' final statement in his paper he said of the writer that in testifying against polygraph admissibility Abrams has committed "...a grave injustice to the court, the polygraph profession and ultimately society." It would seem that Honts would prefer the publicity of inaccurate findings to damage the reputation of the polygraph profession by subjecting this field to ridicule. It is errors of this nature that cause the courts to reject

polygraph testimony, and worse, they assist a guilty person to go free and further prey on society. This creates a much greater injustice than that of this writer testifying against the admissibility of this manner of polygraph evidence.

Perhaps the damage to the polygraph profession is best demonstrated by the findings of U.S. District Judge Gary Taylor in United States v. Cordoba. "The blanket and non-critical approval of Defendant's test by Dr. Raskin, who is probably the strongest and best informed advocate for polygraph admissibility, illustrates that the polygraph industry lacks sufficient controlling standards to satisfy Daubert. If pro polygraph's best expert declines to find any fault with an obviously faulty examination, that is strong evidence that there are insufficient controlling standards."

That is the type of situation that will create major damage to polygraphy. It is what will keep polygraphy from being admitted into court and hurt society, not this writer's actions to keep inadequate, and insufficiently researched approaches from being admitted into evidence and later found to be inaccurate, or worse, result in the release of a guilty person.

References

- Abrams, S. (1991) The directed lie control question. *Polygraph*, 20, 26-31.
- Anderson v. Samrock & Bernalillo County Sheriff's Office, 2nd Judicial District Court, Case No. CV-93-7878, Albuquerque, NM, 1995.
- Dollins, A.B. (1998). A guide to Department of Defense Polygraph Institute research interests. *Polygraph*, 27, 89-95.
- Fuse, L.E. (1982). Directed lie control testing (DLCT). *APA Newsletter*, 1982.
- Gelb, E. (1994). Paper presented at the American Polygraph Association 29th Annual Seminar, Nashville Tennessee.
- Griffith, et al., v. Muscle Improvement, Inc., et al. Superior Court for the State of California for the County of Los Angeles, No. BC215080, 1998.
- Griffith V. Milgaar. Boise, Idaho MB&C, File No. 4502.00, 1996.
- Hofmann, M, Naifeh, S. & White, G.W. (1989). *The Mormon Murders*. Onyx Books.

- Honts, C.R. (1999). The discussion of questions between test repetitions (charts) is associated with increased test accuracy. *Polygraph*, 28(2), 117-123.
- Honts, C.R. & Perry, M.V. (1992). Polygraph admissibility: Changes and challenges. *Law and Human Behavior*, 16, 357-379.
- Honts, C.R. & 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.
- Honts, C.R., Raskin, D.C., & Kircher, J.C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, 79, 152-159.
- Horowitz, W. W., Kircher, J.C., Honts, C.R., & Raskin, D.C. (1997). The role of comparison questions in physiological detection of deception, *Psychophysiology*, 39, 108-115.
- House, G. (15 June 1994). *Albuquerque Tribune*. Albuquerque, New Mexico.
- Idaho v. Kildare. (1984) Prosecutor Keith Roark. Personal communication with author.
- MacDonald, J., & Potter, J. (1995). "Fatal Justice." Norton Press.
- Mak, W. (23 June 1983). *Seattle Times*. Seattle, Washington.
- Montana v. Gordon. Montana Fifth Judicial Court, Jefferson County. Case No. DC-96-1541A, 1997.
- New Mexico v. Martins. (1994). Assistant Attorney General Denise Brown Kulseth. Personal communication with author.
- New Mexico v. Mead. County of San Juan in the District Court. No. CR 92-502-3, 1994.
- New Mexico v. Raebuck. (1983). Prosecutor Bill Bonem. Personal communication with author.
- New Mexico v. Wilson. (1986), County of San Juan in the District Court. Prosecutor's Office. Personal communication with author.
- Wyoming v. Reno. (1985). Roger Cowan of the Cambell County Attorney's Office. Personal communication with author.
- U.S. v. Clayton and Dalley. U.S. District Court, District of Arizona. No. 92-374-PCT-RCB, 1994.
- U.S. v. Freedman. (1983). R. Fishkin of the Miami Strike Force. Personal communication with author.
- U.S. v Gilliard. U.S. District Court, District of Georgia. No 196-19, 1996.

Fiscal Year 1998 Report to Congress on the Department of Defense Polygraph Program

I. DoD Use of Polygraph Examinations

The Department of Defense has used the polygraph for almost half a century. It is used in criminal investigations, counterintelligence cases, foreign intelligence and counterintelligence operations, exculpation requests, and as a condition for access to certain

positions or information. The polygraph is a tool that enhances the interview and interrogation process. Often it is the only investigative technique capable of providing essential information to resolve national security issues and criminal investigations. The use of the polygraph as a condition for access is limited by a statutory quota for counterintelligence scope polygraph (CSP) examinations.

The following table reflects Department of Defense Polygraph Program statistics for Fiscal Year 1998.

Criminal	2,417	22.1%
Exculpatory	537	4.9%
CI Scope	7,461	68.1%
All Others*	539	4.9%
Total**	10,954	100%

* Includes examinations conducted in support of personnel security investigations, counterintelligence and intelligence operations, and polygraph assistance to non-DoD federal agencies.

** Does not include polygraph examinations conducted by the National Security Agency (NSA). Nor does it include polygraph examinations conducted by the National Reconnaissance Office, which are conducted under the authority of the Director of Central Intelligence (DCI).

II. Fiscal Year 1998 Counterintelligence-Scope Polygraph Examinations

Section 1121 of the National Defense Authorization Act for Fiscal Years 1988 and 1989 (Public Law 100-180, December 4, 1987; 101 Stat. at 1147) authorizes the Department of Defense to conduct CSP examinations as a condition for access to certain information.

The purpose of the CSP Program is to deter and detect espionage, sabotage, and terrorism. The following topics are covered during the CSP examination: (1) Involvement with a foreign intelligence/security service,

involvement in espionage; (2) Involvement in terrorism; (3) Unauthorized foreign contacts; (4) Deliberate failure to protect classified information; and (5) Damaging/sabotaging government information systems, clandestine collection, or defense systems. These CSP topics meet the needs of both DoD and the Intelligence Community facilitating the transfer of security clearances.

In Fiscal Year 1998, the Department published a handbook for federal polygraph examiners standardizing techniques and procedures for conducting polygraph examinations. The handbook also outlines the Quality Assurance Polygraph (QAP) wherein DoDPI inspects federal polygraph programs to

ensure compliance with the techniques and procedures taught at the Institute. DoDPI trains all federal polygraph examiners. This allows for standardization and promotes reciprocity, thus eliminating unnecessary polygraph examinations. Memorandums of Agreement are being staffed, by federal agencies that have polygraph programs, to obtain their concurrence with the provisions contained in the handbook.

Public Law 100-180 authorizes DoD to administer CSP examinations to persons whose duties involve access to information that has been classified at the level of TOP SECRET or designated as being within a special access program under section 4.2(a) of Executive Order 12356 (superseded by Executive Order 12958). This includes military and civilian personnel of the Department and personnel of defense contractors. The

number of CSP examinations has been limited to 5,000 per fiscal year since Fiscal Year 1991. For Fiscal Year 1998 through 1990 the ceiling was 10,000. The quota reduction took place two years after new exemptions for cryptographic and reconnaissance programs were adopted. Public Law 100-180 exempts certain intelligence agencies and functions from the 5,000 quota: (1) individuals assigned, detailed or under contract with the Central Intelligence Agency, (2) persons employed, assigned, detailed, under contract or applying for a position in the National Security Agency, (3) persons assigned to a space where sensitive cryptographic information is produced, processed, or stored, and (4) persons employed by, assigned or detailed to, an office within the Department of Defense for the collection of specialized national foreign intelligence through reconnaissance programs or a contractor of such an office.

The following table reflects CSP examinations conducted by the Department of Defense in accordance with Public Law 100-180.

(1)	Special Access Programs	1,680
(2)	DIA Critical Intelligence Positions	1,054
(3)	TOP SECRET	0
(4)	Examinations for Interim Access to Sensitive Compartmented Information	2
Total Examinations Conducted Under the Congressional Ceiling		2,736
Exempted Examinations*		4,725
DoD Counterintelligence-Scope Polygraph Program TOTAL**		7,461

*NOTE: Includes detailees to CIA and NSA, assignees to cryptographic information processing spaces, persons in non-NRO reconnaissance programs.

**NOTE: Does not include polygraph examinations conducted by NSA. Nor does it include examinations conducted by the National Reconnaissance Office (NRO), which are conducted under the authority of the Director of Central Intelligence.

CSP Refusals

In Fiscal Year 1998, nobody declined CSP testing required as a condition of access to certain information. Department of Defense policy states those persons who decline to take

the examination are denied access to the classified material in question, but are retained in their positions or transferred to other positions in the organization of equal pay and responsibility, commensurate with the clearance level held before the declination.

Specific CSP Examination Results

The polygraph examination results for the 7,461 individuals tested under the Department of Defense Counterintelligence-Scope Polygraph Program are as follows:

Two hundred and eight individuals required more than two series (a series is defined as the collection of at least two polygraph charts on an examinee). A total of 72 examinations required more than one day to complete.

There were 7,334 individuals whose polygraph examination results were evaluated as no significant physiological response (non-deceptive).

An additional 110 individuals made admissions relevant to the issues being tested, and through further testing, the examiner was able to resolve all relevant issues favorably to the subject.

After reviewing the psychological data, the polygraph examiner was unable to render an opinion for two individuals. One of these individuals made admissions relevant to the issues being tested.

There were four individuals whose polygraph examination results were evaluated as significant physiological response (SPR) (deceptive) and who made no admissions to the relevant issues.

Eleven individuals made admissions relevant to the issues being tested but continued to be evaluated as significant physiological response (deceptive) during further testing.

Of the 127 individuals whose examination results were evaluated as yielding significant physiological responses, or evaluated as inconclusive and/or provided substantive information, 106 received a favorable adjudication, seven are still pending adjudication, 13 are pending investigation, and one individual received adverse action denying or withholding access.

Significant Information Developed

The following cases reflect significant information developed during DoD CSP examinations covered by this report. It should be noted that all these individuals had been interviewed previously by security professionals and investigated by other means without any discovery of the information obtained by the polygraph examination procedure. In most cases the information was elicited from the subject in discussion with the examiner.

Most of the information developed during CSP examinations relates to the removal of classified material and computer media to residences. Classified material was sometimes commingled with personal papers, and often when discovered was either destroyed or returned to government control. Sometimes the classified material was deliberately taken home to prepare a briefing or to meet a deadline. Classified material was also left in hotel rooms on TDY trips. These types of admissions were followed up with additional polygraph testing to determine whether the material was compromised, or if any other material was still outside of government control.

* * * * *

During examinee's initial CSP examination, he exhibited significant physiological responses to relevant questions concerning the mishandling of classified material. The examinee provided a post-test admission regarding deliberately removing government computer disks from his overseas military assignment and transporting them to his parents' home in his accompanied baggage. Examinee indicated that he had erased all the classified data and had placed personal files, i.e. letters on them. The examinee voluntarily returned the disks to government control. The examinee denied placing the classified disks into any other computer system or using them on a personal home computer. The examinee also denied copying any other classified material to the disks or onto a personal computer. Further investigation determined that the examinee had degaussed the disks prior to returning

them to government control. This examinee refused further interview and requested a lawyer. The intent is to interview the examinee with his lawyer followed by a confirmatory polygraph test.

