Polygraph

VOLUME 29 2000 NUMBER 3

Contents

Department of Defense Polygraph Program Annual Report to Congress - Fiscal Year 1999 209

A Field Validity Study of the Integrated Zone Comparison Technique 220
Nathan J. Gordon, William L. Fleisher, Hisham Morsie, Walid Habib & Khaled Salah

The Debate Over Polygraph in Poland 226
Aleksander Krzyścin

Computer Algorithm Comparison 237
Andrew B. Dollins, Donald J. Krapohl & Donnie W. Dutton

Identifi Comments on Dollins et al's Computer Algorithm Comparison 248
W. Keith Hedges

CPS Comments on Dollins et al's Computer Algorithm Comparison 250
John C. Kircher & David C. Raskin

JHU/APL Comments on Dollins et al's Computer Algorithm Comparison 253
John C. Harris

A Response to Comments 256
Donald J. Krapohl & Donnie W. Dutton

Polygraph in the Crimea: First Results and Future Perspectives 258
Gennady G. Moskal & Vitaliy I. Egorov

A Critical Analysis of Amsel's Comparative Study of the Exclusive v. Nonexclusive Comparison Question 261
James Allan Matte & Cleve Backster

Addendum to 1990 Field Study of the Friendly Polygrapher Hypothesis 267
James Allan Matte & Thomas E. Armitage

Bibliography - The Published Works of Norman Ansley 270
Genni Arledge

Published Quarterly

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

Department of Defense Polygraph Program Annual Report to Congress - Fiscal Year 1999

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 CSP examinations.

The following table reflects Department of Defense Polygraph Program statistics for Fiscal Year 1999.

Criminal	2,183	18.9%
Exculpatory	510	4.4%
CI Scope (CSP)	8,289	71.8%
All Others*	564	4.9%
Total**	11,546	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). A breakout of polygraph examinations conducted by NSA is contained in a classified table submitted with this report. 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 1999 Counterintelligence-Scope (CSP) 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.

The Department published a handbook for federal polygraph examiners standardizing

techniques and procedures for conducting polygraph examinations. The handbook also outlines the Quality Assurance Program (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 4.4 of 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 Years 1988 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 (CIA), (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	2,070
(2)	DIA Critical Intelligence Positions	1,174
(3)	TOP SECRET	0
(4)	Examinations for Interim Access to Sensitive Compartmented Information	0
Total Examinations Conducted Under the Congressional Ceiling		3,244
Exempted Examinations*		5,045
DoD Counterintelligence-Scope Polygraph Program TOTAL**		8,289

*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. A table of polygraph examinations conducted by NSA is contained in a classified annex to this report. Nor does it include examinations conducted by the National Reconnaissance Office (NRO), which are conducted under the authority of the Director Central Intelligence.

CSP Refusals

In Fiscal Year 1999, only one person 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 position 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 8,289 individuals tested under the Department of Defense Counterintelligence-Scope Polygraph Program are as follows:

One hundred and ninety six 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 66 examinations required more than one day to complete.

There were 8,088 individuals whose polygraph examination results were evaluated as no significant physiological responses (non-deceptive). The remaining 201 individuals yielded significant physiological responses or were evaluated as inconclusive and/or provided substantive information.

One hundred and eighty-nine individuals made admissions relevant to the issues being tested, and through further testing, the examination was able to resolve all relevant issues favorably to the subject.

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

There was one individual whose polygraph examination result was evaluated as significant physiological response (deceptive) and who made no admissions to the relevant issues.

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

Of the 201 individuals whose examination results were evaluated as yielding significant physiological responses, or evaluated as inconclusive and/or provided substantive information, 187 received a favorable adjudication, five are still pending adjudication, eight 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 and unauthorized foreign contacts. In some of those cases, classified material was commingled with personal papers, and often when discovered, was either destroyed or returned to government control. In some cases, the classified material was deliberately taken home to prepare a briefing or to meet a deadline. Admissions of removal of classified material 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, or if the extent and nature of the foreign contacts were different than reported. Following subsequent polygraph testing the results were provided to appropriate security officials for adjudication.

During a CSP examination, the examinee, who had traveled to a sensitive location, disclosed that he had befriended a foreign national from his target country. The

association involved an exchange of money to the individual and the individual's family for artwork, and additional money to support the foreign national's lifestyle. The examinee also expressed a desire to defect to Cuba. He also admitted that he had manipulated evaluation reports on the target country due to personal sympathies he had for the country.

During CSP testing, the examinee, who previously had been evaluated as deceptive on four CSP examinations given by another U.S. Government agency, admitted to providing prohibited collection export technology to a foreign country. The examinee indicated that he provided the information in order to obtain the business before obtaining an export license.

During a CSP examination, the examinee exhibited "Significant Response" 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. 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 Response" to the above topics, and the matter was referred for additional investigation.

During CSP testing, examinee admitted disclosing classified information to unauthorized persons. An investigation was conducted concerning the unauthorized disclosures and additional polygraph testing was requested. During the subsequent polygraph testing the examinee confessed to providing additional classified information not previously reported and to telling additional unauthorized individuals classified information at a previous duty station. The individual was processed for discharge from the Air Force.

During CSP testing, examinee disclosed during the post-test polygraph interview, that five years earlier he had thoughts of toppling the U.S. Government and was advocating anarchy. He reportedly believed if he could create a race war and destroy transportation and electrical systems, he would obtain this goal. He was denied a security clearance.

During CSP testing, examinee admitted disclosing to uncleared friends and co-workers specific classified details about her employment at another federal agency. These details included target and intercept methodology. The individual was denied access to sensitive information.

III. Utility of the Investigative Polygraph

During the Fiscal Year 1999, DoD investigations obtained unique and significant information from interviews conducted with the aid of the polygraph. In all illustrated instances, the polygraph examination process produced significant security or criminal information, which would not otherwise have been secured for the specific investigation. The polygraph examination process was also valuable in helping to establish the innocence of persons charged with serious infractions.

A transport truck was stopped and searched under the General Framework Agreement for Peace (GFAP) in Bosnia-Herzegovina. During the search, illegal weapons of war in contravention to GFAP were discovered and confiscated. A passenger in the vehicle was interviewed and denied knowledge of the weapons or illegal activities. The individual agreed to undergo a polygraph examination to confirm his story. The examinee was evaluated as deceptive and admitted withholding significant information regarding the weapons being transported. The examinee disclosed the location of a large weapons cache located in Bijeljina, Republia Srpska; identified a Republia Srpska Army (VRS) General and other personnel involved in the illegal weapons transportation; and confirmed that he was a soldier in the VRS. The examinee's information led NATO forces to a warehouse where a major cache of weapons was stored.

An Army member was suspected of having unreported contact with Russian military personnel during a deployment to Bosnia. During the investigation, the member

understated her involvement with Russian military personnel. She agreed to undergo a polygraph examination during which she admitted to an intimate relationship with a Russian Warrant Officer, and that she had continued associations with Russian military personnel during her subsequent assignment to an NSA facility. She denied having contact with any foreign intelligence personnel or having provided details of her military assignment or classified information to any polygraph testing with no deception indicated. This information was referred to the Army Central Clearance Facility for adjudication of her security clearance.

During a background investigation, an employment reference from an insurance company where the subject was previously employed, stated that subject had embezzled money from the cash payments made to him by policyholders. Subject denied the allegation and agreed to undergo a polygraph examination. During the pretest, he admitted to stealing the money and agreed to make restitution to his former employer.

During a background investigation, information was developed that subject was trafficking in drugs on a military base. Subject denied any involvement with illegal drugs since 1958. He agreed to undergo a polygraph examination to support his claim. During the pretest interview, subject admitted that he has been using marijuana, cocaine, LSD, opium and several other drugs since 1980. He also admitted to selling illegal drugs and stated that he intends to continue using and selling illegal drugs. This information was referred to law enforcement authorities.

An investigation was initiated based on a complaint by an Army trainee that she had been raped by her drill instructor. The drill instructor was interviewed and denied any sexual contact with the trainee and agreed to undergo a polygraph examination. The results of the polygraph examination indicated no deception. Based upon the results of the polygraph examination, the alleged victim was re-interviewed and admitted she had fabricated the entire story. The trainee was administratively discharged from the military.

An investigation was initiated regarding the death of an Air Force Noncommissioned Officer. A suspect was developed based on his association with the NCO's wife and access to the residence. The suspect was interviewed and denied any knowledge or involvement in the death of the NCO. The suspect agreed to undergo a polygraph examination to confirm his denials. The examination results indicated deception. During a subsequent interview, the suspect confessed that he planned the NCO's death with the NCO's spouse and committed the murder. The suspect is awaiting trial.

An investigation was initiated regarding the arson of four military recruiting offices in a shopping mall. The primary suspect was a night janitor/security person. This individual agreed to undergo a polygraph examination. The results of the polygraph examination indicated no deception. Next, the military recruiters were polygraphed. The polygraph examination results of the Army recruiter indicated deception. During the posttest interview, the Army recruiter confessed to setting the fire in his office to avoid a command inspection scheduled for the following day. He also confessed to setting the other fires to make it look like kids had done it, thus drawing suspicion away from himself. The recruiter is awaiting trial.

An investigation was initiated on a military officer when a bag of suspected methamphetamine was found in his gym bag during a routine gate check of his vehicle. The officer was interviewed and admitted the gym bag was his, but that he had no knowledge or involvement with the methamphetamine. The officer agreed to undergo a polygraph examination the results of which were evaluated as non-deceptive. Subsequently, a polygraph examination was conducted on the officer's son. The results of this polygraph examination indicated deception. Subsequently, the son admitted that a friend of his had placed the methamphetamine in the gym bag.

An investigation was initiated regarding the reported theft of $877,367.82 from an MCI center located on a military installation. There were no signs of forced entry to the building; therefore, it was suspected that an employee might be involved. Shortly after the theft, an employee quit work at the facility. The

employee was interviewed and denied any criminal knowledge or involvement in the theft and agreed to undergo a polygraph examination. The results of the polygraph examination indicated deception. Subsequently, the employee admitted that she provided the key for the facility to an individual who made a copy of the key and paid her $500 to keep quiet about the theft. The employee received 12 months probation.

An Air Force member was suspected of physically abusing his four-year-old daughter. The abuse resulted in fractures to both of her legs, her right wrist, and several ribs. The member denied any knowledge or involvement in the abuse and agreed to undergo a polygraph examination to support his denial. The results of the polygraph examination indicated deception. During the posttest interview, the member admitted to physically abusing his daughter. Prosecution is pending.

During a background investigation for a DoD contractor, information was developed indicating subject was experiencing financial difficulties due to his excessive gambling. During his interview, subject claimed he attended Gamblers Anonymous in 1992 and 1993, and had not been involved in any illegal gambling activities since 1990. Subject agreed to undergo a polygraph examination to prove his claims. During the pretest interview, he admitted betting money in football pools conducted on U.S. government property, and having friends place between $800 and $1,000 bets for him on jai lai games. The results of the subsequent polygraph examination indicated deception. In a posttest interview, he admitted gambling about $500 a month on Keno games and placing four illegal football game bets through bookmakers. The last of those was made two months prior to the polygraph examination. Adjudication action on his security clearance is pending.

An investigation was initiated after a 5-month-old child was brought into the hospital emergency room with head injuries. It was determined that the injuries consisted of a skull fracture and hemorrhaging of the brain. At the time of the injuries, the child was in the care of his father. The father stated that he and the child had fallen asleep on the couch and the child subsequently fell to the floor. Due to the extent of the injuries, the child was placed on life support which was ultimately discontinued resulting in the child's death. Medical personnel stated that the circumstances described by the father were not sufficient to cause the severe trauma to the child. The father maintained his original statement and agreed to undergo a polygraph examination. The results of the polygraph examination indicated deception. During a subsequent interview, the father admitted that he had shaken the child and hit the child's head on the coffee table because the child would not stop crying. The father is pending court-martial.

A fraud investigation disclosed that a DoD subcontractor provided monetary payment totaling $554,560.00 and two non-interest bearing loans totaling $240,000.00 to a DoD contractor responsible for overseeing engineering contracts at an Air Force Base. The subcontractor agreed to undergo a polygraph examination. During the polygraph examination he admitted to providing gratuities to the government contracting officer in exchange for authorized cost overruns and extensions for completing the contract. All individuals were convicted in court and fined and/or incarcerated.

During a background investigation for a military reservist, information was developed alleging that he molested his former stepdaughters. His former spouse had not reported this information to the police out of fear that it would interfere with ongoing divorce proceedings. During his interview, subject denied ever molesting any children and agreed to undergo a polygraph examination to support his denials. The polygraph examination results indicated deception. During the posttest interview, subject admitted molesting one of his step-daughters. This information was referred to law enforcement authorities. The reservist is awaiting trial.

An investigation was initiated regarding the theft of seven computers valued at $15,125.00. During the investigation, one of the computers was located at the home of an individual who claimed he had bought the computer from another individual. The individual he bought the computer from had

been identified as a suspect in similar cases. The suspect was interviewed and denied any knowledge or involvement in the theft and agreed to undergo a polygraph examination. The results of the examination indicated deception. During the posttest interview, the individual admitted to the theft of the computer and the thefts of five other computers, totaling about $5,000.00. Prosecution is pending.

An Air Force member was suspected in the unattended death of her infant child. The member claimed that she went to sleep at 11.p.m. and her child was asleep in the crib. Sometime during the night the child awoke and the mother moved her to the couch in the living room. She checked on the child at 0830 the following morning and noticed that the child's skin color was blue and her jaw was tightly clenched shut. She called 911 and attempted CPR. She believed the child had a seizure because she had one before. The mother denied doing anything that would have caused the child to stop breathing and agreed to undergo a polygraph examination. The results of the polygraph examination indicated deception. During the posttest interview, the mother admitted that she had held the child's face in the couch cushion until she stopped breathing. The mother was prosecuted and convicted of the death of her child.

IV. Training and Qualification Standards for Department of Defense Forensic Psychophysiologists (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 1999, the Institute trained 49 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 conduct 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 21 difference specialized courses in forensic psychophysiology and related disciplines. In Fiscal Year 1999, approximately 470 student attended the specialized courses.

Department of Defense Forensic Psychophysiologists (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%
1999	144	19%

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 1999, the Institute funded three grant proposals and granted extensions for continued research of two previous grants. Efforts to increase PDD related research have resulted in receipt of no less than 12 new proposals. It is anticipated that DoDPI will be able to fund three or four of the proposals in FY 2000.

The Institute has a prioritized research plan which was approved by 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. This plan has been approved in its entirety by the Personnel Security Research Subcommittee. To organize this aggressive agenda, in FY 1999, the Institute hired a new chief of the Research Division. Efforts to establish strategic partnerships with the university laboratories and increase the breadth of PDD research have begun.

The Research Division at the end of FY 1999 was comprised of two research psychologists, two research assistants, one graduate assistant and a polygraph research officer assigned from a non-DoD agency. Other personnel support for the Research Division was secured by contract, co-op students from a local university, and from the Instruction Division staff at the Institute. The Research Division anticipates hiring three additional research scientists in FY 2000.

Current Research Projects

An Examination of Response Parameters of Electrodermal Recording (EDR) to Standard Stimuli. The objective of this project is to determine if equivalent EDR responses are obtained to equivalent psychological stimuli presented at different EDR tonic levels. The goal of the research is to determine if resistance or conductance is a more accurate measure during PDD examinations.

Effects of Augmented Physiological Feedback on the Detection of Deception. This project will determine if augment feedback improves the accuracy of PDD examinations.

Detecting Stress in the Voice. This is a collaborative study between DoDPI and the Chief, Department of Neuroendocrinology and Neurochemistry, Division of Neuroscience, Walter Reed Institute of Research to determine if stress in the voice is related to biomedical measures of psychological stress. Biomedical measures were shown to be reliable indicators of human stress; however, there was no correlation between those measures and those

measures provided by a computer voice stress analyzer (CVSA), a device currently used by many non-federal law enforcement agencies to detect deception. The utility of other analytic technologies is not ruled out and should be examined. This project is complete and a report will be published in FY 2000.

Thermal Imaging During a PDD Examination. Infrared thermal imaging, a non-intrusive and non-invasive technology, was used to determine if peripheral changes in skin surface temperature (SST) are related to psychological stress. Preliminary studies have shown that stressful tasks cause a decrease in SST on the dorsal surface of the hands. Additional studies need to be conducted to determine if the technology is useful for PDD.

Scaled P300 Scalp Profiles in Detection of Deception. Previous electroencephalographic (EEG) studies of deception have been limited to changes in the amplitude of responses to specific questions. An investigator at Northwestern University has been awarded funds to pursue an innovative and unique measure of deception, the distribution of EEG activity over the scalp. A preliminary report showed favorable results and the contract was extended.