* * * * *

Examinee exhibited significant physiological responses to a question concerning deliberate mishandling of classified material. He admitted that on orders of superiors, while in Bosnia, he provided Russian military personnel with large quantities of classified material that technically should not have been released to non-NATO members. Also, during numerous planning sessions, he took copious notes and retained some of the notes at his residence. He did not believe the notes were classified. He subsequently returned the notes and computer disks to government control. Upon examination, some of the material was classified SECRET. The examinee subsequently successfully completed a polygraph examination addressing having any other classified material outside government control, or providing classified information to other unauthorized persons. This matter was referred for further investigation.

* * * * *

Examinee exhibited significant physiological responses to a question concerning having a secret relationship with, working with, and providing classified information to a non-U.S. Intelligence Service. During five days of interviews and polygraph tests, examinee disclaimed any unreported contact with a non-U.S. Intelligence Service or representative. Between 1985 and 1997, examinee was involved in HUMINT intelligence activities. He admitted security violations related to HUMINT operations, but denied any unauthorized activity with a foreign intelligence service. The examinee continued to exhibit significant physiological responses to the above topics, and the matter was referred for additional investigation.

* * * * *

Examinee exhibited significant physiological responses to a question concerning deliberately mishandling classified material.

Examinee admitted to keeping SECRET material in his home. He kept the material with him through five military assignments, two of which were overseas. This matter was referred for further investigation.

* * * * *

During CSP testing, examinee, who has access to a Special Access Program, admitted to the removal of classified material during the late 1970s, while he was on active duty. The examinee claimed that he removed the material as a memento of his assignment. He was honorably discharged from the military in 1980 and has maintained the classified material in an insecure unauthorized location. During a search of his residence, classified material was recovered. Examinee denied other removals of classified material or the disclosure of classified material to unauthorized persons. Additional polygraph testing will occur after a security review of the material has been conducted.

* * * * *

During CSP testing, examinee admitted intentionally disabling the taxi lights of an aircraft in order to avoid working a weekend period. In addition, he admitted to fabricating a Maintenance Action Form reporting the taxi lights were inoperable so the pilots would be forced to use another aircraft. Following this admission, the examinee was administered additional CSP testing which he successfully passed. An investigation of the wrongful destruction of government property has been initiated.

* * * * *

During CSP testing, examinee was evaluated as showing significant physiological responses, and subsequently made numerous admissions concerning the disclosure of a classified TDY location to several members of his family and his foreign national girlfriend. He also admitted providing his girlfriend classified information about his specialty, unit mission, and location of duty assignment. A counterintelligence investigation has been initiated.

* * * * *

During CSP testing of a military member assigned to NSA, examinee disclosed that he had been having a romantic relationship with a foreign citizen while on an overseas TDY. He further stated that he continued the relationship even though he was aware that the military would not sanction the relationship. He also stated that he was married at the time. Examinee did not successfully complete his polygraph processing and the information obtained during his examination was provided to his parent military organization. He was denied access to the NSA site.

* * * * *

After unsuccessfully completing his periodic CSP examination, an NSA employee admitted that he suspected his wife may be engaged in intelligence activity against the United States. The examinee met his spouse while he was on a cover assignment overseas. Since returning to the United States, his wife received U.S. citizenship and has been operating a daycare facility. Examinee admitted to providing classified information to his wife, and knowledge of her having a close association with a foreign intelligence service. This case has been referred to the FBI for additional investigation.

III. Utility of the Investigative Polygraph

During Fiscal Year 1998, DoD investigations obtained unique and significant information from interviews conducted with the aid of the polygraph. In all illustrated instances, the polygraph examination process produced significant security or criminal information that would not otherwise have been secured for the specific investigation. The polygraph examination process was also valuable in helping to establish the innocence of persons charged with serious infractions.

* * * * *

During a background investigation for a DoD contractor employee, information was received alleging that the employee sexually abused children, assaulted his former wife, raped a female neighbor, and misused prescription drugs. The employee denied the

allegations and agreed to take a polygraph examination in support of his denials. During the pre-test interview, the employee admitted to the use of illegal drugs, abusing prescription drugs, committing the statutory rape of a minor female, sexually abusing two daughters of his current spouse, attempting to kill his former spouse, and attempting to commit suicide. This matter has been referred to appropriate law enforcement agencies.

* * * * *

An investigation was initiated when a buyer for a large aircraft manufacturer told investigators of accepting money from subcontractors in return for orders. The buyer named eight companies and individuals that had paid him kickbacks. The buyer consented to confirming his story by polygraph. The buyer passed the polygraph and began cooperating in the investigation. As a result, polygraph examinations were conducted on 18 individuals. Three of these individuals were found truthful and 15 were found deceptive. Fourteen of these individuals made admissions of wrongdoing and have implicated other individuals and companies in various kickback schemes. As of this time, there have been 31 guilty pleas resulting in convictions in federal court and \$2.8 million in fines and restitution.

* * * * *

During a periodic polygraph examination, an SCI cleared contractor reported that he had "hacked" approximately 1,000 times into a computer that his company used for an NSA project. Access was obtained by writing unauthorized scripts that bypassed the firewall protection. The contractor employee also reported that his employer allowed a non-U.S. citizen to access an NSA SCIF to perform maintenance on a computer used on an NSA contract. The contractor successfully completed additional polygraph testing confirming the information he reported.

* * * * *

During a routine after duty hour's inspection of a NATO facility, classified documents were discovered in a maid's storage room. The maid, a host country civilian

contractor, was interviewed with the help of an interpreter and acknowledged having the paper in the storage area. The maid denied knowing about their sensitivity. She explained that she could neither read nor understand English and that she had obtained the papers from the general trash. Further investigation corroborated the maid's story. It was determined that there was extreme lack of security at the facility and that numerous classified documents had been previously tossed out in the unclassified trash. The maid consented to undergo a polygraph examination to verify her statements. The examination was conducted with the assistance of an interpreter. The maid successfully passed the examination.

* * * * *

During a periodic CSP examination, an employee of an NSA SCI contractor reported information about two foreign nationals who were trying to obtain "extremely proprietary" information for a foreign government. The employee stated that the foreign nationals were able to obtain information relating to computer codes hardware design, engineering design specifications, and sample chips for cloning. Subsequently the employee successfully completed the polygraph examination.

* * * * *

During a periodic polygraph examination of an NSA cleared contractor, the examinee admitted compromising classified information to representatives of a foreign military. The classified information was improperly released during an approved training course. Prior to releasing the information, the contractor was specifically told not to release the information. The contractor employee released the classified information because he felt it was fundamental to the operation of the system for which they were receiving the training. The employee also provided unauthorized classified information to another foreign government official while answering technical questions.

* * * * *

An investigation was initiated when a mother reported that she entrusted her two children into the care of an Army member and

his wife to babysit for a two-week period and upon her return she was told by her four-year-old daughter that the soldier had sexually assaulted her. The soldier denied any sexual contact and agreed to take a polygraph examination. After being evaluated as deceptive, the soldier admitted to sexually fondling the four-year-old daughter.

* * * * *

An investigation was initiated when an Army member reported the theft of his 1995 Jeep Wrangler valued at \$8,000, alleging it was stolen from an on-post parking lot. The vehicle was later discovered destroyed by fire. The stereo equipment, valued at over \$1,400 had been removed from the vehicle and an accelerant had been used to ignite the vehicle. There were no signs of forced entry to the vehicle and a gas can was discovered in the vehicle. The owner was administered a polygraph examination and was evaluated as deceptive. He subsequently admitted to the intentional arson of the vehicle by his wife and a friend in order to defraud both the government and an insurance company.

A background investigation was conducted on an Army member who was born in Armenia and has relatives living there. The member previously served in the Soviet Army as a doctor. While on assignment at Fort Benning, Georgia, he reported having contact with two Russian Army officers who had been attending training at Fort Benning. His background investigation could not be completed, so the Army member agreed to undergo a polygraph examination to verify his statements regarding his contacts with foreign nationals or any foreign intelligence service. The Army member successfully completed his polygraph examination.

* * * * *

A polygraph examination was requested to determine an Air Force civilian employee's involvement, if any, in the theft of a \$12,000 computer system from the base golf course. The computer system was taken from a building located inside a locked maintenance area. The investigation developed a suspect and he agreed to take a polygraph examination. The polygraph examination indicated

deception. The suspect subsequently admitted stealing the computer system and selling it for \$150. The investigation has been referred to local law enforcement authorities.

* * * * *

An Air Force civilian employee, who occupied a position with access to sensitive classified information, was removed from that position for suspicious behavior relating to unexplained absences from work. During the pretest interview, examinee admitted to an addiction to crack cocaine. The examinee agreed to under polygraph testing to determine if she had compromised any classified information entrusted to her. The examinee successfully completed her polygraph testing with no deception indicated.

* * * * *

During an NSA applicant polygraph examination, the examinee admitted that she had a friendship with a foreign intelligence officer. She advised that she met the intelligence officer at a math conference in San Diego, California and began a six-month relationship with him. The intelligence officer told the applicant that she should not mention anything about the intelligence officer since NSA would not hire her if the relationship became known. The applicant's testing indicated specific physiological responses and she subsequently admitted that the relationship was sexual in nature and that she has continued contact with the intelligence officer.

* * * * *

An investigation was initiated based on allegations that a government contractor had submitted false statements and progress reports on two government contracts resulting in the payment of over \$2.5 million on an Army contract and \$3.1 million on a Navy contract. Additionally, the government paid for supplies in excess of \$30,000. The systems manager of the firm was administered a polygraph examination after denying any knowledge or participation in the fraud. The polygraph examination results were evaluated as deception indicated. The manager subsequently admitted to receiving kickbacks,

disposal of the property in question, and filing fraudulent claims and progress reports.

* * * * *

An investigation was initiated when a United Services Organization (USO) in Germany discovered the theft of over \$6,000. The civilian manager reported the theft, and indicated that 30-50 people had access to the funds. The manager agreed to take a polygraph examination. The results indicated deception. The manager subsequently admitted that she had a gambling problem and had stolen the money to support that habit. The investigation is still pending.

* * * * *

A baby was admitted to the hospital with a broken leg. The father stated that he broke the baby's leg accidentally when the baby fell from the father's lap. The father agreed to take a polygraph examination. The results of the polygraph examination indicated deception. During the posttest interview, the father admitted to physically abusing the child.

* * * * *

An investigation revealed deliberate tampering with the window in the nose landing gear housing on an operational C141 aircraft. After numerous interviews, two individuals were identified as suspects. Both suspects agreed to take a polygraph examination. One of the suspects was evaluated as nondeceptive on his polygraph examination. The other suspect was evaluated as deceptive on his examination and confessed to removing screws from the window of the nose landing gear housing.

* * * * *

Local law enforcement authorities requested investigative assistance after a Navy member was discovered dead in the Norfolk, Virginia area. The Navy member died as a result of several bullet wounds. A suspect was identified and he agreed to take a polygraph examination. The results of the polygraph examination indicated deception. The suspect subsequently admitted to being a passenger in a vehicle when another individual in the

vehicle shot and robbed the Navy member. The suspect also admitted to helping dispose of the body.

IV. Training and Qualification Standards for Department of Defense Forensic Psychophysiology (Polygraph Examiners)

The Department of Defense maintains very stringent standards for polygraph examiners. The 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 DoD polygraph positions 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. Two years of comparable experience may be

substituted for the requirement of investigative experience 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 DoD-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 DoDPI. In Fiscal Year 1998, the Institute trained 59 new polygraph examiners. After completing the basic polygraph training, DoD personnel must serve an internship consisting of a minimum of six months on-the-job training and the conduct of at least 25 polygraph examinations under the supervision of a certified polygraph examiner before being certified as a DoD polygraph examiner. In addition, DoD polygraph examiners are required to complete 80 hours of continuing education every two years. To help meet this requirement, the Institute offers 19 different specialized courses in forensic psychophysiology. In Fiscal Year 1998, approximately 550 students attended the specialized courses.

Department of Defense Forensic Psychophysiology (Polygraph Examiners)

Fiscal Year	Average Number of Examiners	Attrition Rate
1994	192	19%
1995	176	18%
1996	164	19%
1997	153	18%
1998	147	15%

V. Polygraph (Forensic Psychophysiology) Research

Mandated by Congress, the research program at the Institute is focused on: (1) developing new psychophysiological detection of deception (PDD) techniques, instrumentation and analytic methods to improve PDD technology; (2) conducting research on PDD countermeasures; and (3) evaluating the validity of PDD techniques.

To facilitate the research, a small grant program was established in Fiscal Year 1992. In Fiscal Year 1998, the Institute funded four grant proposals. An additional proposal is currently under review. Another two proposals were funded by contract.

During Fiscal Year 1998, the Institute developed a prioritized research plan and presented it to the Security Policy Board. This plan describes a series of projects to be completed in support of PDD research. Its successful completion is dependent on the availability of resources. The plan has been approved in its entirety by the Personnel Security Research Subcommittee.

The Institute began listing research abstracts on the DoDPI World Wide Web page in Fiscal Year 1998 to make them more available to the public. The site includes links to the Defense Technical Information Center (DTIC) where the full text of the reports can be obtained. DTIC took over the publication of DoDPI research reports in Fiscal Year 1998.

Research Projects

Oculomotor and Pupil Analysis for PDD. Studies are ongoing to examine changes in pupil size and eye movement during a PDD examination. Some of the studies indicate changes in pupil diameter can be used to detect deception.

Improvement of the Johns Hopkins University Applied Physics Laboratory Automated PDD Examination Scoring System (POLYSCORE). The Institute has contracted with the Johns Hopkins University Applied Physics Laboratory to update and improve their POLYSCORE computer program.

POLYSCORE is a computer program designed to evaluate data collected during a PDD examination. Several data analysis techniques have been evaluated and those which improve POLYSCORE's accuracy have been incorporated into a new version. POLYSCORE continues to be evaluated so that additional enhancements can be made.

Detecting Stress in the Voice. The Institute, in collaboration with the Chief, Department of Neuroendocrinology and Neurochemistry, Division of the Neurosciences, Walter Reed Army Institute of Research, is measuring the human voice during stressful circumstances to determine if there are characteristic changes associated with stress. Although the biological measures were shown to be reliable indicators of human stress, there was no correlation between those measures and measures provided by a computer voice stress analyzer (CVSA), a device currently being used by many non-federal law enforcement agencies. The utility of other voice stress analytic techniques is not ruled out.

Vagal Tone Monitor/ARIS. This project was designed to investigate the feasibility of using a Vagal Tone Monitor and Autonomic Response Indicator System (ARIS) software to monitor changes in cardiovascular activity during a PDD examination. The Vagal Tone Monitor and ARIS software measure the direct influence of the vagal nerve on heart rate. Data collection has been completed and an analysis is underway.

A Field Study to Test the Validity and Comparative Accuracy of Voice Stress Analysis as Measured by the Computerized Voice Stress Analyzer: In a Psychological Context. Funds were awarded to an investigator employed by the Michigan State Police Polygraph Unit to assess the validity of the computerized voice stress analyzer using subjects who are being tested for actual crimes. Data collection has been completed for his project and a final report is expected in Fiscal Year 1999.

Thermal Imaging During a PDD Examination. This project is designed to examine the efficacy of thermal imaging technology as a measure of deception.

Infrared thermal imaging, a non-intrusive and non-invasive technology, will be used to determine if peripheral changes in skin surface temperature occur during a PDD examination, and if such changes are indicative of deception.

Test of Mock Theft Scenario Incorporating Judgment, Choice, and Consequences. This study is another in a series of research projects being conducted to delineate a standardized scenario for laboratory studies. The goal is to implement a reliable standardized scenario for PDD testing in laboratory studies to further the development of new and improved technology and procedures for PDD testing. A research protocol has been drafted and submitted for peer review.

Polyscore 3.3 and Psychophysiological Detection of Deception Examiner Accuracy Rates when Scoring Examinations from Actual Criminal Investigations. This study was designed to examine polygraph and human examiner accuracy rates using data collected during actual criminal investigations. The final report is being reviewed prior to publication.

Rank Order Assessment. Two computerized PDD scoring algorithms and two trained examiners using a ratio-based manual scoring process were evaluated. The report will be submitted for final review prior to publication in Fiscal Year 1999.

Effects of Prior Demonstrations of Polygraph Accuracy on Outcomes of Probable Lie and Directed Lie Polygraph Tests. Funds were awarded to investigators at the University of Utah to examine the usefulness of administering an acquaintance test during a PDD examination. A known solution peak of tension, or acquaintance, test is used to demonstrate to examinees that the PDD procedures can accurately detect deception. In addition, the investigators will compare the accuracy of tests administered using directed, versus probable, lie comparison questions. Data collection has begun and a final report is anticipated in Fiscal Year 1999.

Validity of Outside Issue Questions in the Comparison Question Test. Outside issue questions are those which address topics that are not included in the relevant and comparison questions (i.e., Do you believe I will only ask you the questions we reviewed?, Is there something else that you are afraid I will ask you questions about?) Investigators at Boise State University have proposed, and have been funded, to determine the validity of outside issue questions. A final report is anticipated during Fiscal Year 1999.

The Frequency of Appearance of Evaluative Criteria in Polygraph Charts. The Institute teaches that 23 unique physiological responses may be predictive of deception during a PDD examination. Investigators at Forensic Research Inc. have been funded to determine the frequency and predictive value of these criteria in actual PDD examinations. Analysis has begun and a final report is anticipated during Fiscal Year 1999.

Evaluation of DoDPI Evaluation Techniques. Several laboratory studies, published by the University of Utah, have reported high accuracy rates during PDD chart evaluations. Data from the University of Utah studies are currently being evaluated by federal examiners. A final report comparing the decision accuracy of the university and federal examiners is anticipated in Fiscal Year 1999.

A Comparison of Decision Accuracy Rates Obtained by Computer Programs Designed to Evaluate Examination Data. Four vendors currently sell computer programs purported to evaluate accurately PDD examination data. The Institute is currently conducting an independent evaluation of these computer programs and a report is anticipated in Fiscal Year 1999.

A Compilation of Studies on the Effectiveness of Event-Related Stimuli as a Control Procedure in the Psychophysiological Detection of Deception (PDD). A typical PDD examination involves the use of two or three different types of questions. The two types of question that are evaluated are the comparison and the relevant questions. The relevant questions are those regarding the specific issue during the examination (i.e., Did you steal the money?) The comparison

questions are similar in nature to the issue. Responses from the comparison questions are compared to those of the relevant. The event-related test uses only relevant questions. This was designed to evaluate the effectiveness of PDD examinations that only use relevant questions. Work on this project continues in Fiscal Year 1999.

Exploration of Manual and Automated Scoring Methods for Relevant/Irrelevant Multiple Issue Screening Examinations.

The Relevant/Irrelevant (RI) PDD screening format is used by some federal agencies for applicant screening and employee vetting. The development of an objective means of interpretation of the polygraph recordings in this role is the goal of this project. Live cases conducted by a contractor for which ground truth was established independently have been subjected to automated and human scoring systems. The data are being evaluated to determine which series of decision rules maximize decision accuracy. The outcome will help identify the best means of interpretation for these types of polygraph data. A report is expected in Fiscal Year 1999.

Pretest Interview Project. DoDPI is currently designing research to identify important variables in the interview that precedes PDD examinations, the pretest interview.

Effects of Question Phrasing, Complexity and Length on Examinee Performance: Development of a Comprehensive Index.

The Institute is currently evaluating a proposal to study the effects of question phrasing, complexity and length on PDD examination decision accuracy in order to create a comprehensive index for future use.

Phasic Electrodermal Reactivity to Equivalent Psychological Stimuli Presented During Varying Electrodermal Baseline Levels. The electrodermal response measured by most analog polygraphs is skin resistance

while the response measured by most digital instruments is skin conductance. It is not known which, if either, measure provides a more accurate presentation of the electrodermal responses to equivalent stimuli at different electrodermal baseline levels. Proposals to resolve this question will be considered during Fiscal Year 1999.

Other Activities

International Use of PDD. The Institute maintains contacts with PDD examiners in other countries to keep abreast of polygraph developments around the world. The Institute issues periodic reports summarizing international polygraph activity.

A Computer Program for Generating and Modifying Polygraph Charts.

DoDPI instructors currently use numerous polygraph diagnostic features in the tracings. It is recognized that instruction would be more effective if the characteristics of the tracings could be modified by instructors during class to show various forms and combinations of features. Such a capability would better prepare students to evaluate the complex psychological patterns they will later encounter when they perform PDD tests in the field. DoDPI is developing a statement of work for a system that will allow instructors to tailor tracings in real time for this purpose.

Presentations. DoDPI researchers presented results of their research in several fora in Fiscal Year 1998, including the Federal Interagency Polygraph Seminar, the annual seminar for the Society for Psychophysiological Research, the annual seminar for the American Polygraph Association, and the International Organization of Psychophysiology. The researchers have also provided formal instruction to federal examiners at the course sponsored by the University of Virginia, and the basic and advanced courses at DoDPI.

The Relative Utility of the Forensic Disciplines

Gary D. Light and John R. Schwartz

Abstract

The efficacy of the forensic disciplines in felony criminal investigations was assessed. Reports and investigations of the findings of 1,069 forensic examinations reviewed involved 920 felony investigations conducted between 1 July and 30 December 1990 by the United States Army Criminal Investigation Command (USACIDC). The traditional laboratory disciplines combined conducted 584 (55%) and the psychophysiological detection of deception (PDD) discipline conducted 485 (45%) of the examinations. The PDD discipline provided the investigator with 432 (89%) opinions that contained positive results and the laboratory disciplines provided positive results in 431 (74%) examinations. In all categories assessed, regardless of type of crime, a higher solve rate was achieved for USACIDC when multiple forensic disciplines were utilized. The PDD discipline was the most utilized and effective of the individual disciplines, but all forensic disciplines demonstrated a high degree of utility in specific criminal offense categories. Of the 1,069 examinations reviewed, there were no instances in which the findings of one discipline contradicted the results of any other discipline.