A Field Study to Test the Validity and Comparative Accuracy of Voice Stress Analysis Measured by the Computerized Voice Stress Analyzer: In a Psychophysiological Context. DoDPI awarded funds 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. The data has been collected and a report is anticipated in early FY 2000.

Vagal Tone Monitor/ARIS. This project is designed to determine 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 the ARIS software are designed to measure the direct influence of the vagal nerve on heart rate. Data collection has been completed and an analysis is being done. The final report is anticipated in early FY 2000.

PolyScore 3.3 and Psychophysiological Detection of Deception Examiners when Scoring Examinations from Actual Criminal Investigations. This study was designed to examine PolyScore and human examiner accuracy rate using data collected during actual criminal investigations. A set of 100 Zone Comparison Test (ZCT) and a set of 100 Modified General Question Test (MGQT) format examinations were scored by six examiners (three per test format) using a 7-position scoring scale. The data were later converted to a 3-position scale for comparison. The examinations were also scored using a computerized scoring algorithm, PolyScore 3.3. The results showed that the examiner decisions were more accurate when using a 7-position scale, and they generated more correct decisions for the MGQT format than for the ZCT format. PolyScore was more accurate when scoring ZCT examinations, as opposed to MGQT examinations. PolyScore was also more accurate than the examiners when scoring ZCT examinations. The final report has been completed.

Intrarater Agreement of Psychophysiological Detection of Deception Examiners when Scoring Examinations from Actual Criminal Investigations. This is a follow-up study to the PolyScore study above. The examiners who participated in the Intrarater study were asked to score the same ZCT and MGQT examinations in order to assess the reliability of their decisions and scoring procedures. The data has been evaluated and a report will be published in FY 2000.

Effects of Prior Demonstrations of Polygraph Accuracy on Outcomes of Probable Lie and Directed Lie Polygraph Tests. Investigators at the University of Utah are examining the usefulness of administering an acquaintance test during a PDD examination. A known solution peak of tension, or acquaintance test, is used to reassure examinees that the PDD procedure can accurately detect deception. Unfortunately there have been no systematic studies to determine the validity or efficacy of this procedure. This study will examine the usefulness of the acquaintance test and also compare the directed versus probable lie comparison questions. The report will be completed in FY 2000.

Validity of Outside-Issue Questions in the Control Question Test. Investigators at Boise State University are studying the validity of Outside-Issue questions in the Control Question Test. Outside-Issue questions are those which address topics that are not included in relevant comparison questions (i.e. Do you believe I will only ask you the questions we have reviewed? Is there something else you are afraid I will ask you about?) A report is expected in FY 2000.

A Comparison of Decision Accuracy Rates Obtained by Computer Programs Designed to Evaluate PDD Examination Data. Four vendors currently sell computer programs purported to accurately evaluate PDD examination data. The accuracy of the computer programs has not been independently evaluated. This study will be completed in FY 2000.

Psychophysiological Detection of Deception (PDD) Accuracy Rates Obtained Using Test for Espionage and Sabotage: A Replication. DoDPI developed a new security screening examination procedure in the early 1990s, the Test for Espionage and Sabotage. While high accuracy rates were obtained, the number of observations per sample cell was relatively small. This is a replication study with a larger sample to validate the previous results.

Anticipated Projects for Year 2000

Research Training in Cognitive Psychophysiology and Detection of Deception. A strategic partnership agreement with the University of South Carolina is expected next year. This partnership will provide research in cognitive psychophysiology emphasizing brain process and the detection of deception and research training in these areas. More specifically this project will be comprised of research into the cognitive process occurring in the detection of deception, the brain areas underlying the cognitive activity, and the cognition-brain-deception relationship. This research will be conducted using high-density EEG/ERP recordings. Moreover, this effort will investigate and localize deception-specific critical sources and the effects of deception on ERP topography. In a subsequent phase, this research will investigate and correlate these findings with current autonomic nervous system recordings during a PDD examination. Specialized training in cognitive neuroscience will be provided to the DoDPI staff and the PDD community.

Remote Sensing of Emotion and Stress Using Laser Doppler Vibrometry. This multi-disciplinary project involving investigators from the medical, psychological, computer engineering and physics disciplines will use emerging technologies to develop methods for deriving simultaneous information from the Laser Doppler signal regarding multiple psychological functions including body tremor, respiration, cardiac function, muscle contraction, and sweating. Laser Doppler Vibrometry recording methods do not require the attachment of physical transducers and could be adapted to multiple examination settings. This technology, if successful, would be immune to many of the artifacts that are problematic during traditional PDD examinations.

Laboratory Assessment of the Accuracy of the CQT: The Effect of Culture on the PDD Examination Process. Security organizations in the future will need to accept employees from diverse ethnic backgrounds. The critical issue is how to best evaluate risk given the potential employee's background. This project will investigate the cross-cultural effects on the outcome of PDD examinations administered by both Chinese-American examiners and the natural American examiners on examinees of both cultures. These languages and cultural effects have the potential to modify the training of federal polygraph examiners and could suggest test format modifications for the examination of diverse populations.

"Polyplot" – A Computer Program for Generating and Modifying Polygraph Charts. DoDPI instructors currently use polygraph charts produced during field cases for instructional purposes. These polygraph charts do not always provide textbook examples of key physiological patterns nor do they allow the instructor and students to consider subtle but important variations in the tracings. In response to this problem, DoDPI has developed a statement of work for a computer program that allows DoDPI

instructors to generate and modify hypothetical polygraph tracings.

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 is being subjected to statistical approaches 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.

Pretest Interview Project. This project will identify important variables in the interview that precedes PDD examinations, the pretest interview. The pretest interview is an essential part of the PDD process; however, it has not previously been subjected to scientific analysis. Videotapes of pretest interviews are being coded to record the behavior of examiners and examinees during the pretest interview to identify variables, which may be used to predict the validity of the PDD examination.

Other Activities

International Use of Psychophysiological Detection of Deception. The Institute maintains contact with PDD examiners in other countries to keep abreast of PDD development around the world. The Institute issues periodic reports summarizing international PDD activity. The use of polygraph in other countries has increased dramatically.

Presentations. DoDPI presented results of their research to several audiences in FY 1999, including the Federal Interagency Polygraph Seminar, the annual seminar of the Society of Psychophysiological Research, and the annual seminar of the American Polygraph Association. The Institute has also provided formal instruction to federal examiners at courses sponsored by the University of Virginia, the FBI National Academy and the basic and advanced courses at DoDPI.

The DoDPI staff taught a total of 23 courses to more than 600 students within the federal polygraph community. Additionally, the DoDPI staff provided PDD courses and seminars to more than 300 students in the local and state law enforcement community.

VI. Plans for Expanded Use of the Polygraph

In an effort to eliminate or reduce the number of unauthorized disclosures of classified information to the media, we plan to implement a new policy. As a condition for access to Top Secret, SCI or higher information, DoD military, civilian personnel and contractor employees will sign a form certifying that they are willing to undergo a specific issue polygraph examination if classified information they had access to has been leaked. We believe this will serve as a deterrent to individuals who may be considering leaking classified information.

A Field Validity Study of the Integrated Zone Comparison Technique

Nathan J. Gordon, William L. Fleisher, Hisham Morsie, Walid Habib and Khaled Salah

Abstract

This field study tested the validity of the Integrated Zone Comparison Technique (IZCT) designed for specific issue tests, utilizing 309 confirmed field cases by examiners of the Egyptian Government. During 1998 and 1999 the IZCT correctly identified 100% of the innocent examinees and 99.5% of the guilty examinees, excluding Inconclusives, or 94.8% of innocent examinees and 90.5% of the deceptive examinees, including Inconclusives. A detailed explanation of the technique is included in this report.

Key words: Egypt, field study, Integrated Zone Comparison Technique, validity

This field study is the first published research on the polygraph Integrated Zone Comparison Technique (IZCT). Its theory and methodology were published in the textbook, *Forensic Psychophysiology; Use of The Polygraph* (Matte, 1996). The IZCT is a modification of the Backster Zone Comparison Technique. The IZCT's structure is comprised of three probable-lie comparison questions and three relevant questions, similar to the technique used and validated by the University of Utah (Raskin, Barland, & Podlesney, 1977). In the Raskin et al. study the experimenters reported 95% accuracy for the technique in correctly identifying truth from deception. Like the Utah technique, the IZCT also starts with a number test, followed by a minimum of three crime charts. The IZCT has been taught at the Academy for Scientific Investigative Training since 1987, and is currently being used by law enforcement, government, and private examiners in the United States, Egypt, Israel, Switzerland, South Africa and France. The results of this study apply only to the IZCT when used without deviation. The IZCT may be used in single-issue, as well as multi-issue tests.

Originally, Academy students were taught that, when they had an examination which only involved a single issue, they should use the Backster You Phase Zone Comparison Technique. If they had an examination which lent itself to a multi-issue format, they were instructed to use a Backster S-K-Y, Reid, MGQT, or Arther technique.

In 1987, the IZCT was originated by Nathan J. Gordon, Dr. William Waid, and Philip Cochetti. Dr. Waid suggested that a technique be designed which utilized sound formatting principles, but allowed the examiner the flexibility to use a common test structure in both single-issue and multi-issue cases.

The format for the IZCT is as follows:

1. Irrelevant	Is your first name _____?
2. Outside Issue	Do you understand that I will only ask the questions I reviewed?

Nathan Gordon is the owner and principal instructor for the National Academy for Scientific Investigative Training, and regular contributor to *Polygraph*. Reprint requests should be sent to Nathan Gordon, Academy for Scientific Investigative Training, 1704 Locust, 2nd Floor, Philadelphia, PA 19103, or by e-mail to truthdoctor@polygraph-training.com.

3. Weak relevant — Do you intend to deliberately lie to any test question?
4. Irrelevant — Were you born in ___?
5. Comparison — During the first (-2 years from age at time crime was committed) years of your life, did you ever _____?
6. Flexible Relevant
7. Optional Irrelevant — Is your last name ____?
8. Comparison — In your entire life, did you ever (similar comparison issue as utilized in question 5)?
9. Flexible Relevant
10. Optional Irrelevant — Is today _____?
11. Comparison — During the (-2 years from age at time crime was committed) years of your life, did you ever (similar comparison issue as utilized in question 5)?
12. Flexible Relevant
13. Countermeasure — Did you deliberately do anything to try and beat this test?

Technique Description

The following describes the reasoning behind the questions utilized in the IZCT:

Question 1: Neutral question, used to acclimate the examinee to the test and to establish a norm.

Question 2: We believe that the traditional symptomatic questions (Q25: Do you believe me when I promise you I won't ask you a question we haven't gone over word for word?; Q26: Even though I promised I would not - are you afraid I'll ask you a question we haven't gone over word for word?) developed by Backster (1979)) were very cumbersome. Gordon et al believe that reviewing two symptomatic questions may in some cases actually created distrust of the examiner, as the examinee contemplated the rationale for the questions. To address this problem, a simply worded outside issue question is asked at the beginning of the test as a safeguard against super dampening.

Question 3: In a polygraph examination we are attempting to monitor the flow of the examinee's psychological set to identify those question stimuli, comparison or relevant, that threaten the examinee the most. Since the problem is we sometimes have truthful people presenting deceptive charts (false positives), Gordon believe that starting the examination by directing the examinee only to the relevant questions with the traditional Backster Q39 (Regarding the relevant issue, do you intend to answer each question truthfully about that?) was inappropriately directing the examinee to focus only on the red zone (relevant questions). The reworded question in the IZCT (Do you intend to deliberately lie to any test question?), by its inclusive structure, forces everyone to lie equally to that question. Thus, Q3 allows examinees to self-set to their greatest threat, be it the red zone (relevant) or green zone (comparison), as they themselves perceive it.

Question 4: Used to reestablish a norm, after question number 3, and before the salient phase (comparison and relevant questions) of the test begins.

Question 5: Because the first chart of the examination employs a green - red format, an exclusive comparison question is used (Gordon, 1999), preventing the deceptive examinee from perceiving it as a relevant question. Going back two years in age, from the time of the crime, maximizes the comparison question time period, while separating it in time from the relevant issue time period.

Question 6: Depending on the needs of the case, the examiner can use a strong or medium relevant question which deals with either direct or secondary involvement.

Question 7: Reviewed irrelevant question, but not used unless needed to reestablish the norm during the test following a continuing physiological reaction or artifact.

Question 8: Inclusive comparison question is asked, since a relevant question precedes it. This comparison question may be used as a competitive comparison question, which is extremely useful in employee theft type examinations.

Question 9: Examiner can use a strong or medium relevant question that deals with either direct or secondary involvement, depending on the needs of the case.

Question 10: Reviewed irrelevant question, which is not used unless needed to reestablish a norm someplace during the test.

Question 11: Exclusive comparison question that maximizes the time period covered, ensuring a valid probable-lie issue.

Question 12: Examiner can use a strong or medium relevant question that deals with either direct or secondary involvement, depending on the needs of the case.

Question 13: Used as weak relevant to withdraw the examinee from the test mode, but often provides useful information for the examiner.

IZCT Sequence

The examination begins with a known-number demonstration (stim) test. The examinee is asked to pick a number between 2 and 5, and reveal the choice. The examiner then administers a single chart instructing the examinee to answer "no" to every question from 1 to 6, including the number actually selected. The examiner then explains that this allows for evaluation to ensure that if he or she lies, the polygraph procedure can detect it.

The IZCT sequence is then administered as follows:

Chart 1. Silent answer format
1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13

Chart 2. Mixed out loud sequence, rotating the positions of the relevant questions
7, 2, 5, 12, 8, 6, 11, 9, 3, 13

Chart 3. Reversal of the comparison-relevant (green-red) sequence, to a relevant-comparison (red-green) sequence
1, 2, 3, 9, 5, 12, 8, 6, 11, 13

Silent Answer Test

The Silent Answer Test (SAT) format, originated by Reid and Horvath (1972), is used in Chart 1 with the IZCT. In the Reid technique this format is only used in the latter part of the testing process when the examiner is having problems making a decision. Reid and Horvath reported among other advantages, that the SAT produced enhanced electrodermal reactions, and, "even if the subject failed to react significantly on the SAT, it tends to induce greater responses on the later tests." The Utah technique also utilized the SAT on the fourth crime chart if the first three charts were not conclusive (Raskin, Barland & Podlesney, 1977.)

In the Integrated Zone Comparison Technique, the SAT is introduced as the first crime chart in the following manner:

> "In this first test I am going to ask you the questions I just reviewed with you. During the test I don't want you to answer out loud. I just want you to listen to the questions one more time, get used to being attached to the instrument, and having me ask you questions. It is to make sure you have understood all of the questions, feel comfortable with them, and most importantly, that you have answered every question truthfully. If you remember anything you haven't told me about, you can tell me as soon as the test is over, but don't say anything out loud during this first test: just listen."

The purpose of using the SAT in the first chart is that most examinees do not consciously perceive the chart as a threat, since they are not answering out loud, and lying. Thus, they rarely attempt any type of mental or physical countermeasures. This not only results in an excellent chart of

physiological tracings, but also excellent reactions to the appropriate zone of questions where deception will be attempted. It is not the utterance of "yes" or "no" that creates sympathetic nervous system arousal. It is the examinee's cognitive recognition of the threat the questions pose to his or her well being in a format to which they will attempt deception.

Golden, in the Listen-Answer Technique (1992), hypothesized that maximum psychophysiological stress would be generated during the presentation of an incriminating question, when the person was instructed just to listen, and that vocalization to the same question would actually allow a degree of psychophysiological relief. He made the analogy that for the deceptive person, not to be able to utter his or her lie was like a person stubbing a toe and not being able to yell out in pain, thus resulting in greater psychophysiological reactions to take occur.

In Chart 2, the examinee is instructed to answer out loud each question truthfully. The examiner further instructs that lying to any question, regardless of which question it is, could result in the examinee failing the test. This verbal stimulation further helps self-set examinees to the zone that poses to them their greatest threat.

During this chart the relevant question positions are rotated to allow each relevant question to be next to a different comparison question. This will ultimately pair each of the relevant questions with each of the comparison questions once during the three-chart examination. Mixing of the question order is done as a safeguard against habituation and anticipation.

In Chart 3, the sequence of the chart is reversed from a green-red format (comparison-relevant), to a red-green format (relevant-comparison). Gordon and Cochetti (1982), assert that green-red testing formats leaned a test toward truthfulness, and red-green testing formats leaned a test toward deception. In the IZCT, the first two charts are biased toward a truthful outcome, and the third chart is biased toward a deceptive outcome. We believe that this reversal of the question order in the IZCT safeguards against both false positives and false negatives, giving the overall process a more accurate and balanced conclusion. Furthermore, we do not believe that the IZCT test structure has any impact on the number of inconclusive outcomes.