Key words: forensic disciplines, forensic psychophysiology, polygraph, psychophysiological detection of deception (PDD), utility

The support of the criminal investigator through the use of scientific evidence continues to be a multidisciplinary approach encompassing law, science, and technology (Moenssens, Inbau, & Starrs, 1986). The forensic sciences have traditionally provided the criminal investigator with expert opinions to provide links between the suspect and the crime. The forensic disciplines have been instrumental in resolving criminal investigations. However, a paucity of research exists which provides insight as to the effectiveness and utility of these forensic tools. The purpose of this study was to investigate the impact of these disciplines on criminal investigations utilizing forensic reports of examinations that supported field investigations.

An important concept in employing any forensic discipline is that the findings must be

of such a recognized standard to be able to withstand the scrutiny inherent in the adversary procedures of the U.S. Criminal Justice System (Moenssens et al., 1986). Much emphasis is placed upon the use of forensic disciplines in the courtroom environment. In fact, most texts, in discussing the forensic sciences, delve extensively into "expert testimony, rules of evidence, etc...." It is understood that the utilization of forensic findings in a court of law is an important end product of any forensic examination. However, the courts are not the primary users of these findings.

The criminal justice system has evolved into a process in which the vast majority of criminal offenses involving criminal suspects are adjudicated prior to the onset of actual courtroom proceedings (Cole, 1983). This

Acknowledgments

The authors wish to express special thanks to Wilbur Hardy, Betty Dempsey, and Samantha Edgar, U.S. Army Crime Records Center, and Bob Thibault, U.S. Army Criminal Investigations Laboratory (USACIL) Continental United States (CONUS), for their assistance throughout data acquisition. Also, the authors wish to thank Dr. Bill Yankee, Department of Defense Polygraph Institute, for his expert guidance and insight. This research was supported by DoDPI93-P-0027 project funds from the Department of Defense Polygraph Institute. 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.

process requires findings based upon information that speaks to the trier of the fact (whomever has assumed this role) in a clear, concise, and expeditious manner. The traditional laboratory disciplines (as detailed later) in the analysis of physical evidence, provide the trier of fact in some instances with real evidence which speaks for itself to relate an impression upon the mind of the trier of fact (Donigan, Fisher, Hugel, Reeder & Williams, 1980). In other cases, the laboratory disciplines provide circumstantial evidence with which an inference can be logically drawn from the known facts.

Forensic psychophysiology (Yankee, 1992), which incorporates the psychophysiological detection of deception (PDD), is included in this research to provide a comprehensive review of all the forensic support utilized by the USACIDC field element. The PDD¹ discipline is unique in that the nature of the findings and its by-products (confession/admission) are circumstantial (in the form of expert opinion), allowing the trier of fact to make an inference about the involvement of the suspect in the offense (Donigan et al., 1980). However, in most instances (72%), when deceptive findings are rendered, they are associated with legally sufficient admissions or confessions that provide corroboration for the PDD examination and testimonial evidence for the trier of fact.

Forensic disciplines have resolved countless criminal investigations. The present research was designed to be a comprehensive review of specific forensic disciplines over an extended time period in an attempt to ascertain the utility of each forensic discipline for the field investigative element. Further, the research was designed to demonstrate the impact the findings of disciplines had on the

types of crimes for which the disciplines are utilized.

Method

This research study involved a total of 1,077² forensic examinations conducted between 1 July and 30 December 1990. These examinations were conducted subsequent to requests for forensic support from USACIDC special agents. This study utilized a total of 584 laboratory forensic examinations which represented all analyses completed at the United States Army Criminal Investigation Laboratory (USACIL), Ft. Gillem, Georgia during the applicable time period. The laboratory system of USACIL has three laboratories supporting the U.S. Army worldwide. The USACIL laboratory supporting the Continental United States (CONUS) was utilized for this study.

During this same time period, 493 PDD examinations were completed by USACIDC PDD examiners in support of the USACIDC criminal investigative mission in CONUS. All PDD examination reports are maintained at the United States Army Crime Records Center (USACRC), Baltimore, MD. All reports reflecting analyses completed by the forensic disciplines are retained at the United States Army Crime Records Center. The case folders for these forensic reports (and all investigative efforts) are maintained by USACIDC and are referred to as a Report of Investigation (ROI). All 1,069 forensic reports utilized for this research were obtained from these ROIs (case folders) and reflect all forensic reports that could be located for this time period investigated within CONUS by USACIDC.

A review of all forensic examinations identified eight primary forensic disciplines

¹ For the purposes of this paper the term PDD will be used to refer to a body of techniques that is popularly known as polygraphy. The term polygraph, historically has engendered the perception of a mechanical method utilized for detecting deception. However, PDD is deemed appropriate as a more accurate description of the psychophysiological processes involved in the procedures utilized by the United States Army Criminal Investigation Command in support of their criminal investigative mission.

² The number 1,077 represents the computer-generated total of forensic examinations conducted during this time period in which results could be determined utilizing the database. However, since some investigations initiated in 1990 have not yet been closed, only 1,069 of the reports could be located.

that were utilized in support of the USACIDC criminal investigative mission. The following eight disciplines were identified:

Firearms
Illicit Drugs
Latent Prints
Psychophysiological Detection of Deception
Questioned Documents
Serology
Trace Evidence
Photographic

A description of each of the forensic disciplines utilized for this study is provided in Appendix A.

The results provided to the investigative field elements from these eight disciplines were evaluated utilizing the following topics: The "Utility of the Forensic Principles"; the "Impact of the Forensic Disciplines"; and an "Interdisciplinary Comparison of the Results of the Forensic Disciplines."

Utility of the Forensic Disciplines

This analysis of the forensic disciplines identified how often the individual disciplines were utilized and the quantification of positive information provided to the investigative field element. This analysis involved three steps. (1) The number of forensic examinations completed in support of the USACIDC field element. (2) The impact each report had upon the investigation. This was determined by categorizing the results of those forensic analyses into the following categories:

Positive Result
Negative Result
Positive Result With Subject
Positive Result With No Subject

For a detailed discussion of the process for determining these outcomes refer to Appendix B.

(3) The percentage of each forensic discipline that met the criteria for the above listed categories.

Impact of the Forensic Disciplines

This analysis of the forensic disciplines determined if a correlation existed between solve rates of felony investigations and the use of the forensic disciplines. This analysis involved: (1) identifying investigations which had been solved wherein a forensic discipline had been utilized (a discussion of the process for determining the resolution of an investigation is contained in Appendix C) and, (2) determining the number of instances each forensic discipline was involved in the resolution of an investigation based on the category of the crime. The categories of crime utilized were: crimes against property, crimes against persons, and drug investigations.

Interdisciplinary Comparison of the Results of the Forensic Disciplines

This analysis determined if reliability existed in those investigations in which a PDD examination and a laboratory examination were completed during the same investigation. The comparison of the results rendered by PDD and other disciplines was made since the USACIDC Polygraph Program is headquartered at USACRC, Baltimore, MD, while the USACIL is located at Ft. Gillem, GA. Laboratory reports are generated at the USACIL, while PDD reports are initiated by the field examiner who is conducting the PDD examinations in the geographical area where the suspect is located. The PDD report is subsequently forwarded to the quality control section of the USACIDC Polygraph Program, USACRC, where it is filed pending receipt of the final ROI at USACRC.

The review of all 32 ROIs (case folders) used in the interdisciplinary comparison affirmed that the laboratory results were not known when the PDD examinations were completed. In all instances, laboratory discipline examinations were completed weeks after PDD examinations. Additionally, the review of all requests for laboratory examinations at USACIL did not reflect the fact that a PDD examination had been completed. The 32 ROIs that met this criteria were reviewed to determine the number of instances in which contradictions between the results of the PDD and other disciplines occurred.

Results

Table 1. Utility of the Forensic Disciplines

Discipline	Positive Results	Positive Results With Subject	Positive Results Without Subject	Negative Results	Total #s Percentage
PDD	432 (89%)	289 (60%)	143 (29%)	53 (11%)	485 (100%)
Latent Fingerprints	90 (59%)	31 (20%)	59 (38%)	64 (42%)	154 (100%)
Questioned Documents	105 (72%)	66 (45%)	39 (27%)	40 (28%)	145 (100%)
Illicit Drugs	125 (93%)	106 (80%)	19 (14%)	8 (6%)	133 (100%)
Firearms	39 (76%)	14 (27%)	25 (49%)	12 (24%)	51 (100%)
Trace Evidence	33 (65%)	14 (28%)	19 (37%)	18 (35%)	51 (100%)
Serology	34 (85%)	20 (50%)	14 (35%)	6 (15%)	40 (100%)
Photographic	5 (50%)	1 (10%)	4 (40%)	5 (50%)	10 (100%)

The Utility of the Forensic Disciplines

A total of 1,069 forensic examinations were reviewed during this portion of the study. The PDD discipline conducted 485 (45%) of all forensic examinations completed, while the latent fingerprint and the questioned document disciplines were each utilized in 14% of the examinations.

The individual disciplines demonstrated that "positive results" were provided to the field elements most often by the illicit drug discipline (93%), followed by PDD (89%), and serology (85%). A "positive result with subject" was demonstrated most often by the illicit drug discipline (80%), followed by the PDD

discipline (60%). The latent fingerprint discipline provided "positive results with subject" in 20% of the examinations provided to the field element.

In the "positive results without subject" category, the firearms discipline provided a laboratory report meeting this criteria in 49% of the laboratory examinations, while the latent fingerprint discipline provided a report in this category in 38% of the examinations. In the "negative results" category, the latent fingerprint discipline provided 42% of the reports in this category. Trace evidence provided "negative results" in 35% of the reports in this category.

Table 2. Impact of the Forensic Disciplines

Total Exams	Laboratory Exams	PDD Exams	USACIDC Solve Rate	Inter-Discipline Solve Rate	Non-Discipline Solve Rate	PDD Solve Rate	Laboratory Solve Rate
914	421	493	81%	86%	78%	82%	81%

Impact of the Forensic Disciplines

In reviewing all case folders available at the USACRC, 914 forensic examinations were identified for use in this study. The 914 case folders (ROIs) were located to determine if the investigation had been solved. As a result, 421 examinations completed by USACIL and 493 of the PDD examinations completed by the USACIDC were located.

In this study an investigation is determined to be resolved when a person is titled in the investigation. (See Appendix C for detailed discussion of this process.) The frequency with which crimes are resolved will indicate the effectiveness of the investigative

process. The ratio for solved investigations is determined by dividing the number of investigations resolved by the number of investigations completed. This solve rate will indicate the effectiveness of the forensic disciplines. The interdisciplinary solve rate (ROIs involving both PDD and laboratory examinations) was 86%. The solve rate for investigations involving only the laboratory disciplines was 81%. The solve rate for investigations involving only the PDD discipline was 82%. The overall USACIDC solve rate (all investigations) was 81%. The solve rate in which USACIDC utilized no forensic discipline during an investigation was 78%.

Table 3. Crimes Against Property Solve Rate

Crimes Against Property		Laboratory Exams		PDD Exams		Latent Finger Prints		Questioned Documents		All Other Laboratory Disciplines	
# of Exams	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate
473	71%	234	75%	239	67%	95	69%	102	77%	37	80%

The impact of the forensic disciplines upon specific categories of crimes is as follows:

Crimes Against Property

The solve rate for all USACIDC investigations which occurred during the applicable time period in this category was 56%. The solve rate for those investigations in which a forensic discipline was utilized was 71%. The solve rate for those investigations in

which a laboratory discipline (completed at USACIL-CONUS) was involved was 75%, while the PDD discipline was 67%.

In reviewing the ROIs (case folders) in this category involving only solved crimes in which a forensic discipline was utilized, a total of 334 examinations were identified. The forensic disciplines of PDD, questioned documents, and latent fingerprints were

involved in solving 91% (304 of 334) of those examinations. PDD was involved in resolving 47% (159 of 334 examinations) of the property crimes. Questioned documents were involved in 24%. (79 of 334 examinations) of those

solved investigations. Latent fingerprints were involved in the resolution of 20% (66 of 334) of those investigations. The remaining disciplines combined were involved in 9% (30 of 334) of the remaining investigations.

Table 4. Crimes Against Persons Solve Rate

Crimes Against Persons		Laboratory Exams		PDD Exams		Latent Finger Prints		Trace Evidence		Serology	
# of Exams	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate
309	93%	124	91%	185	94%	30	90%	30	97%	26	90%

Crimes Against Persons

The solve rate for all USACIDC investigations that occurred during the applicable time period in this category was 89%. The solve rate for those investigations in which a forensic discipline was utilized was 93%. The solve rate for those investigations in which a laboratory discipline (completed at USACIL-CONUS) was involved was 91%, while the PDD discipline was 94%.

In reviewing the case folders in this category involving only solved crimes in which the forensic disciplines were utilized, a total of 282 examinations were located. The

disciplines of PDD, trace evidence, latent prints, and serology were involved in 264 (94%) of the examinations. The PDD discipline was utilized in 62% (175 of 282 examinations) of those crimes against persons. The trace evidence discipline was utilized in 10% (29 of 282 examinations) of those investigations. Latent fingerprints were involved in the resolution of 10% (27 of 282 examinations) of those investigations. The serology discipline was utilized in 9% (26 of 282 examinations) of those investigations. The remaining disciplines combined are responsible for the resolution of 9% (25 of 282 examinations) of the investigations.

Table 5. Illicit Drugs Solve Rate

Total Exams: Illicit Drugs	Laboratory Exams		PDD Exams	
	# of Exams Conducted	Solve Rate	# of Exams Conducted	Solve Rate
132	63	100%	69	97%

Illicit Drugs

The solve rate for all USACIDC investigations which occurred during the applicable time period in this category was 99%. The solve rate for those investigations in

which a forensic discipline was utilized was 99%. The illicit drug discipline was utilized in 63 investigations and all investigations were solved. The PDD discipline was utilized in 69 examinations and 67 were solved.

Interdisciplinary Comparison

In reviewing all ROIs at the USACRC, Baltimore, MD, during the applicable time period for this category, a total of 32 investigations were identified in which at least one laboratory forensic examination (completed at USACIL-CONUS) and one PDD examination was completed during the same investigation. Further, 47 laboratory examinations were conducted in support of those 32 investigations, while the PDD discipline completed 32 examinations in support of the same 32 investigations. In comparing the results of all 79 forensic examinations identified in this study, in no instance did any forensic finding contradict another discipline.

Of the 47 laboratory examinations completed, 33 rendered a conclusive opinion in support of 25 investigations. Of the 32 PDD examinations, 29 resulted in conclusive opinions being rendered. A total of 24 investigations were identified in which conclusive opinions were rendered by PDD and the laboratory disciplines in the same investigation. In all 24 investigations, the laboratory and PDD rendered the same opinion pertaining to the same subject of that investigation.

In considering the impact of the PDD and laboratory disciplines in resolving these investigations it was found that in 20 ROIs a subject was identified by the laboratory discipline. Of the 29 positive results obtain by the PDD discipline, 26 of the examinations resulted in a subject being identified. The PDD discipline rendered 3 no opinion findings, in comparison to the 14 no opinion examinations rendered by the laboratory disciplines.

As demonstrated, in 24 out of 32 ROIs, PDD and the forensic disciplines concurred in positive findings and a subject was titled in those investigations. In 20 of the 32 investigations, confessions/admissions were obtained during the course of the investigations. The confessions/admissions were obtained in 19 of the 20 PDD examinations and one in support of the laboratory examinations without a PDD confession. It should be noted that in 2 of the 24 ROIs the laboratory and PDD disciplines concurred that

an individual was not involved in the incident. The validity of the PDD opinions was confirmed in all 24 ROIs. In 20 of the 24 ROIs both a laboratory discipline and a confession confirmed the PDD result and in the remaining 4 ROIs the PDD opinion was confirmed by the laboratory discipline.

Discussion

To our knowledge, this study represents the first comprehensive research effort in which actual field data have been utilized to compare the effectiveness and utility of the findings routinely provided by a major crime laboratory in support of felony investigations. The effect of the forensic disciplines on solving investigations held consistent throughout this study. The forensic disciplines were found to be more (or less) effective based upon the category as well as the specific type of crime involved. For example, the illicit drug discipline was extremely effective in drug investigations, but was not utilized in crimes against property or crimes against persons investigations. The questioned documents discipline was seldom used in crimes against persons investigations, and the serology discipline was seldom appropriate in crimes against property investigations. In contrast, PDD was used effectively in every type of crime that was investigated. For instance, in the 24 child abuse cases included in this study, PDD was the only forensic discipline utilized. The findings indicate that if both PDD and the laboratory forensic disciplines are utilized, the solve rate for the investigative field element is significantly increased. When the forensic disciplines are utilized separately in crimes against property, the laboratory disciplines have a significant positive impact in increasing the USACIDC solve rate. In crimes against persons, the solve rate for USACIDC was increased when PDD was utilized exclusively.

The findings of greater utility with certain disciplines based upon the examination and analysis of various physical evidence types is consistent with the findings of Widacki and Horvath (1978). When utilizing an analog study, they found a 100% accuracy rate with fingerprint examinations but found that fingerprints could only render a positive opinion with subjects in 20% of examinations.

Widacki and Horvath also found PDD and handwriting examinations resulted in a high accuracy (90% and 85%, respectively) while rendering a positive result with subjects in 95% and 94% of the examinations. The fingerprint, handwriting, and PDD results were consistent with the results in this study. In Widacki and Horvath (1978), the examiners in all forensic disciplines were experienced, well trained, and employed standardized procedures. Other research (Raskin & Podlesney, 1979; Patrick & Iacono, 1988) also found the validity of PDD to be over 90%, when professionally trained PDD examiners utilizing standardized procedures conducted a PDD examination.

The Laboratory Proficiency Testing Research Program (1979) demonstrated that certain laboratories that lacked trained personnel and suffered from budgetary constraints produced significant "unacceptable responses" in examinations involving blood and paint **samples**³. The present research project identified no discipline in which an error by a laboratory discipline occurred. However, there were a significant number of no opinion and "positive without subject" findings, which is also consistent with the findings of Widacki and Horvath (1978).

The Laboratory Proficiency Testing Research Program, under certain conditions, recognized an inconclusive opinion as an error. This research project did not consider an inconclusive opinion as an error. The use of an inconclusive opinion allows the forensic professional the right to say, "I don't know" (Willard, 1982). Without this ability, forced erroneous opinions would result.

PDD appears to be the most robust of all forensic disciplines due to its applicability to more types of criminal investigations. PDD examinations alone accounted for 45% of all forensic examinations. The findings of utility with the laboratory disciplines are consistent with crime resolution. The disciplines are most often in those types of crimes in which they have higher solve rates. Therein lies the rationale for the significantly higher utility of

PDD. The laboratory disciplines depend on the existence of traditional physical evidence that can be examined and analyzed. These circumstances do not normally pose a problem for the forensic psychophysicist. A PDD examination is conducted based upon the individual's concealed knowledge of the criminal event -- unlike the requirements physical evidence impose on the laboratory disciplines. With PDD, the evidence linking the suspect to the crime exists in the mind of the perpetrator for every criminal offense.

The final comparison of this research project was originally designed to ascertain in what context do discrepancies between the forensic disciplines occur. In no instance were any contradictions between laboratory disciplines identified. This is particularly significant when considering that all of the findings of the PDD and laboratory examinations were achieved autonomously. In every instance, the PDD examination was completed before the results of the laboratory examination had been completed.

Most opponents of PDD procedures have argued that the high accuracy of PDD examinations in some laboratory studies cannot be generalized to the field environment (Office of Technological Assessment [OTA], 1983). This is due largely to the fact that ground truth is difficult to establish in the field environment (Lykken, 1979). The findings of this comparison support other studies that utilized the confession as ground truth (Barland Raskin, 1976; Patrick & Iacono, 1988). In these studies, PDD was found to have been over 90% accurate in the field setting. Iacono (1991) asserts that sample bias creates "substantial methodological shortcomings" (p.201), and that PDD examinations which historically have been selected for studies based on confessions are a select group and reflect bias in favor of PDD. While there can be no question that examinations verified by confession are a unique subset of PDD examinations, this study indicates that this bias has a minimal impact, and confession-based samples would accurately reflect the overall population.

³ USACIL was not among the laboratories which were reviewed by the Laboratory Proficiency Testing Research Program.

This assertion is further substantiated by a study conducted by Mason (1988), wherein 111 PDD examinations were conducted in which ground truth was ascertained by urinalysis examinations. The validity of PDD (verified by these biomedical tests) was in excess of 95% and if utilizing confessions in conjunction with the urinalysis forensic discipline, accuracy of that confession subset would be over 98%.

Another means of reviewing PDD results was utilized by Peters (1982) in which he reviewed 220 PDD examinations. After stipulation for admittance into a state court a PDD examination was conducted. Based upon the outcome of the judicial proceedings, Peters found "the vast majority of settled cases were resolved in a manner consistent" with the PDD

results (93.1%)(p. 164). Further, 98.8% of the defendants had the charges dropped when they were opined to have been truthful during a stipulated examination.

Additional research relating to the interdisciplinary comparison of forensic disciplines is warranted. This methodology not only addresses the reliability between forensic disciplines, but also provides an excellent process with which to address the issue of ground truth in a field situation by means other than confession. This process does involve a particular subset of criminal investigations (i.e. only those investigations involving multiple forensic examinations). However, this subset cannot be considered as having a bias favoring PDD, as is sometimes argued.

References

- Barland, G. H., & Raskin, D. C., *Validity & Reliability of Polygraph Examinations of Criminal Suspects* (Report No. 76-1. Contract No. 75-NI-99-0001) Washington, DC: National Institute of Justice.
- Cole, G. F. (1986). *The American system of criminal Justice* (4th ed.). Monterey CA: Brooks/Cole Publishing Co.
- Donigan, R. L., Fischer, E. C., Hugel, D. H., Reeder, R. H., Williams, R. N. (1980). *The evidence handbook* (4th ed.). Evanston, IL: Traffic Institute, Northwestern University.
- Horvath, F. (1977). The effect of selected variables on interpretation of polygraph records. *Journal of Applied Psychology*, 52 (2), 127-136.
- Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests? *Psychophysiology*, 4, 201-207.
- Keeler, L. (1935). From personal papers and archival material donated to the Department of Defense Polygraph Institute Library, Fort McClellan, AL.
- Lykken, D. T. (1979). The detection of deception. *Psychological Bulletin*, 86, 47-53.
- Mason, P. (1991). Association between positive urinalysis drug tests and exculpatory examinations Unpublished doctoral dissertation. United States Army Criminal Investigation Command, Baltimore, MD.
- Moenssens, A. A., Inbau, F. E., & Starrs, J. E. (1986). *Scientific evidence in criminal cases* (3rd ed.). Mineola, NY: Foundation Press.