Procedure

As pointed out in Matte (1996), there is very little literature on validity and reliability of polygraph techniques in field cases. All polygraph examinations used in the present study were conducted by four examiners of the Egyptian Government from 1998 to 1999. All four examiners were trained at the Academy for Scientific Investigative Training. All of the examinations were single-issue tests, utilizing the format cited. Charts were scored by the Horizontal Scoring System (Gordon & Cochetti, 1987) and the Academy's Algorithm for Manual Scoring (Gordon, 1999).

The polygraph utilized in the examinations was the Lafayette LX 2000 Computerized System, which monitored thoracic and diaphragmatic breathing, electrodermal response, and cardiovascular activity.

Results

Five hundred seventy-six examinations were conducted, 47 of which were re-examinations due to initial Inconclusive results. The 47 Inconclusive examinations had been classified as such because the numerical scores had not met the threshold established for the Horizontal Scoring System of a $\pm$13 (Honts & Driscoll, 1987), for a single-issue examination consisting of three relevant questions, administered over three charts. Of these 576 examinations 309 were confirmed by confession, or judicial convictions. Among the 309 verified cases, 288 were confirmed by confession, and 21 by judicial conviction. It should be noted that in the examinations confirmed by conviction, polygraph results played no part in the judicial decision.

Using the 309 confirmed cases, the decisions of the polygraphists were 210 Deception Indicated (DI), 74 No Deception Indicated (NDI), and 25 Inconclusive (Table 1). Of the 232 examinees that were later confirmed by confession or judicial conviction to be guilty, 210 were determined by the

polygraph examiners to be deceptive, 21 were called Inconclusive, and 1 was determined be truthful. The polygraph examiners correctly identified 99.5% of the deceptive examinees in this sample, excluding Inconclusives, and 90.5% including Inconclusives. Of the 77 confirmed truthful examinees, 73 were called NDI by the polygraph examiners, and 4 were called Inconclusive. The polygraph examiners correctly classified 100% of the truthful examinees when Inconclusives were excluded, and 94.8% including Inconclusives.

One hundred three of the 309 examinations had documentation of the results of a second examiner's blind numerical analysis of the polygraph charts using either the Horizontal Scoring System, or the 7-position scale as utilized in the Utah System (Weaver, 1980). These blind evaluators reached the same conclusion as the original examiner 96 out of the 103 re-evaluations, or 93.2%.

Table 1. Ground truth and polygraph decisions for 309 confirmed field cases.

		Ground Truth		
		Deceptive	Nondeceptive	Total
Decisions	DI	210 (90.5%)	0 (0.0%)	210
	Inconclusive	21 (9.1%)	4 (5.2%)	25
	NDI	1 (.4%)	73 (94.8)	74
	Total	232 (100%)	77 (100%)	309

Conclusion

This validation study demonstrates the efficacy of the Integrated Zone Comparison Technique. The IZCT is an innovative and powerful technique which ethically directs the examinee to his or her proper zone of greatest threat, resulting in accurate determinations of truth or deception.

Because the present study was conducted within the Egyptian Government, raw data were not readily available to independent evaluators. Future research could address the IZCT in a non-military setting. In addition, studies should also be undertaken to compare results with the IZCT testing method using scoring systems other than the Academy's Horizontal Scoring System, and Examiner Algorithm for Manual Scoring. In the interest of scientific replication, manuals explaining the IZCT and Academy's scoring systems are available through the Academy for Scientific Investigative Training, 1704 Locust Street, Philadelphia, PA 19103.

References

Backster, C. (1979) *Standarized Polygraph Notepack and Technique Guide*.

Golden, R. (1992). *Ready, Aim, Hire; The Riddle of the Lie; A Scientific Approach To Lie Detection* Pages 81 - 101. PerSysCo Publishing, Oak Brook, IL.

Gordon, N.J. (1999) The Academy for Scientific Investigative Training's Horizontal Scoring System and Examiner's Algorithm for Chart Interpretation. *Polygraph* 28(1), 56-64.

Gordon, N.J. (1999) Inclusive vs. Exclusive Controls. *American Association of Police Polygraphists' Journal.*

Gordon, N.J. and Cochetti, P.M. (1982). The Positive Control Concept and Technique. *Polygraph* 11(4), 330 - 342.

Gordon, N.J., & Cochetti, P.M. (1987). The Horizontal Scoring System. *Polygraph* 16(2), 116-125.

Honts, C.R., & Driscoll, L.N. (1987). An evaluation of the reliability and validity of rank order scoring and standard numerical scoring of polygraph charts. *Polygraph* 16(4) 241-257.

Horvath, F.S. and Reid, J. E. (1972). The polygraph silent answer test. *Journal of Criminal Law, Criminology and Police Science* 63(2), 285 - 293.

Matte, J.A.(1996). *Forensic Psychophysiology Using The Polygraph.* Williamsville, NY, J.A.M. Publications.

Raskin, D.C., Barland, G.H., & Podlesny, J.A. (1977). Validity and reliability of detection of deception. *Polygraph* 6(1), 1 - 39.

Weaver, R.S. (1980). The numerical evaluation of polygraph charts: Evolution and comparison of three major systems. *Polygraph* 9(2), 94 -108.

The Debate Over Polygraph In Poland

Aleksander Krzyœcin

Abstract

This paper is a history of polygraph examinations in Poland in the years 1950-1999. It will address basic assumptions of the original Polish concept of revealing memory traces of punishable acts. Polemics and the views of Polish opponents of using psychophysiological polygraph examinations in criminal procedure are discussed. There is particular emphasis on such problems as: consciousness in polygraph examinations, and role of interfering variables, freedom of speech of the examinee, importance of polygraph for defense of an innocently accused person, reliability of examinations and possibility of preventing mistakes in court.

Key words: history, law, memory trace, Poland, polygraph, theory

Historical Perspective

Poland has had almost fifty years of experience with conducting various kinds of examinations with the use of polygraph. In the early 1950s, Polish intelligence became interested in this instrument after it was established that American intelligence tested Polish citizens by means of a lie detector before beginning any cooperation with them. They managed to buy a Keeler polygraph. After gathering the necessary experience, they started using it in foreign operations. It was extremely hard to master the examination procedures, as the American authorities imposed an embargo on the polygraph. In 1975 the editorial board of the journal *Polygraph* still rejected a request for a subscription, explaining their policy not to send their publications across the Iron Curtain. The situation did not change until 1994 when a delegation of the American Polygraph Association paid a visit to Poland in order to start contacts with our experts.

In 1969 the Polish military police began using polygraph examinations in criminal cases, mainly those involving theft of weapons. The experience with the examinations was positive, and contributed to clearing many soldiers from unjust suspicions and to revealing actual perpetrators of crimes, who were later convicted by courts. In the years from 1969 to 1998, military bodies performed such examinations in 1180 cases, in which a total number of 5243 persons were tested. Those examinations were based on the technique of probable-lie comparison questions developed by J. E. Reid, and this method was successful in our country. It was confirmed by a study (Krzyœcin, 1979) in which verified results of polygraph examinations were analyzed and compared against independent criteria.

The results of examinations were used in evidence proceedings before courts. The Supreme Court of Poland expressed its view on the admissibility of polygraph in criminal proceedings on several occasions. For example, in September 1976, the Supreme Court decided in case No. II KR 171/76 that this kind of evidence could not be used as independent evidence providing the grounds for making definite decisions. Thus, such evidence is of an auxiliary nature. Polygraphic examination, according to this court, is concerned primarily with showing that the examinee is emotionally linked with the given event.

Nowadays, polygraphs are being used in five agencies of the Polish Government: the Office of State Protection, Military Information Service, Military Police, Frontier Guard and Chief Customs Inspection. The main bulk of these examinations are tests checking applicants for jobs. The Polish Police do not

For reprints, contact Dr. Aleksander Krzyœcin,Bulgarska 34 Street, 03-971 Warsaw, Poland

have their own experts. Sporadically, in the most difficult criminal cases, mainly cases of murder, the police order such examinations to be conducted by experts from other institutions.

There are three chairs of crime detection studies in Poland, which conduct research in the field of psychophysiological polygraph examinations: the Silesian University in Katowice, the University of Wroclaw and the Copernicus University in Torun. Prof. M. Kulicki who works in the last of these institutions, is a supporter of D. T. Lykken's theory. He rejects the comparison question tests, claiming the absolute reliability of the test of hidden knowledge of the act.

Since 1990, private detective agencies have tried in vain to draw the attention of businesses to polygraph. Nowadays, there are only a handful of retired experts who perform polygraph examinations on request of private businesses. Most of these examinations are not personality-type tests connected with employment, but examinations related to offenses committed by employees to the detriment of the employer.

In recent years some theoreticians and practitioners in the field of criminal procedural law have tried to eliminate the polygraph from court proceedings. They attempted to do so by inserting an article in the new Code of Criminal Procedure, which prohibited the use of this instrument for the purposes of evidence. The codification commission working on the draft of this code did not consider the opinions of experts in the field of psychophysiological polygraph examinations.

The Polish National Prosecutor's Office is definitely against the use of polygraph in the operations and recognition activities of the police. This view was expressed, inter alia, in the answer to an inquiry sent from the Police Headquarters in 1994. It was justified in the following way: "if the procedural law establishes some safeguards limiting the duty to provide information to a State authority ... it follows from this that outside of the trial, in the course of operations, one cannot do what the safeguards do not permit."

There are fundamental procedural differences between the Polish criminal procedure and the American one. We do not have a court with a jury in our country. Prior to court proceedings, the prosecutor conducts preparatory proceedings in the form of investigation. Its purpose is to verify whether a crime has been committed, find the culprit and collect the evidence for the court issuing the judgment. In Poland, it is impossible to claim that the aim of polygraphic examination is to assess the credibility of a statement of the accused or the testimony of a witness. The court will not allow evidence in the form of a polygraph expertise formulated in terms of deception - nondeception.

Concept of Revealing Memory Traces

After 1975, a group of Polish crime detection researchers and experts in this field concluded on the basis of their experiments that the American theory of detection of deception did not provide sufficient explanation for the physiological phenomena registered during the tests. Consequently, this solution was rejected. During the past 25 years, a new concept of psychophysiological polygraphic examination (PPE) was invented and developed. The model of this kind of examination consists of four basic elements:

1. The character of this examination is to reproduce memory traces;

2. Examination procedure takes account of the principles used in psychological experiments;

3. The examination is a method of criminological identification;

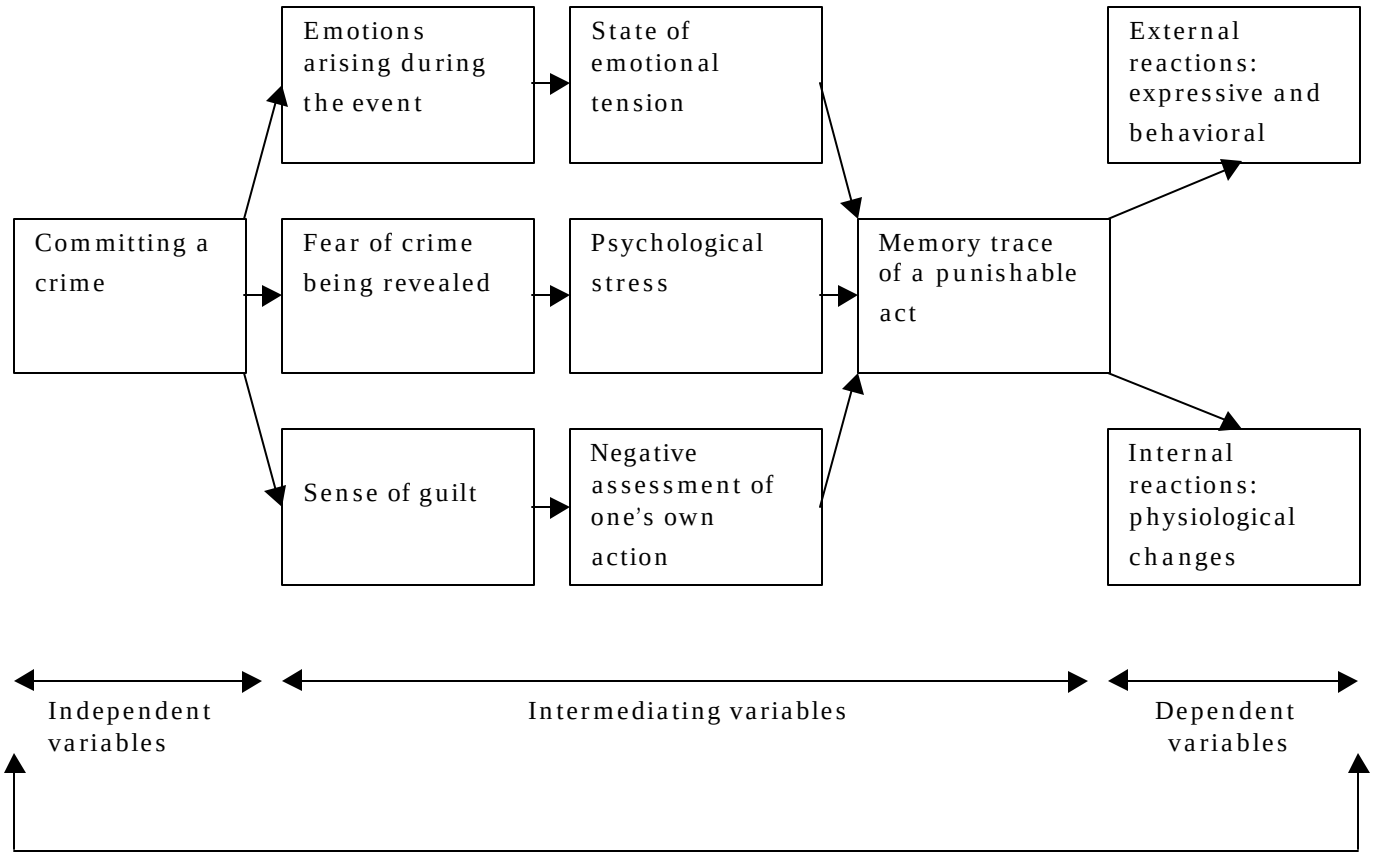
4. The examination is aimed at retrieving information needed by law enforcement agencies.

The above concept is based on the assumption of revealing memory traces of criminal offences. The starting point is the thesis that each experience a human being has gone through leaves its own separate trace in the nervous system. The information about delinquent activities and the strong negative emotions accompanying them are recorded in the long-term memory of the perpetrator. The

fact of committing a serious crime (and only such crimes should be the object of polygraph examination) is, as a rule, connected with shock or strong emotions, which reside in a special place in the memory. Many of such shocking experiences remain fresh in the memory, particularly in the memory of an active participant of the event. The memory trace of an act corresponds to the criteria that should characterize a criminological trace, namely it is invariable, unique and allows identification. The examinations carried on in our country by the majority of experts comply with the high methodological requirements set for experimental research in psychology.

Since memory traces cannot be directly observed, it was accepted that the indicatium is the existence or lack of a trace of a punishable act in the brain of the person undergoing examination. The observable indication is the changes in the functioning of various systems of the human body. The methodology of examination consists in manipulating with the psychological variables and observing their influence on the physiological variables, which are objectively measurable, just like it happens in any psychophysiological experiment. There are three types of variables that occur in such examination. The independent variable is the link between the examinee and a given criminal event or lack of such link. The dependent variable denotes changes in physiological parameters, which are provoked or not by specific test questions. In the polygraph examination, like in any psychological experiment, there occur interfering variables. Figure 1 below describes the mechanism of origination and consolidation of the memory trace of a specific punishable act and how it is revealed during a polygraph examination.

Figure 1. Memory trace model for polygraph testing.



We have to emphasize that this illustration is not an exhaustive one and it serves only as an example of the factors contributing to the creation of a memory trace of an act. It is possible that in a given offender, the emotion of sense of guilt may never arise, while other emotions not shown in this chart may still arise. The correct diagnosis has a dichotomic nature. It is formulated in terms of positive identification - negative elimination. Thus an expert does not reveal lies or evaluate the reliability of explanations or testimonies.

Polygraph in the 1997 Code of Criminal Procedure

Article 171 paragraph 4 subparagraph 2 the Act of 6 June 1997 - Code of Criminal Procedure (Official Gazette 1997, No. 89, item 555) introduced, inter alia, a prohibition against the use of technical means aimed at verifying unconscious reactions occurring in the body of the interrogated person. The authors of the governmental justification of a draft of CCP (updated in 1997 by the Department of Training and Human Resources of the Ministry of Justice) concluded in their interpretation of this prohibition that it applied to testing with the use of polygraph, mistakenly referred to by some authors as the variograph.