- Office of Technological Assessment. (1983). *Scientific validity of polygraph testing: A research review and evaluation - A technical memorandum* (U.S. Congress, OTA-TM-H-15). Washington, DC: U.S. Government Printing Office.
- Parker, B. (1963). The status of forensic science in the administration of criminal justice, *Rev. Jur. U.P.R.*, XXXII, No. 2, 414-417.
- Patrick, C. J. & Iacono, W. G. (1987). Validity and reliability of the control question polygraph test: A scientific investigation. A paper presented at the Society for Psychophysiological Research Meeting, Amsterdam.
- Perry, M. V. (1990). Proposed Military Rule of Evidence 707, polygraph evidence in the military courts: solution by edict is no solution. *Polygraph*, 19(1), 12-61.
- Peters, R. B. (1982). A survey of polygraph evidence on criminal trials, *American Bar Association Journal*, 69, 161-165.
- Peterson, J. L., Fabricant, E. L., Field, K. S. (1979). Laboratory proficiency testing research program (Final Report). Rockville, MD: Forensic Sciences Foundation.
- Raskin, D. C. & Podlesney, J. A. (1979). Truth and deception: A reply to Lykken. *Psychological Bulletin*, 86, 54-59.
- Saferstein, R. (Ed.) (1982). *Forensic science handbook*. Englewood Cliffs, NJ: Prentice-Hall.
- Widacki, J., & Horvath, F. (1978). An experimental investigation of the relative validity and utility of the polygraph technique and three other common methods of criminal identification. *Journal of Forensic Sciences*, 23, 596-601.
- Willard, R. (1982). The value of the inconclusive opinion. Prepared speech presented at the Federal Inter-Agency Seminar, Quantico, VA.
- Yankee, W. (1992). A case for forensic psychophysiology and other changes in terminology. Memorandum prepared at the Department of Defense Polygraph Institute, Fort McClellan, AL.

Appendix A

Discipline Descriptions

The eight primary disciplines utilized in this study are depicted below. This list is meant to describe, for the purposes of this research, the primary types of analyses conducted by the different laboratory divisions within USACIL-CONUS and the PDD conducted by the USACIDC Polygraph Program. The list is a guide for the reader in understanding how the different disciplines were categorized based upon the analyses completed during the reviewed time period.

Questioned Documents

The review of evidence in this discipline included, but was not limited to, the forensic examination of handwriting and hand printing identifications, typewriting identifications and comparisons, erasures, obliterations, alterations, composition of major types of writing inks, paper comparisons and datings, charred documents, evidence of alterations, writing with the unaccustomed hand, analysis of inks, and imprinting or stamping machines.

Illicit Drugs

The review of evidence in this discipline involved the forensic examination of substances submitted to USACIL-CONUS suspected of being controlled substances as defined by the Comprehensive Drug Abuse Prevention and Control Act. Street drugs and prescription drugs of abuse are included in these definitions.

Latent Fingerprints

The review of evidence in this discipline involved the forensic examination of evidence submitted in order to identify if latent fingerprints were found at the scene and subsequently to ascertain if identifiable latent prints could be matched to a suspect, a victim, or other persons.

Serology

The review of evidence in this discipline included, but was not limited to, the forensic analysis of blood evidence (the identification of bloodstains, determination of species origin, techniques for the determination of blood groups, etc.); other biological matter (sperm cells, saliva, perspiration, etc.); and DNA referrals to other laboratories with the requisite capabilities.

Trace Evidence

The review of evidence in this discipline included, but was not limited to, the forensic analysis of hair; fibers; paint; flammables; soil; dust; and the identification and comparison of other minute particles, objects, and liquids.

Firearms

The review of evidence by this discipline included, but was not limited to, the forensic analysis of firearms, bullets, cartridge casings, the operational capability of a weapon, gunshot residue, tool mark identification, and the operational nature of locking devices.

Photographic Division

This discipline involved few (10) actual forensic examinations. However, photographic support is included as one of the seven laboratory division categories for this project as it was the most

appropriate category for these particular examinations. It is noted that the photography division of any criminalistics laboratory provides a number of special processes for the criminalist lab and the selection of the particular task depends upon the type of evidence involved and the result sought. Examples of specific photographic support to other disciplines would be: filters to emphasize certain colors or suppress others; infrared photography to assist in the discovery of erasures on documents; reveal blood stains, etc. The particular examinations completed for this project included, but were not limited to, videotape restoration, enlargements, or specialized film development.

Psychophysiological Detection of Deception

The review of evidence by this discipline included the process of determining if a person is attempting to deceive or is being truthful to an issue in question. This conclusion is arrived at by considering: Stimuli (questions) are presented and are psychologically evaluated by the examinee. The subjective interpretation of the stimuli will affect the activity levels of the selected physiological functions that are recorded (Yankee, 1992). These recordings are quantified and the diagnosis rendered.

Appendix B

Data Collection Procedures

The following is a list of the four possible outcomes of each laboratory discipline determined by opinions rendered in the USACIL-CONUS laboratory reports:

Positive Result

The results of the laboratory analysis were positive if they provided the user with investigative information that might lead to the resolution of the criminal investigation.

Positive Result with Subject

The results of the laboratory analysis were positive and the information from that analysis provided sufficient information amounting to a reasonable belief that the suspect committed the criminal offense that was the source of the laboratory request.

Positive Result with No Subject

The results of the laboratory analysis were positive if they provided the user with an investigative clue that might lead to the resolution of the criminal investigation, but the information from that analysis did not provide sufficient specific information about an individual to amount to a reasonable belief that the person was involved in the commission of the criminal offense that was the source of the laboratory request.

Negative Result

The results of the laboratory analysis provided no positive information that would assist the investigator with an investigative lead or to identify the perpetrator of the criminal offense that was the basis for the laboratory request.

Outcomes by Discipline

The following are examples of criteria for determining the four outcomes for each discipline:

Questioned Documents

Positive Result

1. The forensic examination revealed that an individual was possibly the author of the questioned documents.
2. The document submitted was determined to have been altered or that the victim did not author the questioned writings.

Positive Result With Subject

1. The forensic examination revealed that an individual was the author of all or part of the writings on the questioned document.
2. The forensic examination revealed that the document had been altered, and the victim or custodian had alleged otherwise.

Positive Result With No Subject

1. The forensic examination revealed that the document had been altered but could not identify or eliminate any individuals as having made any of the writings on the questioned document.
2. The forensic examination revealed that the printing on the questioned document had been typed by a specific typewriter but no specific subject was identified.

Negative Result

1. Forensic examination revealed that no one could be identified or eliminated as having been the author of the questioned documents.
2. Forensic examination revealed that the word processor suspected to have been utilized to create the questioned document could not be identified or eliminated.

Illicit Drugs**Positive Result**

1. Forensic examination revealed that the substance submitted was determined to be a controlled substance.
2. Forensic examination revealed that the substance submitted was determined not to be a controlled substance and the substance was identified.

Positive Result With Subject

Forensic examination revealed that the substance submitted was determined to be a controlled substance and the substance was linked to an individual.

Positive Result With No Subject

1. Forensic examination revealed that the substance submitted was determined to be a controlled substance and the substance was not linked to an individual.
2. Forensic examination revealed that the substance submitted was determined not to be a controlled substance.

Negative Result

Forensic examination revealed that no determination could be made about the substance submitted.

Latent Fingerprints**Positive Result**

Forensic examination revealed that the material submitted contained latent fingerprints suitable for identification.

Positive Result With Subject

Forensic examination revealed that the material submitted contained latent fingerprints suitable for identification and that the latent fingerprints matched those of a suspect.

Positive Result With No Subject

Forensic examination revealed that the material submitted contained latent fingerprints suitable for identification but that the latent fingerprints did not match any suspect fingerprints submitted.

Negative Result

Forensic examination revealed that the material submitted did not contain any latent fingerprints suitable for identification.

Serology

Positive Result

1. Forensic examination revealed that the substance submitted was determined to be human blood.
2. Forensic examination revealed that the substance submitted was determined to contain sperm cells.

Positive Result With Subject

1. Forensic examination revealed that the substance submitted was determined to be a blood type which is the same type as the suspect. This type of blood is found in a certain percentage of people.
2. Forensic examination revealed that the substance submitted was determined to contain human saliva. The blood type of the saliva is the same as the suspect. This type of blood is found in a certain percentage of people.

Positive Result With No Subject

1. Forensic examination revealed that the substance submitted was determined to be blood, but not enough of a sample was submitted to further identify the sample.
2. Forensic examination revealed that the substance submitted was determined to contain human saliva. No further identification of the sample could be completed.

Negative Result

Forensic examination revealed that no determination could be made about the substance submitted.

Trace Evidence

Positive Result

1. Forensic examination revealed that the material submitted contained a substance identified as an accelerant.
2. Forensic examination revealed that the material submitted contained paint fragments which could be identified as being similar to the type related to the suspect.

Positive Result With Subject

1. Forensic examination revealed that the material submitted from the suspect's clothing contained a substance identified as an accelerant.
2. Forensic examination revealed that the material submitted contained paint and glass fragments which could be identified as having originated from the suspect vehicle.

Positive Result With No Subject

1. Forensic examination revealed that the material submitted contained a substance identified as an accelerant.
2. Forensic examination revealed that the material submitted contained paint fragments which could not be identified or eliminated as matching the suspect vehicle.

Negative Result

Forensic examination revealed that the material submitted could not be identified.

Firearms

Positive Result

1. Forensic examination revealed that the projectile submitted was fired by the weapon submitted.
2. Forensic examination revealed that the locking device submitted was in operational order and no signs of tampering were noted.

Positive Result With Subject

1. Forensic examination revealed that the projectile submitted was fired by the suspect's weapon.
2. Forensic examination revealed that the locking device submitted had been cut by the bolt cutters found in the possession of the suspect.

Positive Result With No Subject

1. Forensic examination revealed that the weapon submitted was operational, but the projectile could not be identified or eliminated as having been fired by the weapon submitted.
2. Forensic examination revealed that the lock submitted was cut by a device such as bolt cutters; however, the bolt cutters submitted could not be identified or eliminated as having cut the submitted locking device.

Negative Result

1. Forensic examination revealed that the projectile submitted could not be identified or eliminated as having been fired by the weapon submitted.
2. Forensic examination revealed that the bolt cutters submitted could not be identified or eliminated as having cut the submitted locking device.

Photograph Division

Positive Result

1. Forensic examination revealed that the negatives submitted were restored and prints were successfully developed.
2. Forensic examination revealed that still photographs were developed and enlarged from the videotape submitted.

Positive Result With Subject

Forensic examination revealed that the negatives were developed and photographs with negatives of the suspect's likeness were successfully developed.