The above-mentioned document formulated a number of objections against polygraph testing. They are as follows:

1. The object of examination is the issue of reliability of explanations or testimonies, that is the assessment of evidence, which is undoubtedly an exclusive task of the court;

2. The examination is an external intervention in the domain of subconsciousness;

3. When one is charged with a serious crime, it is a stressful situation, in which the accused, in spite of being innocent, is going to react to statements connected with the act for which he or she is incriminated;

4. The examination is contrary to the humane assumptions of criminal procedure, its objectives and principles, as well as safeguards of the rights of an individual in criminal proceedings;

5. Theoreticians and practitioners of the criminal procedure present in principle a uniform negative attitude to such examinations;

6. By definition, these examinations are connected, at least indirectly, with the law enforcement agencies, and our society, still remembering those not so remote times, is particularly sensitive to that.

Many authors have criticized the ideas expressed in the justification of the governmental draft of CCP. Sobolewski (1998) states that the force of the arguments quoted above is so small that one can leave them aside. Wojcikiewicz (1995) holds that the arguments brought in favor of the absolute ban on using polygraph are not always substantively correct. Jaworski (1999) is of the opinion that the solutions provided in CCP as regards polygraph testing are based on dubious, completely mistaken assumptions.

Not all representatives of the legal science agree with the argumentation of the authors of CCP. Waltos (1998) believes that the new CCP does not prohibit the use of polygraph. It is only forbidden to use this instrument during interrogations. Bulsiewicz (1998) emphasizes the fact that in the draft of CCP of 1994, article 194 paragraph 1 stipulated that "in the examinations carried out by court experts it shall be inadmissible to use the means and methods specified in article 168 paragraph 4" (now article 171 paragraph 4) and was deleted. This confirms that the 1997 CCP does not prohibit conducting polygraph examinations. Gruza (1999) argues that under the new provisions contained in CCP polygraph testing is still an admissible means of evidence.

Widacki (1999) writes in his commentary to the decision of the Supreme Court of 21 December 1998 (IV KO 101/98) that it was the first time the Supreme Court expressed its opinion about the permissibility of polygraph (variographic) examinations since the entry into force of the new CCP. The Supreme Court was unanimous in stating that polygraph expertise may be included in the

evidences and may be one of the grounds for ascertaining guilt.

The Appellate Prosecutor in Lublin, A. Witkowski, in an interview for *Zycie*, No. 24 (708) published in January 1999, stated that he could not imagine his work without the polygraph helping to adopt the right direction of investigation. His practical experience shows that thanks to the polygraph it was possible to bring murderers before the court or clear from allegations persons charged with crimes. On 16 October 1998, already under the new CCP, the Voivodship Court of Radom in case No. II 93/97 following a request of the accused person, appointed a specialist in the field of polygraph testing to determine whether or not the nervous system of the accused had recorded any memory traces connected with the act he was charged with. The nature of this article does not allow us to quote the opinions of other supporters of polygraph, however, even these examples disqualify the thesis that the theoreticians and practitioners of criminal procedure represent a unanimous anti-polygraph attitude. The statement that the majority of this group believe that the use of polygraph in proceedings should be inadmissible cannot be a criterion for measuring the value of polygraph expertise. In science it is not the majority of votes that determines what is true. The truth of a scientific thesis must be proved empirically.

The above set of arguments concerning polygraph seems rational. However, an analysis thereof shows that these theses are just verbal, not based on results of scientific research. Thus they have to be treated as unverified hypotheses. Jaworski (1999) presents a similar view stating that any objections raised against the polygraph are made in a categorical form, which makes one think they are based on solid scientific grounds. As a matter of fact, when we try to find justification for these objections, they prove to be based predominantly on introspection. There is no scientific research to support them.

First of all, it must be emphasized that so far nobody has succeeded in proving that it is the unconscious body reactions of the examinee that are controlled during a polygraph examination. On the basis of modern achievements of such disciplines as neurophysiology and psychology, it may be possible to prove the above thesis to be untrue. The opinion that the body reactions of the examinee are conscious has good scientific grounds. The authors trying to discredit this instrument do not take the trouble to objectively consider the arguments of their opponents, simply keeping quiet about them and ignoring them.

The commentaries to CCP from the period before 1997 represent identical opinions as those expressed in the governmental justification. For instance, the commentary Bratoszewski et al. (1998), reads "the legislator prohibited the use of polygraph, thus resolving the dispute over the admissibility of this instrument in the negative". We have to emphasize the fact that the name "polygraph" is not used in the statute.

The opponents of using polygraph in trials form their opinions about it on the basis of their own false visions of these examinations and not on the foundations of the actual state of affairs.

The Problem of Consciousness in Polygraph Examinations

In a two-volume commentary to CCP edited by Hofmanski et al. (1999), we find that "there is no doubt that the essence of variographic examination lies in controlling the organic reactions which the examinee is completely unaware of and which are beyond their control." In my view, this opinion is based on a traditional psychological thesis, which is now a historical one. Kofta (1979) wrote that Hilgard (1967 publication in Poland, 1953 in UK) pointed out that there was at least one aspect of emotional states, which was not subject to arbitrary control and that this was the set of physiological symptoms of emotions. Kofta believes that Hilgard's thesis cannot remain in such a radical tenor. In his view a human being can easily learn to control their physiological functions.

Psychologists Lindsay and Norman (1972) argue that the latest research leads to rejecting the opinion that people have no

influence over what is happening in their bodies. In their view, a person whose nervous system was aroused may observe all changes occurring in their body and conclude that they experience fear. LeDoux (1998), a neurophysiologist, holds that the emotions characteristic of a given individual, expressing his experiences of past situations as well as anticipations regarding the current situation, remain under the control of their will. The type of reactions LeDoux refers to are characteristic of the agent, as they are connected with past experiences acquired at the time of committing the crime. Poeppel (1989) stated that "if a certain state of affairs cannot become conscious, which signifies that it is closed from the consciousness, then there is no possibility to test it at all" (underlined by the author). It follows from the above opinion that if the polygraph indeed checked the unconscious processes or dug as deep as the subconsciousness, it would be completely impossible to carry out such examination.

The empirical evidences gathered by scientists in the recent years indicate beyond a shadow of doubt that the reactions of persons tested by means of polygraph are conscious. This thesis can be proved, among other things, thanks to latest methods of observing brain activity. They give insight into the processes occurring in this organ. Even the potentials method used in electroencephalography (EEG) provide evidence to support that the processes occurring during a polygraph examination are absolutely conscious. The examinee consciously receives a verbal stimulus in the form of a test question, consciously experiences its content and is fully aware of the changes occurring in his or her body after being exposed to critical questions. Examinees are able to describe such changes, which is completely impossible in the case of unconscious or subconscious reactions.

The commentary to CCP edited by Hofmanski et al. (1999) contains a statement that the use of polygraph examinations in the form of expert evidence is an attempt to circumvent a provision, which has the nature of a safeguard. Regarding this issue, Widacki (1999) justly points out that "arguing that the prohibition expressed in article 171 paragraph 4 subparagraph 2 covers polygraph examinations conducted as part of expert evidence would involve the need to adopt a wider interpretation and to assume that what is forbidden in relation to interrogation is also - automatically - forbidden during the preparation of expert evidence". We may add that in Poland the polygraph has never been and is not used in the course of interrogation.

Emotions and the Possibility of Controlling Them

The objectives of PPE are the emotional processes and their material correlates, that is physiological reactions. Nowadays, emotion researchers answer the question "are we able to control emotions?" in the affirmative. They believe people are able to control their inappropriate emotions quite successfully. Episodes in which a human being is completely deprived of control over emotions are rare. A lot of recent studies are devoted to the ability of individuals to control their body functions. One of the elements of emotional development is the ability to control emotions appearing as early as the age of two or three years. A grownup can control emotions in a number of ways, such as avoiding averse stimuli, inhibiting emotional expression, using external and internal distractors, applying perseverance. Reykowski (1992) writes that "human beings may influence their own emotions by the fact that they take into consideration different aspects of the situation and analyze them in different ways". Clore (1998) argues that "unconsciousness of emotions is impossible, because emotions are involved in experience, and one cannot experience anything without experiencing it". The fact that the examinees are conscious both during the tests of emotional reactions and their sources and that they feel the changes in the functioning of their bodies is a proof that the polygraph is not one of the technical means, referred to in article 171 paragraph 4 subparagraph 2, as the authors of the commentary to CCP claim.

The Role of Interfering Variables in the Examinations

Polish opponents of polygraph maintain that psychophysiological polygraph examinations are worthless and useless in criminal proceedings, as they are subject to interfering variables. They claim that the

symptoms occurring in the examination may follow from intrinsic physiological disturbances not having any link with the physical phenomena. Emotional reactions may also be caused by different psychological experiences, ones which are neutral from the point of view of criminal proceedings. The results of the examination themselves may not only be the results of emotional processes linked with the sense of guilt of the perpetrator of a given crime, but may also be induced in a person experiencing fear of being unjustifiably accused of having committed this crime or even fear of revealing a different act prohibited by law.

The governmental comment to the Code of Criminal Procedure (CCP) states that an innocent person accused of a severe crime will react to statements connected with the act they are incriminated with. This theory is a paraphrase of the opinion presented by Gutenkunst (1965) concerning the factors interfering with the examination. The bibliography shows that the opinion on this matter was formulated on the basis of information contained in *Psychology: US Armed Forces Manual*, edited by E. G. Boring (1943). This work states that "the difficulties in applying the lie-detection method consist in that fact that it is inefficient against an experienced liar, who shows no signs of excitement when lying, but, on the other hand, may give over for a decent man who cannot help being moved simply because he is being examined." The source for this thesis is the literature published before 1939. The Polish publisher of the American manual (1960) warned the readers that certain concepts presented in the manual were erroneous in light of the developments of contemporary science, which was particularly applicable to physiology. As we know, the contemporary PPE has nothing in common with detecting lies. Since WWII, the US Defense Department has completely reviewed its opinion on polygraph examinations. Now, it is a main advocate of the polygraph. If the procedure of these examinations did not actually give the possibility of differentiating between trace reactions of crime perpetrators and the nerves and stress experienced by an innocent person, who happens to be unjustifiably accused, then polygraph tests would be valueless and would have been discredited by now.

The essence of the problem is not the fact that interfering factors may occur in the examination, but the fact whether the examiner is able to reveal them and minimize or exclude their impact on the test results. The whole procedure of examination, and in particular the rules of analyzing and interpreting the charts, prevent us from making a mistaken diagnosis, as a result of the influence of interfering variables.

Freedom of Speech in Tests

Grzegorczyk (1998) emphasized that article 171 of this Code makes the use of technical means, such a polygraph, illegitimate because they exclude, by definition, freedom of speech. The authors discussing this issue usually quote Doda and Gaberle (1995) who claimed that the lack of freedom of speech in polygraph examinations was obvious, as this was the effect of the manner of conducting the tests. If the examinees have to limit their utterances to "yes" or "no", there can be no mention of freedom of speech whatsoever. This view is mistaken. It displays ignorance of the basic elements of examination procedure. The methodology of examination gives the examinees many possibilities of unrestricted expression as regards the acts they are accused of. They are able to present their own version of the event, evaluate the methods of testing and the behavior of the examiner, etc. In each examination, which lasts on the average between two and four hours, 20 percent of the whole time is assigned for tests in which the examinees can only answer "yes" or "no". Also, they may not give any verbal answers when they are instructed to simply listen to the questions they are asked or to repeat the last word of the question. The remaining time is a dialogue between the examiner and the examinee, which is extremely important in the testing procedure. The important thing is that what the examinees say is taken into consideration during the analysis and interpretation of charts, so they do have influence over the results of examinations.

As an illustration we can use the case involving rape and murder (file No. II K 112/96) in which an examination was conducted in Plock in July 1996. This

examination consisted, inter alia, of a test of hidden knowledge. The questions were concerned with the kind of sexual intercourse with the victim. In this test, the strongest stimulus was a statement regarding oral intercourse. When asked about it during the posttest interview the accused explained that he was aroused by the question about oral sex because he had been placed in a reformatory establishment for having raped a ten-year-old girl in that way. Naturally, in this context the changes recorded by the polygraph were not an indicator of a trace reaction. As we can see, the procedure in PPE is a communicative process. The examinee, being an active party in the examination system, brings in his life experience that can modify test results.

Polygraph in the Defense of an Innocently Accused Person

The authors raising objections against the polygraph assume that the results of testing are used only as inculpatory evidence. In reality things are different. As a result of the tests, an overwhelming majority of those examined turn out to be unjustly suspected. The size of this group is differently assessed by various authors, ranging from 80 to 90 percent of all examinees. So the expert polygraph evidence is predominantly used as exculpatory evidence. At a conference on the application of polygraph testing in criminal cases in Katowice in 1978, Waltos (1978) reminded the opinion of Daszkiewicz (1974) "there can be no mention of any coercion in relation to the accused where it is the accused himself who asks for the polygraph to be used ... Can we also say that applying the polygraph in this situation would be a contradiction of human dignity? Would it be moral to refuse to conduct such an examination?" Nowadays, our courts receive more and more requests for conducting a polygraph examination. In a case heard by the District Court in Warsaw, file No. VII K 666/96, the accused wrote: "Dear Sirs, I cannot prove that I am innocent, therefore I decided to request that the Court direct me to an examination of my truthfulness. I am 100 percent sure I want to take part in the test ... The reason why I mention such examination is because this is my last chance". The incorrect term used in this request is the result of the fact that Polish press commonly uses the wrong name "lie detector." In an examination of unjustly suspected persons it is of vital importance that in their memory there are no recorded traces which would link them with an act committed by somebody else. We are dealing here with a situation, which was well characterized by Poeppel (1989) "in order for something to be transferred from memory to consciousness, this something must be provoked by the given situation. If there is no semantic reference to the past event, if nothing in our memory brings the past events, the content of the memory remains silent". Therefore, for an unjustly suspected person, during the examination, questions connected with the case will not be emotiogenic stimuli, causing significant changes in the record of physiological parameters. We can also come across an opinion that considerable emotional traces are formed in an innocent person who is charged with a crime, interrogated and put under arrest. Yet we cannot treat the memory trace of a specific punishable act registered in long-term memory of its perpetrator and the traces of mental and emotional experiences in the same way. These are two separate types of memory traces.

Reliability of Polygraph Examinations

Kruszynski (1998), expressing his approval for article 171 CCP, which prohibits the use of polygraph, claims that polygraph tests are not reliable. Beginning from 1978, this author has stated on numerous occasions that PPEs are uncertain and cannot be relied upon. However, he has never cited any scientific evidence to support the thesis about their "great deceptiveness." Examinations carried out with the help of polygraph are one of those methods of criminal identification whose diagnostic value is subject to most detailed checks. Gradually, as more technically advanced instruments are built and diagnostic methods developed, year-by-year the precision of tests increases. As it is known to the author, there were no persons (among all persons tested in Poland in the period 1969-1997) officially expressing his objections to the results of the polygraph testing (Krzyœcin, 1997).

Since 1980 more than 250 well-documented scientific studies on the adequacy and reliability of PPE have been published in the world. They prove that the results of such

examinations are very precise provided that they are performed by well-trained and experienced specialists. The above-mentioned studies were prepared by psychologists and crime detection researchers. They concerned both experiments involving revealing "perpetrators" of simulated crimes and analyses of results of experiments in criminal cases, compared against independent criteria. In this article, however, we have to limit ourselves to the presentation of a single study, which fortunately concerns a statistically significant population of the examinees. Mason (1998) presented a comparison of diagnoses prepared by experts in the field of PPE with the positive results of chemical tests determining traces of drugs in urine. The proceedings concerned with taking cocaine and marijuana by soldiers were conducted by the army police. After exclusion of unresolved examinations (5%) the reliability of examinations carried on 1920 suspects was as high as 99 percent. On the basis of vast literature we can argue that the reliability of PPE is not lower than the value of evidence obtained by other methods of crime detection, generally accepted by courts.

Possibility of Preventing Mistakes in Courts

Supporters of the prohibition of using polygraph in proceedings seem not to notice the fact that an examination carried out in preparatory proceedings may prevent later mistakes in the court. According to Daszkiewicz (1974) "the difficulty in separating truth from false can often act to the detriment of innocent persons as it can - and sometimes does - lead to unjust charges and unjust convictions. Developing a certain technical means verifying the truthfulness of one's words and using it in the proceedings would help in eliminating such mistakes or at least reduce the risk." Our experts can provide a lot of examples showing that the results of tests helped to clear from allegations unjustly suspected persons in situations where the law enforcement agencies were in possession of considerable presumptive evidence and sometimes even material evidence which - if presented during a trial before court - could be the grounds for conviction. Jaworski (1999) describes four examples of polygraph examinations he had conducted, which were in clear contrast with a significant collection of presumptive evidence and personal versions of prosecutors. Taking the above into consideration, one can prove that PPEs are in line with the humane assumptions of procedural law, its objectives, principles and the safeguards of the rights of an individual. One must bear in mind that if the court rejects an evidence request from the suspect (accused) or their counsel for such examinations to be conducted with the aim of checking whether in their consciousness there is a record of any memory traces of the act they are accused of, this may be the grounds for a justified complaint to the Constitutional Tribunal regarding an infringement of the constitutional right of defense.

Some lawyers, who are prejudiced against PPE, associate it with forcing people to admit guilt. We cannot rule out this danger in cases where examinations are conducted by persons without the requisite knowledge, but this risk can be eliminated by introducing a duty of documenting the examinations in the form of video recordings and by ordering the quality of performed tests to be assessed by experts possessing the highest qualifications. PPE can be fully controlled by trial bodies. An important role in this respect may be played by the courts' developing the principles of admissibility of expert polygraph evidence. There are five basic conditions to be met by such examination to be admitted in evidence proceedings:

1. The examination was performed with the use of a standardized instrument recording at least three physiological parameters.

2. The examination was performed by a specialist of high professional qualifications and ethics.

3. The examiner applied a verified examination procedure, accepted by the majority of experts in this field.

4. The physical and mental condition of the examinee was sufficient for conducting this kind of tests.

5. The expert presented well-documented results of the examination.

The lawyers presenting negative opinions about PPE find it difficult to overcome the well-established mental schemata derived from the concept of "lie detection", which in our country was abandoned many years ago. If the critics of polygraph were really right, such examinations would have been falsified and rejected a long time ago. The history of crime detection studies has seen many methods, which were eliminated because they did not prove correct in the practice of law enforcement agencies, while the polygraph has been proved practically useful for 80 years already. In 1988, polygraphs were effectively used in 57 countries on different continents.

Polish achievements in this field are valued abroad. Matte (1996) wrote, "polygraph technique is successful in culturally different countries, which is confirmed by studies carried out in Poland, Iceland, Israel and USA." However, at present the problem of polygraph in our country is a theoretical one. Its application in criminal cases is insignificant. There are just a few experts who can perform examinations for the needs of criminal law enforcement agencies and administration of justice. It takes at least three years to train an expert in this field. No governmental agency in Poland is interested in applying the polygraph in the fight against increasing criminality.

References

Boring, E.G. (1943). *Psychology: US Armed Forces Manual*, E. G. Boring (Ed.)

Bratoszewski, J., Gardocki, L., Gostynski, Z., Przyjemski, S.M., Stefanski, R.A., & Zablocki, S. (1998). *Code of Criminal Procedure - A Commentary* (in Polish), ABC: Warszawa.

Gruza, E. (1999). Controversial evidences in new Criminal Procedure (in Polish), *Edukacja Prawnicza*, 6.

Bulsiewicz, A. (1998). Polygraph examinations in view of criminal procedure (in Polish) in M. Kulicki (Ed.) *Variograph Use in Criminology*, Wyzsza Szkola Policji: Szczytno.

Clore, D.L. (1998). Why are emotions always conscious?, in *The Nature of Emotion*, Ekman P & Davidson R.J. (Eds.) (Polish edition), Gdanskie Wydawnictwo Psychologiczne: Gdansk.

Daszkiewicz, W. (1974). Self-accusation and right to the silence (in Polish), *Panstwo i Prawo*, 2.

Doda, Z., & Gaberle, A. (1995). Evidences in the court procedure (in Polish), ABC: Warszawa.

Grzegorczyk, T. (1998). *Code of Criminal Procedure: A Commentary* (In Polish), Zakamycze: Kraków.

Gutenkunst, W. (1965). *Criminology - Systematic Lecture Outline*, Polish Edition, Wydawnictwa Prawnicze: Warszawa.

Hilgard, E.R. (1967). *Introduction to Psychology*, Polish edition, PWN: Warszawa.

Hofmanski, P., Sadzik, E., & Zgryzek, K. (1999). *Code of Criminal Procedure - A Commentary* (in Polish), C.H. Beck: Warszawa.

Jaworski, R. (1999). Opinion of expert polygraphic evidence as exculpatory evidence (in Polish), habilitation dissertation: Wroclaw.

Kofta, M. (1979). *Self-Control and Emotion* (in Polish), PWN: Warszawa.

Kruszynski, P. (1998). New Legal Regulations (in Polish), *New Code of Criminal Procedure*, Skretowicz, Zakamycze (Ed.)

Krzyœcin, A. (1979). Examinations of perpetrators by the Reid's technique (in Polish), *Zeszyty Naukowe ASW*, 25.

Krzyœcin, A (1997). Reliability of the psychophysiological diagnoses by means of the polygraph (in Polish), *Wojskowy Przeglad Prawniczy*, 3-4.

LeDoux, J.E. (1998). Self-control of emotions depends on type of the reaction system, in *The Nature of Emotion*, Ekman, P. and Davidson, R.J. (Eds.) (Polish edition), Gdanskie Wydawnictwo Psychologiczne: Gdansk.

Lindsay, P.H., & Norman, D. (1972). *Human Information Processing - An Introduction to Psychology*, Academic Press, Inc.: New York.

Mason, P.L. (1998). Polygraph validity with urinalysis results as ground truth. *Polygraph*, 27(1),45-48.

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

.

Poeppel, E. (1989), *Limits of Conscience* (Polish Edition), PIW: Warszawa.

Reykowski, J. (1992). *Emotional Processes, Motivation and Personality* (in Polish), PWN: Warszawa.

Sobolewski, Z. (1998). Controversy of prohibition some sources of the evidences in the new Code of Criminal Procedure (in Polish), in *New Code of Criminal Procedure*, E. Skretowicz, Zakamycze (Ed.)

Waltos, S. (1978). Polygraph examinations in Poland as seen by the law, the Supreme Court, and science, (in Polish). *Proceedings of the Conference -Application of the Polygraph Examinations*: Katowice.

Waltos, S. (1998). Code of Criminal Procedure - Between Traditional and Modern Needs (in Polish), in *New Code of Criminal Procedure*, E. Skretowicz, Zakamycze (Ed.)

Widacki, J. (1999). Glosa (in Polish), *Palestra*, 3-4.

Wojcikiewicz, J. (1999). Glosa (in Polish), *Palestra*, 3-4.

Computer Algorithm Comparison

Andrew B. Dollins, Donald J. Krapohl & Donnie W. Dutton

Abstract

This preliminary study was completed to determine if accuracy differences exist among computer programs designed to render decisions using data collected during psychophysiological detection of deception (PDD) examinations. The vendors of AXCON version 1.2, Chart Analysis version 5.1-090-17-097, Computerized Polygraph System version 2.2, Identifi version 1.43, and PolyScore version 4.0 participated in the project. The data from 97 PDD examinations (56 deceptive and 41 nondeceptive), for which examinee veracity was known only to the investigators, were sent to software vendors with the request that a decision of deception indicated, no deception indicated, or no opinion (e.g., inconclusive) be returned for each examination--using the vendors' software. The proportion of correct, erroneous, and no opinion decisions rendered by the five evaluated computer programs ranged from .71 to .77, .07 to .10, and .13 to .21, respectively. When no opinion decisions were excluded, the proportion of correct decisions ranged from .88 to .91. There were no statistically significant differences among the frequency of correct decisions, erroneous decisions, and no opinion decisions rendered by the five computer programs. There was no single examination in the data set that all of the computer programs classified as no opinion. Finally, all of the computer programs, except the Computerized Polygraph System, erroneously classified more nondeceptive examinees as deceptive (i.e., false positive) than deceptive examinees as nondeceptive (i.e., false negative).

Key words: accuracy, algorithm, Chart Analysis, comparison, computer program, CPS, Identifi, lie detection, PDD, PolyScore, software, veracity.

One of the advantages of using a computerized polygraph to collect data during a psychophysiological detection of deception (PDD) examination is that computer software can be used to evaluate the physiological data. Today's examiners have several such programs which are designed to evaluate data collected during a PDD examination to choose from. The question then becomes one of determining which, if any, computer program provides the greatest accuracy when evaluating PDD examinations.

Method

This project was completed to compare the accuracy of computer programs designed to evaluate PDD examinations. In late 1997 the physiological data from 103 PDD examinations were selected from a database of

Acknowledgments

The authors wish to express thanks to the federal, state, and county agencies that supplied the psychophysiological detection of deception examinations which were evaluated, and to the software vendors who participated by evaluating those examinations. Special thanks to: N. Joan Blackwell; Robert P. Cole; William J. Gaschler, Jr.; Linda A. Landy; Jimmie W. Swinford; Rose M. Swinford; Earl D. Taylor, and Dan V. Weatherman of the Department of Defense Polygraph Institute for their contributions to this project. This project was funded by the Department of Defense Polygraph Institute as DoDPI96-P-0006. 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.

Dr. Andrew Dollins is a research psychologist with the Research Division of the U.S. Department of Defense Polygraph Institute (DoDPI). Donald Krapohl is a federal examiner assigned to that division. Donnie Dutton is an instructor with the Instruction Division of DoDPI. Requests for reprints should be sent to Dr. Andrew Dollins, DoDPI, 7540 Pickens St., Ft. Jackson, SC 29207-5000, or by e-mail to dollinsa@Jackson-dpi.army.mil

confirmed cases maintained by the Department of Defense Polygraph Institute. All of the data were collected using an Axciton computerized polygraph system (Axciton Systems, Inc., Houston, TX). The Axciton polygraph records respiratory, cardiovascular (via an ascultatory cuff technique), and electrodermal activity. The manufacturer has not published details regarding the instrument, but basic specifications can be found in Cestaro (1997). Each examination consisted of at least three charts, as verified by a certified government examiner. The government examiner also classified the examinations as either Modified General Question Test (MGQT) or Zone Comparison Test (ZCT) format. No attempt was made to delineate among the format variations (e.g., single issue, multi-issue, and multi-facet; US Air Force MGQT vs. US Army MGQT; Backster ZCT vs. DoDPI ZCT; etc.). The examinations were classified as confirmed deceptive if the examinee signed a confession which included a statement describing the details of the crime, or if there was irrefutable corroborating evidence (e.g., positive urinalysis, fingerprints, possession of incident-related paraphernalia, etc.) linking the examinee to the crime under investigation. Examinations were classified as confirmed nondeceptive if the examinee was cleared of involvement in the crime by confession of another person.

The data were sent to vendors of computer programs designed to evaluate PDD examinations. The vendors were asked to evaluate the examinations and return decisions of deception indicated (DI), no deception indicated (NDI), or no opinion (NO). None of the vendors had, to the best of our knowledge, seen the 103 examinations before this project and they were not told examinees' veracity until after all vendors had responded. The computer programs included AXCON version 1.2 and Chart Analysis version 5.1-090-17-097 by Axciton Systems, Inc. (Houston, TX); the Computerized Polygraph System (CPS) version 2.2 by the Stoelting Company (Wood Dale, IL); Identifi version 1.43 by Identifi (Olympia, WA); and, PolyScore version 4.0 by the Johns Hopkins University Applied Physics Laboratory (Laurel, MD). The Axciton AXCON program was under development by Axciton Systems, Inc. and was not available to the public when testing was done. The computer programs named Chart Analysis, AXCON, CPS, Identifi, and PolyScore are, to the best of our knowledge, the only software currently available to evaluate and render decisions using data from specific issue PDD examinations.

All of the computer programs tested were able to read the proprietary Axciton data format except the CPS system. The data sent for CPS evaluation were converted to text using a program provided by Axciton Systems, Inc. The CPS analysis differed in that (a) the conversion program rounded question onset marks to the nearest full second (according to Bell, Raskin, Honts, & Kircher, 1999; CPS only scores electrodermal reactions if they begin at least 0.5 seconds after question onset) and (b) the CPS program was developed using a true measure of skin conductance (the Axciton system measures a hybrid of skin conductance and skin resistance). The influence of these differences is not known.

When reviewing the cases, in preparation for reporting results, we observed that the question content of some cases did not correspond to notations in our master database. We thus attempted to reconfirm each of the 103 examinations with the original testing agency. We were unable to confirm three cases because the testing agency had moved and the case files were lost. Three additional cases were discarded due to poor data quality as determined by a panel of DoDPI PDD examiners. The computer programs rendered either correct or NO decisions on each of these cases, so discarding them decreased, rather than increased, the accuracy of decisions reported here.

Of the 97 examinations analyzed, 64 (28 deceptive and 36 nondeceptive) were collected by federal agencies (e.g., US Army Criminal Investigations Division and US Postal Service). The remaining 33 examinations (28 deceptive and 5 nondeceptive) were collected by nonfederal agencies (e.g., Birmingham Police Department, Birmingham, AL; Clayton County Sheriff's Office, Jonesboro, GA; Marion County Sheriff's Office, Ocala, FL, Mobile County Sheriff's Office, Mobile, AL; South Carolina Law Enforcement Division, Columbia, SC).

The reported sample was composed of 44 MGQT examinations (27 deceptive and 17 nondeceptive) and 53 ZCT examinations (29 deceptive and 24 nondeceptive). Eighty-five examinations were of suspects under investigation, eight were of witnesses, and four were of victims. Twenty eight of the examinees were female and 69 were male. Of the 70 examinees for which race was available, 25 were African American, 39 were Caucasian, four were Hispanic, one was a Native American, and one was of South Pacific heritage. The examinee age range, for the 94 examinees for which age data were available, was 14 to 70 years with a mean of 32.6 (SD = 10.83).

Results

The frequency of decisions rendered by each computer program is presented in Table 1. Results of analyses using Cochran's Q statistic (Siegel & Castellan, 1988), which is analogous to a Repeated Measures Analysis of Variance for nonparametric data, indicated that there were no statistically significant differences among the proportion of correct ($Q[4] = 3.243$, $p = .518$), incorrect ($Q[4] = 2.061$, $p = .724$), and NO ($Q[4] = 3.804$, $p = .433$) decisions rendered by the computer programs. The frequency data are presented as proportions in Table 2. The Cochran's Q analyses were only calculated with NO decisions included because it is not, to the best of our knowledge, possible to calculate a repeated measures analysis unless all groups have an equal number of observations. The proportion of correctly identified deceptive examinees was between .732 and .893 (N=56); if NO decisions were excluded the range was .911 to .980. The proportion of correctly identified nondeceptive examinees was between .537 and .683 (N=41); if NO decisions were excluded, the range was .727 to .903. The overall proportion of correct decisions regarding examinee veracity was between .711 and .773 (N=97); if NO decisions were excluded, the range was .881 to .908. The proportion of NO decisions was between .134 and .216 of the 97 decisions rendered. While the number of observations are too few for meaningful analysis, breakout decision frequency tables for test format and testing organization are provided in Appendix A to assist readers in understanding the data. The raw data are provided in Appendix B.

Table 1
Frequency of Decisions by Subject Veracity

	Deceptive (n = 56)			Nondeceptive (n = 41)		
Computer Program	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
AXCON	50	1	5	24	9	8
Chart Analysis	49	2	5	22	8	11
CPS	41	4	11	28	3	10
Identifi	49	1	6	22	8	11
PolyScore	49	1	6	26	7	8

Table 2
Proportions of Correct Decisions (with SEMs)

Computer Program	NO Decisions Included			NO Decisions Excluded			Proportion NO Decisions
	DI	NDI	TOTAL	DI	NDI	TOTAL	
AXCON	.893	.585	.763	.980	.727	.881	.134
	(.041)	(.077)	(.043)	(.020)	(.078)	(.056)	(.035)
Chart Analysis	.875	.537	.732	.961	.733	.877	.165
	(.044)	(.078)	(.045)	(.027)	(.081)	(.060)	(.038)
CPS	.732	.683	.711	.911	.903	.908	.216
	(.059)	(.073)	(.046)	(.041)	(.053)	(.052)	(.042)
Identifi	.875	.537	.732	.980	.733	.888	.175
	(.044)	(.078)	(.045)	(.020)	(.081)	(.058)	(.039)
PolyScore	.875	.634	.773	.980	.788	.904	.144
	(.044)	(.075)	(.043)	(.020)	(.071)	(.051)	(.036)

Note: NO = no opinion.

A test for the significance of proportion differences (Bruning & Kintz, 1987) indicated that the proportion of correctly identified deceptive examinees was significantly greater than the proportion of correctly identified nondeceptive examinees ($p < .005$) for the AXCON, Chart Analysis, Identifi, and PolyScore computer programs. This was true when NO decisions were included and excluded. There were no significant differences between the proportion of deceptive and nondeceptive examinees correctly identified by the CPS computer program.

The proportion of agreement between pairs of computer programs are presented in Table 3. Two evaluation systems were considered to be in agreement if both classified the same examination as DI, NDI, or NO. The proportion of decision agreement between evaluation systems ranged from .722 to .907. Most of the disagreements between pairs of evaluation systems were combinations including NO decisions (i.e., DI and NO or NDI and NO).