Positive Result With No Subject

Forensic examination revealed that the negatives submitted were restored and developed, but no images or likenesses of any persons were observed on the negatives or prints.

Negative Result

1. Forensic examination revealed that the negatives submitted could not be restored or developed.
2. Forensic examination revealed that all attempts to develop still photographs from the submitted videotape were unsuccessful.

Psychophysiological Detection of Deception

Positive Result

The examiner, with subsequent quality control concurrence, rendered an opinion of deception or no deception indicated as a result of the completed PDD examination. These findings provide the user with investigative information that might lead to the resolution of the criminal investigation.

The following definitions are those adhered to within the Department of Defense, in accordance with Department of Defense Directive 5210.48R (Draft) December 1990.

Deception Indicated DI

A PDD examination result based on analysis of PDD charts indicating that an examinee's physiological responses indicated deception when answering relevant questions concerning the matter under investigation.

No Deception Indicated (NDI)

A PDD examination result based on analysis of PDD charts indicating that an examinee's physiological responses did not indicate deception when answering relevant questions concerning the matter under investigation.

Inconclusive

A PDD examination result based on analysis of PDD charts indicating that a conclusive determination (NDI or DI) could not be made.

No opinion

A term used to describe the overall results of a PDD examination wherein circumstances prevent the examiner from obtaining sufficient data to form an opinion.

Positive Result With Subject

The results of the PDD examinations in this category were separated into DI and NDI.

Deceptive (DI)

An opinion of DI accompanied by information received from the individual undergoing the PDD examination that amounted to a statement against his or her self-interest (admission/confession) or the information developed during the PDD examination provided the investigation with a suspect based on this information.

Non-Deceptive (NDI)

An NDI opinion, accompanied by information received from the individual undergoing the PDD examination that another person was identified as the perpetrator of the offense. In order for an examination to be identified within this category, the other person identified in the Report of Investigation (ROI) would have to have been listed in the "Title Section" of the ROI. In order for the other person to have been listed in the title portion of the investigation, a prosecutor would have to opine that enough criminal information and probable cause exists to believe that the other person committed the offense for which the PDD examination was requested.

Positive Result With No Subject

The examiner, with subsequent quality control concurrence, rendered an opinion of DI or NDI, but no individual could be identified as a result of the PDD examination.

Negative Result

The examiner, with subsequent quality control concurrence, rendered an opinion of inconclusive or no opinion as a result of the completed PDD examination.

Appendix C

Resolution Criteria

During the course of an investigation, the investigator relies on differing types and amounts of evidence and/or information to indicate a suspect's involvement or lack thereof, in the criminal offense. The credence or weight attached to the various forensic disciplines by the field investigative element is not known. A method of reviewing the value or weight placed in the forensic sciences by the field element and to further consider the utility of the forensic sciences in the real world setting would be to ascertain the degree to which the forensic disciplines assist in the resolution of the investigation. Of the 914 analyses reviewed during this project, the 920 ROIs that contained these analyses were located for review at USACRC, Baltimore, MD. A portion of this review consisted of determining the number of ROIs that resulted in an individual being officially listed as the person who was criminally involved in the offense. An individual who is identified as having been criminally involved in the offense is listed within the "Title Section" of a USACIDC ROI. In order for a person to be listed in the title section of a USACIDC ROI, an established standard of proof must be met. This standard is: probable cause must exist that a crime was committed and that this individual committed the offense. The process utilized by USACIDC for making this determination is not arbitrary, but one that is relatively consistent throughout USACIDC.

In order to place a person in the title section, initially, the USACIDC special agent who has conducted the investigation or is responsible for the investigation, will attain a level of proof through the investigative process that would cause the special agent to opine that probable cause exists to believe the suspect committed the offense. Once the investigator believes this level of proof has been reached, the special agent refers the ROI to a supervisory special agent who will review the investigation. If the supervisor concurs that a sufficient level of proof exists, the investigative special agent briefs the prosecuting attorney on the investigation to ascertain if the prosecutor believes the probable cause standard has been met. The prosecutor, when making a determination that this standard has been met, understands that this same prosecutor will be expected to assure that subsequent judicial action is taken against that individual. It should be understood that once a person is titled in a USACIDC ROI, upon completion of the investigative and administrative processing of the ROI, the completed ROI is forwarded to the commander or the agency in charge of the person who is titled in the ROI. This commander or agency head is then required (by Army regulation) to take action against the individual or provide written justification as to why no action was taken. The attorney who concurs with the listing of the person in the title section of the ROI is usually the attorney who will be responsible to the commander or the agency head for taking subsequent judicial action against that individual. Therefore, prosecuting attorneys are usually careful when concurring with the USACIDC special agent in placing an individual in the title section of a USACIDC ROI.

It is interesting to note that PDD has not always been considered a traditional forensic discipline. This is ironic since the PDD discipline was an integral part of the first National Crime Laboratory at Northwestern University School of Law (Keeler, 1935). In 1935, a proposed expansion of this laboratory called for "... two psychophysiological laboratories for lie detection.," (Keeler, 1935). This established the National Crime Laboratory as the prototype for many of the current forensic disciplines and laboratories.

As noted with the USACIDC Polygraph Program, this PDD program, as well as many PDD programs, are not physically located within most crime laboratories because of the portability of the PDD instrument and the need for mobility of PDD examiners. Most large investigative organizations assign PDD examiners to different geographical locations to ensure this forensic tool is available to the field investigator. The portability of the instrument, combined with the broad application of PDD examinations to virtually all types and categories of crimes, accounts for the fact that PDD was utilized in a far greater number of investigations than other forensic disciplines.

Practicing forensic psychophysiologicalists have argued that the accuracy of PDD tests will be greater in real life situations than in laboratory studies (OTA, 1983). This phenomenon may occur since arousal associated with the commission of a mock crime in a laboratory study is not likely to produce physiological responses in the guilty subjects as great as the magnitude of the physiological responses of subjects guilty of an actual crime.

Based on the data utilized in this study, the reliability coefficient between the PDD discipline and the other laboratory disciplines is 1.00. This fact certainly establishes the reliability, if not the validity, of the PDD examination in the field. If one were to question the validity of the PDD results, one would also have to question the validity of the corresponding results of the other forensic disciplines. The validity of PDD was further reinforced when utilizing those 22 ROIs in which a confession was obtained. All ROIs in which a confession was elicited confirmed the PDD results. It is also interesting to note, given the data included in this report, that the results of the other forensic disciplines are routinely admitted in court, while the results of PDD examinations are normally excluded from admission as evidence in criminal trials (Perry, 1990).

The U.S. Army CID Command is one of the finest investigative entities in existence in the United States. USACIDC's high solve-rate alone of 81% establishes this fact. Additionally, USACIDC laboratory and PDD examiners receive the most demanding and comprehensive training and education available in their specialties. The quality control standards for all disciplines are extensive and adhered to by both USACIL and the USACIDC Polygraph Program. There are crime laboratories and polygraph activities which fail to maintain such standards. In order for other entities to attain the high rate of correlation and lack of contradiction in forensic results of this research project, those entities must require similar standards of their crime laboratories and PDD programs.

Memory, Reason, Emotion, Anatomy and the Polygraph The Mind-Body Link Offered by Neurobiology and Related Sciences

William B. Anderson

Abstract

The Supreme Court, in Daubert 1993, opened the door to expert witness testimony derived from polygraph. No such witness has passed through-yet. We postulate a lack of scientific foundation as the reason, and we describe current findings from brain scientists as a way to correct that deficiency. They now say with increasing frequency and intensity: We can explain what goes on in the brain regarding memory, reason, emotion and anatomy linkages. These elements evolved together. They are dependent on and modify each other. Emotional memories such as fear are permanently ingrained into the brain, and can be suppressed but never erased. Body is the frame of reference for the mind-brain. Emotional brain circuits are just as tangible as circuits for seeing, hearing, and touching. The biological nature of emotions can now be described. An increasing volume of published material from these scientists generally validates these statements. None refer to or indicate awareness of a polygraph relationship. Nor are polygraph practitioners aware of this significant potential scientific foundation toward admissibility. One key scientist apprised in detail of specific incident polygraph process agrees-there is a theoretical basis for this linkage. An historic review leading to and summarizing brain research is presented along with a proposed method for introducing these two unaware disciplines, brain science and polygraphy, to each other.

The Supreme Court, in Daubert 1993, disposed of per se exclusion of polygraph evidence as expressed in Frye 1923. Federal trial courts must now consider admission, though it has been so far excluded. While there may be hope, there is little tangible basis to expect acceptance. Ample and substantial resistance comes from many directions: judges and prosecutors, the psychology/psychiatric cohort, medicine as expressed in AMA policy, civil libertarians, labor and some political quarters, for starters. Whatever the arguments for admission, they have not been persuasive, arguably because they have contained no acceptable scientific explanation. Such scientific basis for polygraph has its foundation in psychology, a largely behavioral discipline rather than hard science.

Most psychologists derogate polygraphy. (Devitt, Honts and Vondergeest, 1997) A psychiatrist may give expert opinion regarding truthfulness of a witness' court

testimony without having tested or examined that witness. A polygrapher who has examined and tested that witness may not. This conflict, unresolved, is a strong stimulant for the research here proposed.

Is there now a settled scientific basis for polygraphy? It certainly is not visible in the material available to practitioners and interested counsel, or we would have seen it in court reports or *Polygraph*. Is there, anywhere, a potential or possible hard science foundation for polygraph? Perhaps, just perhaps, which underlies this inquiry, review and synthesis.

On December 6, 1994, the *New York Times* Science Section carried an article, "Tracing the Brain's Pathways for Linking Emotion and Reason", which emphasized that an emotion like fear has been conserved through evolution. Polygraphers know as much about the emotion of fear as any group

The author is an APA member and previous contributor to this journal. Requests for reprints should be sent via FAX to (610) 469-8639.

on earth, probably more. It is the emotion with which we work. This article reported as well that emotional memories involving fear permanently ingrained into the brain can be suppressed but never erased. We examiners surely knew that but had never heard it from the scientific community. There was more: That the body is the frame of reference for with which we work. This article reported as what humans experience as mind and our thoughts and actions use the body as a yardstick. (Now that was close to saying cardiovascular and respiratory and electrodermal tracings were the body telling us what is going on in the mind.) Even more, that emotional brain circuits are just as tangible as circuits for seeing, hearing and touching. Further, that the biological nature of emotions and feelings can at last be described.

It is important to know the *New York Times* article never mentioned or referred to polygraph in any way, attributing these fascinating comments to three scientists: Dr. John Allman at Cal Tech, Dr. Joseph LeDoux at New York University and Dr. Antonio Damasio at the University of Iowa Medical School (M.D. Neurologist).

My opinion was then and is now that here emerges a scientific foundation for specific incident comparison question polygraphy. What made it even more valuable was its total independence from polygraphy, as we know it.

Dr. Allman and I later met. We reviewed "The New York Times" article and his research. I then detailed to him the regime in a proper, specific incident polygraph test and asked if he thought my belief these (meaning all three) researches could be a scientific foundation for this process. Answer: "Yes. I think it theoretically entirely reasonable."

In 1998, more time to reflect returned.

I had done little to forward my conviction except to follow published, relevant reports. There were a great many, and they are here briefly summarized, but first, some early history.