Table 3
Proportion of Agreement Between Pairs of Scoring Systems (n = 97)

Computer Program	Chart Analysis	CPS	Identifi	PolyScore
AXCON	.907	.753	.804	.856
Chart Analysis	.742	.784	.804	
CPS	.722	.753		
Identifi	.722			

Table 4
Frequency of Disagreement Between Pairs of Scoring Systems and Between Pairs of Scoring Systems and Ground Truth (n = 97)

Computer Program	Chart Analysis	CPS	Identifi	PolyScore	Ground Truth
AXCON	0	0	1	1	10
Chart Analysis		0	0	1	10
CPS			1	1	7
Identifi				0	9
PolyScore					8

Note. A disagreement is defined as one scoring system classifying an examination as DI while another classifies the same examination as NDI.

The decisions made by the five computer programs are summarized as follows. All five computer programs agreed and correctly classified the veracity of 52 examinees (i.e., 36 deceptive and 16 nondeceptive). All programs agreed and incorrectly classified three nondeceptive examinees as deceptive. Of the remaining 42 examinations: 16 deceptive examinees were classified as DI or NO, 3 deceptive examinees were classified as NDI or NO, 16 nondeceptive examinees were classified as NDI or NO, and three nondeceptive examinees were classified as DI or NO. Four examinees (i.e., 1 deceptive and 3 nondeceptive) received opposite classifications by the computer programs. That is, one or more computer programs classified the examinee as DI while other computer programs classified the same examinee as NDI. The disagreement frequencies are shown in Table 4.

Discussion

Three important conclusions may be drawn from this preliminary study. First, as far as we were able to determine from this sample, there are no statistically significant accuracy differences among the five computer programs evaluated. A second, less apparent, conclusion is that there was no single examination in the data set that all of the computer programs classified as NO. Finally, as may be seen in Tables 1 and 2, all of the computer programs, except CPS, erroneously classified more nondeceptive examinees as deceptive (i.e., false positive) than deceptive examinees as nondeceptive (i.e., false negative).

The most obvious flaw with this report is bias due to sampling error. The data were those sent to the Institute between July and October 1997. Although a few cases were erroneously decided by the original examiner, the majority of the cases were correctly decided-which could have biased the sample. Contributors to the database may have been reluctant to send cases they had missed. All of the cases were confirmed via confession of the examinee or another-which could have further biased the sample because the number of unconfirmed and uncomfirmable examinations is not known. No attempt was made to counterbalance, randomly assign, or otherwise control for bias due to examiner ability, originating agency rules, test format, examinee status (e.g. suspect, witness, victim), examinee or examiner gender, examinee or examiner race, or examinee age. In fact, the only things which can be said about this sample with any degree of certainty are that the sample is composed of actual field data and examinee veracity was not available to the vendors.

There were no experimental controls regarding the software used in testing. Vendors were allowed to test the data and

report results to the best of their ability. Vendors could have performed expert editing, screening, or other manipulations that might not be available to the field examiner. As previously mentioned, all of the data were collected using an Axciton computerized polygraph. If a laboratory grade instrument had been used to collect the test data, the results may have been quite different. (A laboratory grade instrument would allow absolute, rather than relative, measurement of reactions. A laboratory grade instrument would also measure true skin conductance, rather than a hybrid of conductance and resistance.) Differences between the computer programs used in this test were not addressed. Assessments regarding operator training and skill using each computer program would be appropriate topics for future evaluations. In addition, we were unable to control for or assess the accuracy of features or criteria used by the computer programs. Only the Stoelting Company has published the decision criteria used by their software (Kircher & Raskin, 1988). Similar publications by other vendors would allow us to assess the predictive ability of features used by the different computer programs.

The implication of the first conclusion, that there are no statistically significant differences in the tested computer programs' ability to predict examinee veracity, is self explanatory. The differences in decision accuracy observed in this sample could be due to sampling error-and not to the ability of one or more computer programs to predict examinee veracity. The implications of the second conclusion, that all of the computer programs did not agree on a single NO decision are not so self evident. It is sometimes difficult to understand the significance of something that doesn't exist. There were cases where all of the scoring systems made correct and incorrect decisions. There were a few cases where the scoring systems made diametrically opposed decisions. There was, however, no single case that all scoring systems labeled unscorable. It is difficult to justify the continued supposition that a NO decision should be omitted or labeled as "correct" when there is so little agreement regarding when a NO decision should be made. Perhaps the PDD discipline should expend some effort to clearly define the parameters which make a NO decision appropriate.

The third conclusion drawn from this data, that all scoring systems, except the CPS system, have a bias toward predicting that subjects are deceptive was unexpected. The bias could have developed because the algorithm training sets were similar for all except the CPS system. (Chart Analysis and AXCON, Identifi, and PolyScore were developed using at least some confirmed case data supplied by the Department of Defense Polygraph Institute. All of the confirmed case data was collected using an Axciton computerized polygraph. The CPS system was developed using data collected with a Stoelting CPS by the US Secret Service.) The bias could also be an artifact of the data set used in this project. Vendors should be aware of this potential bias-which should be examined in future studies.

In summary, this preliminary study may be flawed due to the failure to manipulate, randomize, or otherwise control for (a) the data sample, (b) the data quality, and (c) operator skill. The field samples do, however, provide an index of how software designed to evaluate PDD examinations and predict subject veracity will perform. The sample size (n = 97) provided the statistical power to detect a 10% difference between scoring systems, had such a difference actually existed, with a probability of .80. We therefore present the conclusion that there are no statistically significant differences in the tested computer programs' ability to correctly predict deception. We suggest that efforts be made to clearly define the parameters necessary for a decision of NO and that future studies assess the possibility that a false positive bias exists among computer programs designed to evaluate PDD examinations.

References

Bell, B. G., Raskin, D. C., Honts, C. R., & Kircher, J. C. (1999). The Utah numerical scoring system. *Polygraph*, 28, 1-9.

Bruning, J. L., & Kintz, B. L. (1987). *Computational handbook of statistics* (3rd. ed.). Glenview, IL: HarperCollins.

Cestaro, V. L. (1998). Instrumentation for presenting a known standard signal to the electrodermal activity channel for assessing response characteristics. *Polygraph*, 27(3), 188-209.

Kircher, J. D., & Raskin, D. C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology*, 73, 291-302.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd. ed.). New York, NY: McGraw-Hill.

Appendix A

Subgroup Decision Frequencies

Decision Frequencies for Federal and Nonfederal Examinations

Computer Program	Deceptive			Nondeceptive		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
Federal Examinations						
AXCON	25	1	2	23	7	6
Chart Analysis	25	1	2	21	6	9
CPS	20	2	6	26	2	8
Identifi	25	0	3	21	6	9
PolyScore	26	1	1	25	4	7
Nonfederal Examinations						
AXCON	25	0	3	1	2	2
Chart Analysis	24	1	3	1	2	2
CPS	21	2	5	2	1	2
Identifi	24	1	3	1	2	2
PolyScore	23	0	5	1	3	1

Decision Frequencies for Zone Comparison Test and Modified General Question Test Examination Formats

Computer Program	Deceptive			Nondeceptive		
	Correct	Incorrect	No Opinion	Correct	Incorrect	No Opinion
Modified General Question Test						
AXCON	25	1	1	8	5	4
Chart Analysis	24	1	2	8	4	5
CPS	21	2	4	9	1	7
Identifi	24	0	3	9	5	3
PolyScore	25	1	1	9	2	6
Zone Comparison Test						
AXCON	25	0	4	16	4	4
Chart Analysis	25	1	3	14	4	6
CPS	20	2	7	19	2	3
Identifi	25	1	3	13	3	8
PolyScore	24	0	5	17	5	2

Appendix B

Raw Data
(A label key is located at the end of the table)

S#	Trth	Axcon	ChtAn	CPS	Idntf	PlySc	Format	Agncy	Sex	Stus	Race
1	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
2	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Cauc
3	DI	DI	DI	NO	DI	DI	ZCT	Fed	F	Spt	Afri
4	NDI	NO	NO	NDI	NDI	NDI	MGQT	Fed	F	Spt	Cauc
5	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	Hisp
6	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
7	NDI	NO	NDI	NDI	NDI	NO	MGQT	Fed	M	Spt	
8	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
9	NDI	NDI	NDI	NO	NO	NDI	MGQT	Fed	M	Spt	
10	NDI	DI	DI	NO	DI	DI	MGQT	Fed	M	Spt	
11	NDI	DI	DI	DI	DI	DI	ZCT	Fed	M	Spt	
12	NDI	DI	DI	NO	DI	DI	ZCT	Fed	M	Spt	
13	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	M	Spt	
14	NDI	NDI	NDI	NDI	NO	NDI	ZCT	Fed	M	Spt	
15	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
16	NDI	NDI	NO	NDI	NO	NDI	ZCT	Fed	M	Spt	
17	NDI	NDI	NO	NDI	NO	NDI	ZCT	Fed	M	Spt	
18	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
19	NDI	NO	NO	NDI	NDI	NO	ZCT	Fed	M	Spt	
20	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Hisp
21	NDI	DI	NO	NO	DI	NO	MGQT	Fed	M	Spt	
22	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
23	NDI	NO	NO	NO	NO	DI	ZCT	NoFed	M	Spt	Cauc
24	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	M	Spt	
25	DI	DI	DI	NO	DI	DI	MGQT	Fed	F	Spt	Afri
26	NDI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	
27	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
28	DI	DI	DI	DI	DI	DI	MGQT	NoFed	M	Spt	Cauc
29	DI	DI	DI	DI	DI	DI	MGQT	NoFed	F	Spt	Afri
30	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
31	DI	DI	DI	DI	DI	DI	MGQT	NoFed	M	Spt	Afri
32	NDI	NDI	NO	NDI	NDI	NDI	MGQT	Fed	M	Spt	
33	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
34	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	M	Spt	
35	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
36	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Afri
37	NDI	NO	NO	NDI	NO	DI	ZCT	NoFed	F	Spt	Cauc
38	DI	DI	DI	NO	DI	DI	MGQT	NoFed	M	Spt	Cauc
39	DI	DI	DI	DI	DI	DI	MGQT	NoFed	M	Spt	Afri
40	NDI	NO	NO	NDI	NO	NDI	ZCT	Fed	M	Spt	
41	DI	DI	NO	DI	DI	DI	MGQT	NoFed	M	Spt	Afri
42	DI	DI	DI	DI	DI	DI	ZCT	Fed	F	Spt	Cauc
43	NDI	NDI	NDI	NDI	NDI	NO	ZCT	Fed	M	Spt	
44	DI	DI	DI	DI	DI	DI	ZCT	Fed	F	Spt	Afri
45	DI	DI	DI	NO	DI	DI	MGQT	Fed	F	Wtn	Afri
46	DI	NO	NO	NDI	DI	NO	MGQT	Fed	F	Spt	Cauc
47	DI	DI	DI	DI	DI	DI	MGQT	Fed	F	Spt	Cauc
48	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	F	Spt	Cauc

S#	Trth	Axcon	ChtAn	CPS	Idntf	PlySc	Format	Agncy	Sex	Stus	Race
49	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Hisp
50	NDI	NO	DI	NO	NO	NO	MGQT	Fed	F	Spt	Cauc
51	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Afri
52	DI	DI	DI	DI	DI	DI	MGQT	Fed	F	Wtn	Cauc
53	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Wtn	
54	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
55	DI	DI	DI	NO	NO	DI	MGQT	Fed	M	Wtn	Cauc
56	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	F	Spt	Afri
57	DI	DI	DI	DI	DI	DI	MGQT	Fed	F	Spt	Cauc
58	DI	DI	DI	NO	DI	DI	ZCT	Fed	F	Spt	Cauc
59	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Spt	
60	DI	NDI	NDI	NDI	NO	NDI	MGQT	Fed	M	Spt	Hisp
61	NDI	NDI	NDI	NDI	NO	NDI	ZCT	Fed	M	Spt	
62	DI	DI	DI	DI	DI	DI	ZCT	Fed	M	Spt	Cauc
63	DI	DI	DI	DI	DI	DI	ZCT	NoFed	F	Spt	Cauc
64	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Afri
65	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
66	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	NoFed	M	Vtm	Afri
67	DI	DI	DI	NO	DI	NO	ZCT	NoFed	M	Spt	Cauc
68	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Afri
69	DI	DI	DI	DI	DI	NO	ZCT	NoFed	F	Vtm	Cauc
70	DI	DI	DI	NO	DI	DI	ZCT	NoFed	M	Spt	Afri
71	DI	DI	DI	NO	DI	DI	ZCT	NoFed	M	Spt	Cauc
72	DI	DI	DI	DI	DI	DI	MGQT	NoFed	M	Spt	Cauc
73	DI	DI	DI	DI	NO	DI	ZCT	NoFed	F	Spt	Cauc
74	NDI	DI	DI	NO	DI	NO	MGQT	NoFed	M	Spt	Afri
75	DI	NO	NDI	NDI	NO	NO	ZCT	NoFed	M	Spt	Afri
76	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
77	NDI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Cauc
78	DI	DI	DI	DI	DI	DI	MGQT	NoFed	M	Wtn	Cauc
79	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Afri
80	DI	NO	NO	NDI	NDI	NO	ZCT	NoFed	M	Spt	Afri
81	DI	NO	NO	NO	NO	DI	ZCT	NoFed	F	Spt	Afri
82	DI	DI	DI	DI	DI	NO	ZCT	NoFed	M	Spt	Afri
83	DI	DI	DI	DI	DI	DI	ZCT	Fed	M	Spt	Cauc
84	NDI	DI	DI	NO	NO	NDI	ZCT	Fed	F	Vtm	Cauc
85	DI	DI	DI	DI	DI	DI	ZCT	Fed	F	Spt	Cauc
86	DI	DI	DI	DI	NO	DI	MGQT	Fed	M	Spt	Asia
87	NDI	NDI	NDI	NDI	NO	NDI	MGQT	Fed	F	Wtn	Afri
88	NDI	NO	NO	NO	DI	NO	MGQT	Fed	F	Wtn	Afri
89	DI	DI	DI	DI	DI	DI	ZCT	Fed	M	Spt	Afri
90	NDI	NDI	NDI	NDI	NDI	NDI	ZCT	Fed	M	Wtn	Cauc
91	DI	NO	NO	NO	DI	DI	ZCT	Fed	F	Spt	Cauc
92	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Cauc
93	NDI	DI	NO	NO	NDI	NO	MGQT	Fed	M	Vtm	Cauc
94	DI	DI	DI	DI	DI	DI	ZCT	Fed	F	Spt	Cauc
95	DI	DI	DI	DI	DI	DI	MGQT	Fed	M	Spt	Afri
96	DI	DI	DI	DI	DI	DI	MGQT	Fed	F	Spt	Nati
97	DI	DI	DI	DI	DI	DI	MGQT	NoFed	F	Spt	Cauc

Omitted Cases

S#	Trth	Axcon	ChtAn	CPS	Idntf	PlySc	Format	Agncy	Sex	Stus	Race
1	DI	DI	DI	NO	DI	DI	MGQT	NoFed	F	Spt	Cauc
2	DI	DI	DI	DI	DI	DI	ZCT	NoFed	M	Spt	Afri
3	NDI	NDI	NDI	NDI	NDI	NDI	MGQT	Fed	M	Wtn	Cauc
4	NDI	NDI	NO	NDI	NDI	NDI	MGQT	Fed	M	Wtn	Cauc
5	NDI	NDI	NO	NDI	NDI	NDI	MGQT	Fed	M	Spt	Hisp
6	NDI	NDI	NO	NDI	NDI	NDI	MGQT	NoFed	M	Spt	Cauc

(Appendix B Label Key)

S# - Subject number
Trth - Ground truth
Axcon - AXCON ver 1.2 (Axciton)
ChtAn - Chart Analysis 5.1 (Axciton)
CPS - Computer Polygraph System 2.2 (Stoelting)
Idntf - Identifi 1.43 (Identifi)
PlySc - PolyScore 4.0 (John Hopkins Univ. Applied Physics Laboratory)
Format
ZCT - Zone Comparison Test
MGQT - Modified General Question Test
Agncy - Agency
Fed - Federal
NoFed - Nonfederal
Sex
M - Male
F - Female
Stus - Status
Spt - Suspect
Wtn - Witness
Vtm - Victim
Race
Afri - African American
Cauc - Caucasian
Hisp - Hispanic
Asia - South Pacific heritage

* * * * * *

Erratum

In the last edition of *Polygraph*, Table 3 (page 240) of the Dollins, Krapohl and Dutton article was formatted incorrectly. The correct information is provided below. We regret the error.

Table 3
Proportion of Agreement Between Pairs of Scoring Systems (n = 97)

Computer Program	Chart Analysis	CPS	Identifi	PolyScore
AXCON	.907	.753	.804	.856
Chart Analysis		.742	.784	.804
CPS			.722	.753
Identifi				.722

Identifi Comments on Dollins et al's Computer Algorithm Comparison

W. Keith Hedges

We were extremely pleased that the authors elected to conduct a study comparing algorithm accuracy. The algorithms were tested on a sample composed of different examination types i.e., Department of Defense Polygraph Institute Zone Comparison Test (ZCT), Backster ZCT, Modified General Question Test (MGQT) and the US Air Force MGQT. While the accuracy demonstrated by the algorithms on a sample of this nature demonstrates their robustness, it did not demonstrate their full potential. As the Identifi developer, I do not want to appear overly critical of the study and at the same time, I do have a strong desire to protect that which I have worked so hard to develop. With that in mind, I want to bring up points that in a large part were addressed by the authors in their summary but require additional emphasis.

The algorithms available to the examiners at the time the sample data was collected could have had a bearing on an individual algorithm's accuracy. If the examiner conducting an examination was using algorithms A and B but not C and at least partially based decisions upon A and B results, the sample could be severely biased. Future studies should include this information and examiners should either have all or none of the algorithms available at the time data is collected.

The comparison question technique variations used in the Dollins et al paper could have had a bearing on the accuracy demonstrated by individual algorithms. Identifi version 1.43 has a history of performing at optimum scoring specific-issue single-incident examinations composed of three relevant and three comparison questions per chart (DoDPI ZCT) and was less accurate analyzing examinations composed of two relevant and two or three comparison questions per chart. Version 1.43 was also less efficient when scoring the US Army MGQT. Future study samples composed of a single examination variation could determine which if any of the algorithms show a distinct advantage over the others for a given format variation.

The Identifi No Opinion / Inconclusive rate was about 14% higher than we normally experience. We would like to see future projects include quality control to detect violations that would reject inferior examinations from the study. The quality of the overall examination and data provided to an algorithm for analysis may have a direct influence on accuracy. It is suggested that a panel classify each examination by data and examination quality to determine if the algorithms perform differently depending upon perceived data and examination quality. This may not be practical, as a substantial increase in sample size would probably be required to obtain meaningful information.

Blind scoring of the examinations by at least three PDD examiners could compare examiner and algorithm accuracy. This would provide extremely useful information to the PDD profession. As a field-examiner, it would be useful to know if examiners are more or less accurate than the algorithm I am using.

Address correspondence to: Keith Hedges, P.O. Box 11422, Olympia, WA 98508

Now that we are finished crying and picking out potential study flaws, we would like to thank the authors very deeply for including Identifi version 1.43 in their study. Algorithms are in their infancy and are constantly improving. Perhaps the only way they will become significantly better will be the results of studies like this one. Competition between the algorithm developers should result in increased accuracy. We look forward to testing version 3.0 in future algorithm comparisons.

CPS Comments on Dollins et al's Computer Algorithm Comparison

John C. Kircher & David C. Raskin

We appreciate the opportunity to comment on the paper by Dollins, Krapohl, and Dutton (2000), and wish only to elaborate on several points they already made.

Convenience Sampling and Validity

The authors noted that the study does not provide unbiased estimates of the validity of any of the tested computer algorithms. Although there was little doubt that subjects coded as deceptive and nondeceptive were in fact deceptive and truthful, cases were voluntarily submitted to the DoDPI by federal and nonfederal agencies. The criteria used by those agencies to select cases for submission is unknown. The authors noted that "very few cases were erroneously decided by the original examiner," which suggests that some agencies or examiners may have systematically selected cases to present themselves in a positive light. Although it is unlikely, it is still possible that the overall accuracy of decisions by the examiners who provided these cases was only 50% (chance), and wittingly or unwittingly, only those cases where the examiner happened to make a correct decision were supplied to DoDPI. Consequently, the accuracy rates for the various algorithms should not be viewed as indicative of their validity in the field.

Relative Validity of Scoring Algorithms

The purpose of the Dollins et al. study was not to assess the accuracy of various computer algorithms in absolute terms. Rather, the expressed purpose of the study was to assess the relative accuracy of computer algorithms, but even that analysis was problematic. Statistical theory predicts that the accuracy of a scoring algorithm will drop to the extent that the database used to develop the algorithm differs from the database that is used to test the algorithm. CPS was uniquely disadvantaged in these comparisons because the characteristics of the databases we used to develop CPS differed from those used by the other vendors in terms of case selection, instrumentation, and event timing information.

Case Selection. CPS was developed from data collected in laboratory experiments and confirmed criminal cases. The field cases were obtained from an exhaustive search of all polygraph tests conducted by the U. S. Secret Service over a period of three years (Raskin, Kircher, Honts, & Horowitz, 1988). Every case confirmed by confession and inconvertible physical evidence was included in the field sample. We did not omit cases because we felt that the quality of the recordings was inadequate, the outcome was unclear, or because the decision by the original examiner was incorrect.

We know of no reports that describe the criteria used by the developers of AXCON, Chart Analysis, and Idenifi to select field cases for developing their scoring algorithms. We assume that they used the same database as was used to develop PolyScore, and problems with that database are described in detail elsewhere (Porges, Johnson, Kircher, & Stern, 1996). As Dollins et al. noted, their sample was biased in part because cases were voluntarily submitted by field examiners. The same sampling bias occurred when the original DoDPI database was compiled. This type of sampling bias was not present in our field database because we had access to every polygraph test administered by the Secret Service examiners, and the examiners did not decide which of their tests we should use. If all algorithms except CPS were originally developed and were subsequently tested using similarly biased sampling procedures, then CPS was uniquely disadvantaged in the comparisons reported by Dollins et al.

In addition to using cases voluntarily submitted by field examiners, the original database on which PolyScore was developed

included some unspecified number of cases in which three examiners independently agreed that the tested individual was truthful or deceptive (Olsen, Harris, Capps, & Ansley, 1997). This one change introduced yet another source of bias in the original database. Wisely, Dollins et al. did not use agreements by polygraph examiners as a criterion for the selection of cases for their study. If one considers that the estimated decision accuracy of PolyScore dropped from nearly 100% (Olsen & Harris, 1994; Olsen et al., 1997) to 79% on confirmed truthful subjects with this one change in case selection criteria, it is clear that sampling bias can have substantial effects on accuracy rates.

Instrumentation. All vendors except us developed their algorithms using charts collected with an Axciton polygraph. Whereas CPS records skin conductance, it is unclear what type of electrodermal activity is recorded by the Axciton. The manufacturer reports that the Axciton records skin conductance. Dollins et al. stated that the electrodermal signals produced by the Axciton are a hybrid of skin conductance and skin resistance. Cestaro (1998), also a research scientist at DoDPI, found that the signals generated by the Axciton did not accurately reproduce known changes in conductance or resistance. Not only were the signals output by the system inaccurate, they were not even monotonically related to the inputs. Our examination of the Axciton charts revealed that the electrodermal signal often showed a small increase followed by a precipitous drop that continued well below the level at response onset. Such changes are not characteristic of skin conductance or resistance responses. CPS expects properly recorded skin conductance, which was poorly represented by the signals output by the Axciton. To make matters worse, CPS weighs increases in skin conductance more heavily than changes in the cardiograph or respiration recordings when it makes its decisions (Kircher & Raskin, 1988). Again, CPS was uniquely disadvantaged because it, unlike the other algorithms, had never been given the opportunity to learn how the signals produced by the Axciton relate to truth and deception.

Event Timing. All algorithms except CPS read the original Axciton data files. In order for us to participate in the Dollins et al. study, the developer of the Axciton wrote a program to convert native Axciton charts to text files that could be read by CPS. Unfortunately, the event marks that indicated question onsets were not accurately reported by this program. Dollins et al. stated that the program rounded the question onset marks to the nearest second. However, in some cases it appeared that the event marks occurred 1 to 2 seconds after the actual presentation of the question. The accuracy of CPS results suffered to the extent that indications of question onset were incorrect. The inaccuracies in event timing introduced small errors in the measurement of all respiration responses, small to moderate-sized errors in some cardiograph measurements, and large errors in some electrodermal measurements. Following conventional scientific practice, CPS automatically rejects electrodermal responses that begin prior to question onset. The program assumes that a response that begins before the question is asked, or within 500 ms of question onset, is not a response to the question. CPS undoubtedly rejected some bona fide electrodermal responses to test questions because the event marks were misplaced and occurred in the middle of the subject's response. Again, this was not a problem for the other algorithms because the event marks were properly recorded in the original Axciton charts.

Decision Accuracy and Inconclusive Outcomes

The unweighted mean accuracy of decisions for the five algorithms ranged from 85% to 91%. Despite several factors that put CPS at a disadvantage, there were no significant differences among the five algorithms in overall decision accuracy. However, only CPS was unbiased. The authors found that the decisions by all algorithms except CPS were significantly less accurate for truthful subjects than for deceptive subjects. Although CPS produced 7% fewer correct decisions on deceptive subjects (91%) than did the other programs (M = 98%), decisions by CPS on truthful subjects (90%) were 15% more accurate (M = 75%).

The mean decision accuracy of CPS was slightly higher than that of the other

algorithms, but it also produced the greatest number of inconclusive outcomes. We recently reviewed the results of seven other laboratory and field studies of CPS and found that an inconclusive rate as high as 22% after three charts is not atypical for CPS (Kircher & Raskin, in press). Had more charts been available, we would expect the inconclusive rate to drop to about 10%. Nevertheless, all of the programs were tested with the same number of charts, and CPS had the greatest number of inconclusive outcomes. The problems associated with the measurement of electrodermal responses could account for the relatively high inconclusive rate. As noted above, CPS was designed to measure increases in skin conductance, it relies heavily on these measurements when it makes a decision, and the signals generated by the Axciton are only remotely related to actual skin conductance. CPS also expected the event marks to correspond to the onset of question presentation, and that was not the case. Errors in the placement of question onsets adversely affected measurements of electrodermal, respiration, and cardiovascular responses. The problems associated with the use of the Axciton and the recording of question onsets were unique to CPS. Since these problems occurred for truthful and deceptive subjects and for comparison and relevant questions, we would expect a general reduction in the percentage of correct classifications by CPS and a relatively high rate of inconclusive outcomes.

That CPS proved to be reasonably accurate is a testimonial to the robustness of the CPS algorithm and its ability to overcome the serious flaws inherent in the Axciton data. We invite a proper scientific comparison of the CPS to the other algorithms that would apply each of them to a set of data generated with scientifically acceptable recording techniques.

References

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.

Kircher, J. C. & Raskin, D. C. (in press). Computer methods for the psychophysiological detection of deception. To appear in a book on the polygraph, M. Kleiner (Ed.), New York: Academic Press.

Olsen, D. E., Harris, J, C. & Chiu, W. W. (1994). The development of physiological detection of deception scoring algorithm. *Psychophysiology*, 31, S11.

Olsen, D. E., Harris, J. C., Capps, M. H., & Ansley, N. (1997). Computerized polygraph scoring system. *Journal of Forensic Sciences*, 42, 61-70.

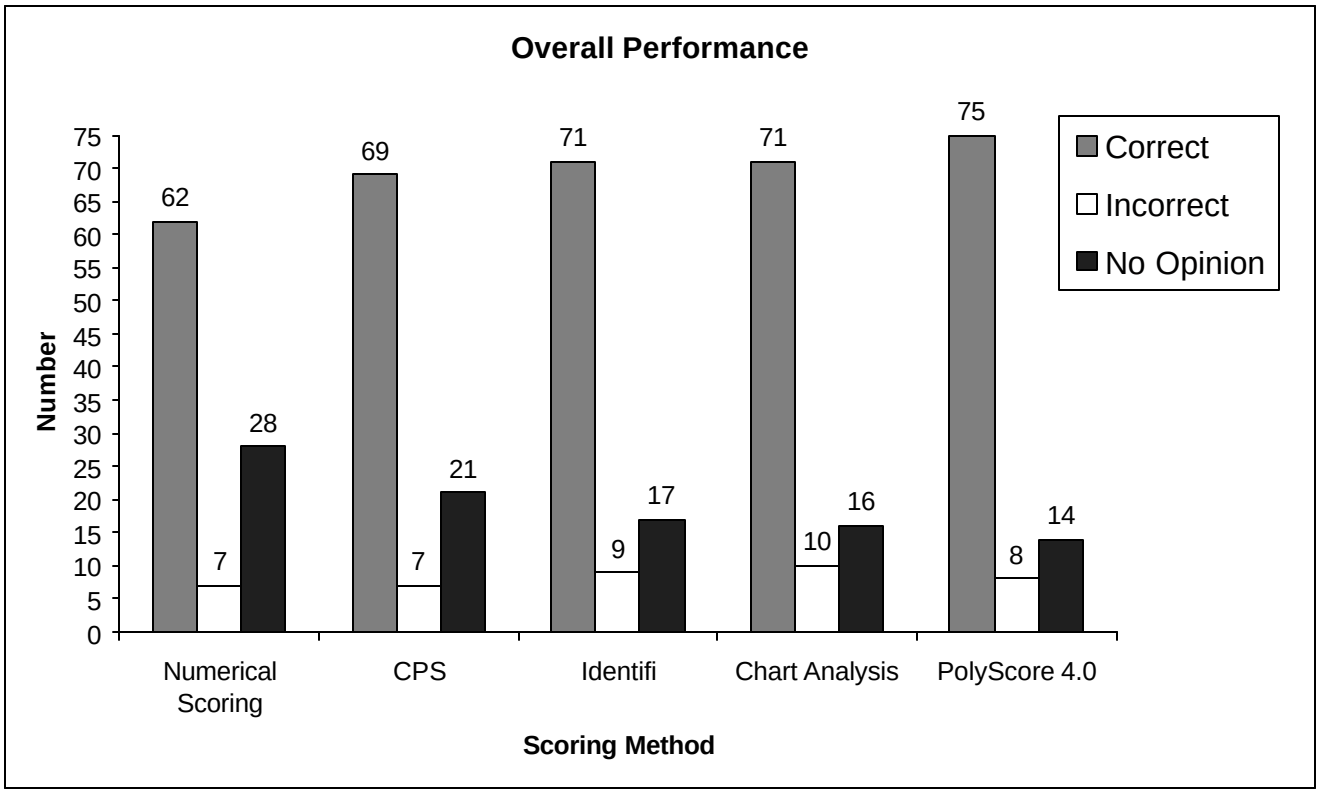
Porges, S. W., Johnson, R., Kircher, J. C., & Stern, J. A. (1996). Report of peer review of Johns Hopkins University/Applied Physics Laboratory. Report to the Central Intelligence Agency. Salt Lake City: University of Utah, Department of Educational Psychology.

Raskin, D. C., Kircher, J. C., Honts, C. R., & Horowitz, S. W. (1988). A study of the validity of polygraph examinations in criminal investigation. Final report to the National Institute of Justice (Grant No. 85-IJ-CX-0040). Salt Lake City: University of Utah, Department of Psychology.

JHU/APL Comments on Dollins et al's Computer Algorithm Comparison

John C. Harris

Over the last several years, The Johns Hopkins University Applied Physics Laboratory has developed and fielded what we firmly believe to be the most accurate polygraph evaluation methodology available in the form of PolyScore and there is nothing in the Dollins et al results that contradicts this. The small sample size of 97 examinations used in the study limits the conclusions that can be drawn, but the data are somewhat more revealing than they may at first seem. As shown in the figure below, one algorithm provided the most correct decisions and the fewest no opinions in the Dollins study--PolyScore. The AXCON algorithm is not included in our comments because it is our understanding that this is a consensus algorithm that makes use of the results of other algorithms including PolyScore. To what degree PolyScore influenced its decisions is unknown, but in any case it is not an independent algorithm.



The Johns Hopkins University, Applied Physics Laboratory
Contact John C. Harris at 240-228-4975, 240-228-6519 (fax) or via e-mail: John.Harris@jhuapl.edu

The fact that PolyScore performed the best on these data is self-evident. The issue is whether this superiority is limited to just these data or applies more generally. Here we are limited by the sample size by what can be inferred. Nonetheless, using a more powerful statistical technique than that used by the algorithm comparison authors, we can reject that all of the computerized algorithms perform the same ($p<0.0001$). In particular, consistent with the above figure, a pairwise comparison of the algorithms reveals that the performance of the CPS algorithm is different from the others. These results are based on Friedman's test, which is a generalization of the Cochran's Q test used by the authors of the algorithm comparison. Friedman's test allows the simultaneous comparison of the algorithms' performances in terms of both accuracy and no opinions. To do this it relies on the intuitively obvious assumption that a no opinion result lies between one that is correct and one that is incorrect, say in terms of either probability of deception or in points.

Comparing the algorithms to perfection is important for validation purposes, but how well they perform against Numerical Scoring is the measure of their field utility. Therefore, we evaluated the Dollins data using the Department of Defense Polygraph Institute scoring criteria and decision rules for Numerical Scoring. The Numerical Scoring results are shown in the preceding figure and in the following table along with those for PolyScore.