In the 17th Century Rene Descartes (1596-1650), French scientist and mathematician, wrote and said forcefully and repeatedly that Body and Soul had no linkage,

were entirely separate and unconnected. By Body he meant human physical anatomy, excluding the brain. By Soul, he meant the higher brain, the mind, emotion, morality, reasoning and thought, and the spiritual. Had Descartes been correct in his assertion there was no linkage between the anatomy and emotion/reasoning, the concept of polygraphy by definition could have had no basis to exist. He was wrong. While Descartes' conviction now seems medieval and irrelevant, it does provide a useful reference point against which to compare modern science.

Charles Darwin, in 1872, wrote *The Expressions of Emotions in Man and Animals*. Little can be said that could be new in regard to his genius. *The Expressions* has gone through three editions, the most recent in 1998 by Oxford University Press. The index does not include the word polygraph, of course, and very little about lies or lying, except a footnote by Dr. Paul Ekman, stating the pitch of the human voice is believed to become higher when people lie.

Of general interest to us is the discussion by Darwin of an extremely wide range of emotional expressions such as suffering, weeping, anxiety, grief, dejection, despair, joy, love, reflection, hatred, anger, disgust, guilt, deceit, surprise, fear and shame, most of which have little meaning in this review. But guilt and deceit do have meaning, and Darwin reports observations which are quite expectable, about the avoidance of looking at the accuser, eyes wavering askance or moving restlessly. In one instance, he reports his detection and denial of an offense by a 2-year 7-month old child (his son) based on unnatural brightness of eye and an odd, affected manner. In an amusing footnote he reports the natives of India are able to, when giving evidence, control the expression of their faces but not the toes, which often reveals the witness is lying.

The meaning? Here is the Great Man telling us from 126 years ago the expression of guilt and deceit emotions could not be suppressed, that Descartes was wrong, and perhaps we are on a reliable track.

The first demonstrable physical linkage between mind and body had its start in 1848

with Phineas Gage, a construction foreman. He was 25 and in charge of a track gang putting down new track for Rutland and Burlington Railroad near Cavendish, Vermont. Gage was in charge of the entire crew's activity, but his particular task was tamping explosive powder into the drillings of interfering rock. On that day, distracted, Gage began tamping too soon. The necessary sand cover over the explosive was not in place and the tamping iron struck sparks into the powder. It exploded, driving the rod through, entirely through Gage's skull. The rod was three feet seven inches long, one and one quarter inches in diameter, tapering to one quarter inch on the end. This end entered his skull first. Gage fell to the ground, terribly injured, but quickly regained consciousness. Within a short time he began to speak, was carried to a cart, thence to a nearby hotel where he was treated by a local, but quite skillful physician. He survived until 1861, 13 years, but as a very different man.

Before the accident, Gage was athletic, vigorous and graceful, efficient and capable, had considerable energy of character, temperate in habit, of a well balanced mind, a shrewd, sharp businessman, and persistent in his plans of action. He was a successful man, especially for age 25.

After the accident, a new and unpleasant personality emerged: fitful, irreverent, profane, having little deference for his fellows, (this old gang had great affection for him before the accident), obstinate, capricious and vacillating--a child in his intellectual capacity.

In the 1840's-60's little significance was given to these personality changes of Gage, though he was ably treated for the injury by physicians of more than average competence. It is only in recent years as specialized brain functions are identified that the specific injury to Gage's brain and its personality consequences make sense to the initiated. They say in effect, "That sort of injury to that segment to his brain ought, from what we know, to have caused those changes reported in his personality." A hundred or more years later, Gage confirms for neurobiologists what current sophisticated

research describes with specificity-the Damasio's particularly.

Dr. Antonio Damasio, author of *Descartes' Error-Emotion, Reason and the Human Brain*, in response to Descartes' separateness of mind and body, says nature appears to have built the apparatus of rationality (mind) not just on top of the apparatus of biological regulation but also from it and with it. There is not separation between mind and body. The body's needs set the pace and indirectly drive the brain's decisions. In the simplest terms, perhaps too simple, without the body there would be no mind.

Mrs. Hanna Damasio, Antonio's wife, using modern neuroanatomy and neuro-imaging technology, examined the skull of Phineas Gage and the tamping iron which did the damage in 1848, and was able to complete an autopsy. Plotting the trajectory of the iron through Gage's skull she could, 120 years later, provide a bridge between Gage and modern research on frontal lobe function. It all matched quite nicely.

Dr. Damasio then describes a patient referred to him, "Elliot," who, after removal of a large non-malignant brain tumor and other tissue, in precisely the area of damage as Gage, exhibited similar inabilities in thought processes as Gage. When a specific physical brain part was removed, the reasoning associated with that part disappeared as well.

Dr. Joseph LeDoux, a professor of neurobiology at New York University, author of *The Emotional Brain*, Weidenfield and Nicolson 1998, is a pioneer in research of neural circuits underlying emotion. Much of his research centers on the basic emotion of fear and the circuitry which supports its expression. Not surprisingly much of this research involved rats, loud noises and a mild electric shock to their feet. According to LeDoux, fear reactions were highly conserved through evolution, a reasoning he links to human fears. Much of LeDoux's research also deals with the amygdala, a tiny structure deep in the brain, crucial for the formation of memories about significant emotional experiences. He found that to damage a rat's

amygdala was to cause its forgetting to be afraid. More, much more interesting, is his finding that over time his rat subjects did lose some of their fear of the sound preceding the shock, but the fear was never eliminated. Moreover, alteration of a small region of the rats' forebrain foreclosed this loss of fear and they then remained afraid much longer. This is construed by LeDoux to be the locus of the reasoning area which overrides fear.

Published material in body/mind research is inevitably random and sporadic. There is much more going on in this field but no one central repository or collector exists. To wait for organization in an independent research atmosphere defeats the urgent needs of the polygraph practitioner and is, as well, unlikely to occur. This synthesis in Step 1, a basic and simplified description of an entirely new direction for practitioners, of an original and expanding explanation of what goes on in the "black box" of the brain when called upon by emotion and/or the need for rational thought. For us, that call is imbedded in the relevant (issue) questions of a specific incident polygraph test. We have been doing laboratory research within the brain-mind of clients, and can show very high accuracy and validity, but the requisite explanation of "how and why" has not been revealed. This paper is the initial presentation to practitioners.

Step 2 must be the full inquiry into the above perceptions with all the identifiable active scientists, comparing the preliminary agreement of Dr. Allman with their experience and thoughts and this polygraph synthesis. Step 2 is easy to state but contains some substantial roadblocks. A cutting edge brain scientist can be, and maybe should be, driven by a dream of personal glory-a Nobel Prize, perhaps. He will not share his hard-found knowledge readily, especially with us. The scientists here referred to will keep their distance, preferring their quiet towers to the give and take of the courtroom and association with deception. No matter, this link is more than worth the effort.

The researchers included in this synthesis are pure scientists devoted to finding and revealing original basic knowledge. They are not subsidized, narrowly focused employees of a large corporation whose ultimate goal is to profit from application of their discoveries. Polygraph practitioners, would draw on this knowledge. This will inevitably include strictures from pure science as well as the courts, and inevitably bring changes yet unknown to the polygraph process. We should welcome changes which make more certain the accuracy, validity and admissibility of a rigorous polygraph process.

References

- Devitt, M.K., Honts, C.R., & Vondergeest, L. (1997). Truth or Just Bias: The treatment of the psychophysiological detection of deception in introductory psychology textbooks. *Journal of Credibility Assessment and Witness Psychology*, 1(1), 9-32. Reprinted in *Polygraph*, 27(2), 123-148.

History and Development of Polygraph Examinations in Bulgaria

Svetoslav Zanev¹

Polygraph examinations were established in Bulgaria in the late sixties. They were born as a result of Cold War competition in the field of intelligence and in an attempt to develop new tools for investigation. Use of the polygraph in the West, beyond the Iron Curtain, stimulated research on polygraph in Bulgarian law-enforcement research centers. In the beginning, research was performed with different medical equipment for studying the influence of the biofeedback and drug substances on physiological reactions. Later, at the Laboratory of Psychology at the Ministry of Interior two Stoelting instruments were received for training and polygraph examinations. Primarily polygraph examiners were medics with experience in psychiatry and similar medical instrumentation. The pioneer of polygraph examinations in Bulgaria is George Boianov, M.D.

Training was carried out by using Reid & Inbau "Truth and Deception" (ed. 1966), translated manuals, articles and shared experience from people who had been tested abroad.

The test formats included the RI question format, POT and stim tests with cards and numbers. RI format was modified by including semi-strong relevant questions, which played the role of comparative (control) questions.

The main paradigm used was the clinical approach (Matte, 1996), where evaluation of behavior was as important as the physiological reactions. Interpretation of the polygraph examination was complemented with the exploration of different nonverbal

cues marking deception. The scoring system was intuitive, with accent on reaction or non-reaction.

The polygraph was also used in self-control and biofeedback training.

In the criminal investigations polygraph examinations were used mainly in cases of conspiracy and espionage.

Usually the polygraph as a special tool was used together with the voice stress analyzer (VSA). The VSA was a preferred tool due to its ease of use and interpretation of the results.

Before 1989 the use of the polygraph was secret, and information regarding examinations was classified and restricted to a small group of people. After 1989, in the time of transition to democracy in the Bulgarian society, the polygraph unit was discharged. Polygraph examinations almost stopped. Polygraphs were used periodically in criminal investigations (murder cases, stealing of weapons, hijacking).

In 1997 the Ministry of Interior purchased a new polygraph for continuing examinations. After having made contacts with Lafayette Instruments, the new computerized instrument was received in the Institute of Psychology (former Laboratory). This started a new turn in polygraph examinations. Mr. Paul Redden from the San Diego Police Department, who brought the LX3000 polygraph system, carried out a short course where the Backster Numeric Score System was introduced.

¹The author is the President of the Bulgarian Polyphysiograph Association and an Intern Member of the American Polygraph Association. Requests for reprints can be sent to Dr. Zanev at: Bulgaria, Sofia 1000, Bulgarian Polyphysiograph Association, P.O. 146, or by e-mail: sis@mobikom.com

The new polygraph became an important stimulus for the use of polygraph examinations by the Bulgarian Ministry of Interior. The Polygraph unit (3 officers) was included in the Department of Expertise and Consulting at the Institute of Psychology. Since receiving the polygraph, more than 200 specific examinations were performed. The examinations were done in murder cases, fraud, corruption and screening procedures for police officers. By the means of the polygraph some difficult murder cases were solved. A new direction of the polygraph was in the fight against corruption.

On April 9, 1998 the Bulgarian Polyphysiograph Association. was established. The Association is a professional institution, and its purpose is to develop and refine polygraph methodology for detection of deception. It is trying to introduce the

polygraph examination into the legal system as a supplementary means for verification of statements, and for evidence investigation. The members of the Association are private and government polygraph examiners.

The Association works closely with research centers connected with polygraph methodology and instrumentation – The Institute of Physiology, Biomedical Engineering Institute, and Institute of Psychology.

Now polygraph examinations are performed both in the private and government sectors. Three computerized Lafayette instruments are used in Bulgaria now. Results of polygraph examinations per se are not admissible at the court. Polygraph examinations are presented as a part of psychological expertise for verification of judgments.

References

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