	Deceptive			Truthful		
	Correct	**Incorrect**	**No Opinion**	**Correct**	**Incorrect**	**No Opinion**
Numerical Scoring	45	0	11	17	7	17
PolyScore® 4.0	49	1	6	26	7	8

As shown, Numerically Scoring these data resulted in the fewest correct decisions and greatest number of no opinions of any evaluation method. In fact, Numerical Scoring's inconclusive rate is twice that of PolyScore's. Despite one more miss on the deceptive cases, PolyScore scored 4 more deceptive cases correctly and 9 more truthful cases correctly than Numerical Scoring. The disparate performance of Numerical Scoring on deceptive and truthful data is consistent with results we presented at the American Polygraph Association Annual Meeting in San Diego on Aug. 7, 1997. Similar results have been reported in *Polygraph* (Blackwell, N. J., PolyScore 3.3 and Psychophysiological Detection of Deception Examiner Rates of Accuracy When Scoring Examinations from Actual Criminal Investigations, *Polygraph*, Vol. 28, No. 2, 1999).

Of the 8 Dollins cases that PolyScore scored incorrectly, only one case was scored correctly by either another algorithm or Numerical Scoring. Of the 14 cases that PolyScore evaluated as No Opinion, 10 were also evaluated as No Opinion by at least one other algorithm. In 6 of these 14 No Opinion cases, when another algorithm made a decision it was always incorrect.

One of the 8 cases that PolyScore scored incorrectly was also scored as deceptive by every algorithm and Numerical Scoring, yet is supposedly confirmed truthful. However, the relevant questions:

R5 HAVE YOU TAKEN PART IN ANY SEXUAL BEHAVIOR WITH A CHILD THAT YOU HAVE FAILED TO TELL ME ABOUT?

R7 HAVE YOU HAD ANY SEXUAL BEHAVIOR WITH A CHILD THAT YOU HAVE NOT DISCUSSED WITH ME?

R10 HAVE YOU WITHELD ANY VICTIMS FROM ME?

appear to be from a post-confession clearing examination. Since these issues seem impossible to confirm as truthful by a confession of others, we suspect a clerical error.

With respect to the issue of having more false positives than false negatives on these data ($p=.006$, if we include the above

suspected clerical error as a true false positive), we can only state that PolyScore is designed to be unbiased. While the authors of the algorithm comparison speculate that this may be related to the training data upon which PolyScore and other algorithms are based, the bias may actually lie outside of the scoring process. That is, it may be that in a failure of the polygraph process, it is more likely for a truthful person to appear deceptive than for a deceptive person to appear truthful. Finally, scoring bias is just one factor to consider in an algorithm's performance. Consider two algorithms. The first, unbiased algorithm has both a 10% false positive rate and a 10% false negative rate. However, an improvement to the first algorithm produces a second algorithm whose false negative rate is only 5%. Clearly the second algorithm is more accurate than the first, although it is no longer unbiased.

In summary, PolyScore 4.0 performed well in this study. It achieved 90.4% accuracy with a 14.4% Inconclusive (No Opinion) rate. No other algorithm or Numerical Scoring scored as many correct or had as few No Opinions. We would like to encourage a follow-on study of this sort using a much larger database and including Numerical Scoring. We are confident that if the trends shown in the figure above continue, PolyScore 4.0 and the forthcoming PolyScore 5.1 will take the cake.

A Response to Comments

Donald J. Krapohl & Donnie W. Dutton

We would like to once again thank those who contributed their time and effort by (a) submitting data for our confirmed case database and by (b) analyzing the data we sent them. We would also like to thank the authors of the Axciton, CPS, Identifi, and PolyScore computer programs for their cogent comments throughout the project. We, too, would like to see further comparisons of this nature and hope to sponsor those comparisons in the future.

We agree completely with comments regarding possible biases within the sample of data that was sent for analysis. The data were voluntarily sent by the identified sources and we relied on the integrity of those sources confirmation of ground truth. The data were definitely not textbook charts and we had discussions both among ourselves and with vendors regarding whether some of the data could be evaluated. We did keep all of the original data with the exception of the six cases for which we were unable to re-confirm ground truth. All evaluation systems made the same decision on the case Mr. Harris believes was misclassified. Thus, all evaluation systems benefited or suffered equally if the case was misclassified. We regret that we were unable to provide a larger sample of homogenous cases. We urge examiners to send data to the Institute using the confirmation form published in the American Polygraph Association *Newsletter*. If we receive enough cases we should be able to support another algorithm comparison with a larger and more homogenous sample.

Mr. Hedges' suggestion of a quality control step to reject inferior examinations from the study is a good one. We were surprised to learn that at least one evaluation system evaluated every case in the sample, that is, each case had at least one definitive decision from an algorithm. We truly expected some of the cases to be classified at No Opinion by all of the computer programs. As most in the profession are painfully aware, there simply are no standardized guidelines regarding the acceptability of recorded data. We would welcome suggestions for objective quality control criteria. Mr. Hedges' suggestion of a comparison between examiner and computer program performance is also well worth considering. The question is, of course, who would the examiners be? Would it be acceptable to elicit the assistance of 3, or 5, or 10 examiners randomly chosen from the American Polygraph Association membership? The readers' input on this question would be appreciated, as well.

Drs. Kircher and Raskin are correct regarding the rounding of the event markers to the nearest second, and we do not know what effect this may have had on the rejection of EDRs with their scoring algorithm. The DoDPI has since funded development of a new conversion program that accurately converts event mark times. The differences in instrumentation with regard to the electrodermal channel of the AXCITON are, thanks to Dr. Cestaro (1998), a matter of record.

We agree with Mr. Harris that the Cochran's Q statistic is not ideal because it is applicable for dichotomous data. We also agree with Mr. Harris' suggestion that the Friedman test is more powerful than the Cochran's Q - when the assumptions for rank order data are met. We depart from Mr. Harris on the methodological question as to whether these data are rank ordered, that a no opinion result lies between one that is correct and one that is incorrect. We chose to report the more conservative, and we believe more appropriate, Cochran's Q statistic.

We did not obtain the same results with a Friedman test as did Mr. Harris, even using the ranking method for levels of accuracy he adopted. We calculated a Friedman Two-Way Analysis of Variance by Ranks as described by Zar (1996) as well as Siegel and Castellan (1988). We first assigned

ranks of 3 to correct decisions, 2 to no opinion decisions, and 1 to incorrect decisions-as described by Mr. Harris. We then ranked the decisions submitted for Chart Analysis, Identifi, CPS, and PolyScore. We calculated Friedman's Test Statistic using the ranked data for these four evaluation methods. The results for this omnibus test were a Friedman Test Statistic of 0.535 ($p = .911$). The hypothesis that there are statistically significant differences among the four evaluation methods would be rejected using a criterion of ($p < .05$).

We appreciate Mr. Harris' contribution of providing the results of the recent Johns Hopkins University Applied Physics Laboratory implementation of the Department of Defense Polygraph Institute scoring criteria and decision rules for Numerical Scoring. We remain unsure, however, of the manner in which the decisions were made (e.g., was this a composite of several evaluators or one individual's decisions), and therefore do not know how to compare them to the decisions of the other systems.

Finally, we would like to once again thank everyone who participated in this project. We realize that we could not have completed the project without the conscientious contributions of many people. We hope the end result is beneficial to us all.

References

Cestaro, V. L. (1998). Instrumentation for presenting a known standard signal to the electrodermal channel for assessing response characteristics. *Polygraph*, 27(3), 188-209.

Siegel, S., & Castellan, N. J. Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Zar, J. H. (1996). *Biostatistical analysis* (3rd ed.). New Jersey: Prentice Hall.

Polygraph in the Crimea: First Results and Future Perspectives

Gennady G. Moskal & Vitaliy I. Egorov

Key words: directive test, guilty knowledge test, non-directive test, peak of tension, searching test, Ukraine

For more than 70 years forensic psychophysiology has made essential advances, and clearly demonstrated its effectiveness in various areas. Unfortunately, despite early positive results obtained by famous Russian psychologists – A.R. Luria and A.N. Leontiev in 1920s – the Soviet government relegated polygraphy to the category of "pseudoscience". The polygraph was introduced to the Ministry of Internal Affairs of Ukraine in the Crimea in 1998 where two policemen were trained at a polygraph school in Krasnodar, Russia. At that time the Russian computerized polygraph "Barrier 14" was bought. Despite the absence of any legislative basis for application of polygraphy in the Ukraine, the first results obtained in the Crimea clearly demonstrated its effectiveness and necessity for extension of PDD examinations in the work of Criminal Investigation, and other departments. Since July 1999, a new Crimean PDD specialist was trained in the Axciton International Academy, and thus, received his education in the heart of polygraph science – the United States.

From November 1999 until the present, we conducted 62 PDD examinations exclusively on criminal cases, and most of them related to severe crimes. During that period we tested 54 males and 8 females. With the polygraph's help we found deception in six criminal cases; four murders, one theft and one illegal keeping of guns. Moreover, in the course of our PDD exams we have uncovered involvement of examinees in the following unreported crimes: keeping of narcotics (nine cases), theft (five cases), one illegal possession of guns, another murder, and one rape. Nondeception was found in 36 examinees.

Equipment and Tests: The Russian Experience

At the present four countries produce computerized polygraphs; the USA, Japan, China and Russia. Despite numerous experiments and attempts to modernize traditional channels that register human psychophysiological responses, four channels remain "classic"; upper and lower breathing, cardiovascular and electrodermal responses. During our examinations we used the Russian Barrier 14 computerized instrument. This polygraph registers two pneumograph channels, pulse rate, arterial pressure, electro-dermal activity, a photoplethysmograph, and tremor. The computer software that scores the physiological data was also developed in Russia, and is called Sheriff 6.0 software.

It is important to point out some principal differences of PDD tests in Russia and the Ukraine. In our experience we use only three types of tests in PDD exams, called "directive", "non-directive" and "searching". Let us briefly describe their principal contents, which may be of some interest to Western forensic psychophysiologists.

Drs. Moskal and Egorov are affiliated with the Main Board of Ministry of Internal Affairs of Ukraine in the Crimea, Laboratory of Forensic Psychophysiology, B. Khmelnizkogo str. 4 Simferopol, Crimea 95000. Ukraine. Request for reprints should be directed to Dr. Vitaliy Egorov, MD, per. Kronshtadtsky 8, Apt. 21, Simferopol, Crimea 95006, Ukraine, or via e-mail to viteg@mail.ru.

The Directive Test (DT)

This test is close to the MGQT. Generally, polygraphs of the Barrier type are set up to use three types of questions; irrelevant (neutral), relevant and control. The question sequence in the DT can be presented as following:

1. Is today Monday? - irrelevant
2. Is your first name ____? - irrelevant
3. Are we at the police station now? – irrelevant
4. Did you participate in the theft of the TV from Pushkina str.? – relevant
5. Do you live in the Ukraine? - irrelevant
6. Did you participate of the theft of TV from Kievskaya str.? – control
7. Is this the month of March? - neutral
8. Do you know who stole the TV from Pushkina str.? – relevant
9. Are you sitting on the chair? – irrelevant
10. Do you know who stole the TV from Kievskays str.? - control
11. Is it now spring? - irrelevant
12. Did you steal the TV from Pushkina str? – relevant
13. Is you last name ___? - irrelevant
14. Did you steal the TV from Kievskaya str.? – control
15. Do you live in the Crimea? - irrelevant

In this test we compare responses on relevant and control questions. The control question is composed to be similar to the actual crime, but where we are sure that the examinee did not participate. It therefore satisfies the scientific definition of a "control." Another important difference from American techniques is that we do not repeat the test three times. It can be used only one time per PDD examination. The other difference is the absence of the so called "no opinion" conclusions. Only "guilty" or "innocent" conclusions have been used in Russia. Also, it is necessary to mention that DT is not mandated for all PDD examinations. Most forensic psychophysiologists in Russia and Ukraine do not use this method regularly.

The Non-directive Test (NDT)

The NDT is the most widely used among Russian forensic psychophysiologists. It consists of two kinds of questions, irrelevant (neutral) and only one relevant. The general number of questions is not strictly limited, and can reach 15-16 questions. For instance, if we test an examinee suspected of the theft of a TV, the question sequence can be presented as the following:

> Regarding that theft, do you know if the criminal stole the following things from Pushkina str. video?

1. money?
2. drugs?
3. icons?
4. TV set? - relevant
5. video camera?
6. books?
7. foreign currency?, etc.

During PDD examination the NDT helps us to determine the knowledge of the examinee about details of a particular crime. By methodology, the NDT is close to the Known Solution POT. As in the DT format, the NDT is used only one time.

Searching Test (ST)

This type of test is close to the Searching POT. The test begins with one or two irrelevant questions, and other questions follow that could be relevant. In practice, any PDD examination starts and finishes with an ST. This type of test is used two or three times during a PDD examination.

Now, some words about the advantages of the Russian polygraph compared to American instruments. The important feature of the Barrier polygraph is the determination of individual levels of psycho-emotional responsiveness before the beginning of examination. Computer-based analysis of the human responses to the first irrelevant ("zero") question will demonstrate the level of emotional readiness – low, normal, high. One other feature is the examiner's opportunity to use particular channels for the examination; all or a selected subset. For example, we can switch off the pneumograph, or the electrodermal channels during the examination. Also, the sensor measuring pulse rate and arterial pressure don't impose physical discomfort by mechanical pressure, as is common for the US polygraphs. Russian software offers the option to select periods between questions, from 5 to 25 seconds, and the intervals of scoring psychophysiological responses, from 5 to 25 seconds.

Conclusion

Forensic psychophysiology has made its first steps in the work of our protective institutions. Most policemen were initially relatively skeptical of a "lie detector" in their professional field. However, early positive results, and other obvious benefits have changed their perspective. From the results of the Criminal Investigation Department, Main Board, Ministry of Internal Affairs of Ukraine in the Crimea, we can prove that forensic psychophysiology has great prospects in the Ukraine. It is extremely important to take into account the high rate of criminality in Ukraine and all countries of the former Soviet Union (FSU), the high level of organization of the criminal structures, and the well developed international network of organized crime. Of course, polygraphy cannot be a panacea for the severe disease called criminality. Nevertheless, forensic psychophysiology is rapidly developing in Russia and the Ukraine, and Russia is annually organizing international conferences on the experience and results of polygraphy in the countries of FSU.

A Critical Analysis of Amsel's Comparative Study of the Exclusive v. Nonexclusive Comparison Question

James Allan Matte and Cleve Backster

Key words: criticism, exclusive comparison question, field study, laboratory study, nonexclusive comparison question, validity

In a field study of the relative effectiveness of exclusive and nonexclusive probable-lie comparison questions, wherein 87 confirmed field examinations used exclusive comparison questions and 143 confirmed field examinations used nonexclusive comparison questions, Amsel (1999) reported that the nonexclusive comparison questions had significantly larger mean numerical scores than the exclusive comparison questions, and the nonexclusive comparison questions had scores significantly more in the correct direction than those produced by exclusive comparison questions. Amsel used a variation of the Backster Zone Comparison Test format. Amsel cited two earlier laboratory studies (Horvath, 1991; Palmatier, 1991) that supported his findings.

Background

The Zone Comparison Technique, developed by Cleve Backster, has historically employed non-current exclusive probable-lie comparison questions[1] that use a time bar that significantly separates the comparison question, usually by several years, from the period that the crime occurred. It has always been Backster's contention that nonexclusive comparison questions, with the guilty subject, on many occasions allowed the comparison questions to act as weak relevant questions, thereby producing reactions competing with the stronger relevant question reactions. Furthermore, the Backster Zone Comparison Technique (ZCT) uses a 7-position scale to numerically score the physiological data, in order to provide degrees of response in each of the three parameters at a given spot location. The sacrifice relevant question used in the Backster ZCT is confined to the precise identification of the specific issue covered by the single-issue relevant test questions, thus also serves as a preparatory relevant question. Finally, the Backster ZCT does not end its test sequence with either a comparison or relevant question, in order to avoid possible end-of-question-series relief on a question that may be used for a spot analysis numerical evaluation.

The nonexclusive comparison question developed by John E. Reid for use in the Reid Technique, is not separated in time from the relevant issue, nor does it exclude the crime or matter contained in the relevant questions. Thus it is an inclusive probable-lie comparison question, but has been named by its employers as a nonexclusive comparison question. Furthermore, the Reid Technique employs two nonexclusive comparison questions against four relevant questions. The Reid technique does not employ a preparatory or sacrifice relevant question, nor symptomatic questions.[2]

Dr. Matte and Cleve Backster are frequent contributors to this journal. Reprint requests should be sent to Dr. James Matte, Matte Polygraph Services, 43 Brookside Drive, Williamsville, NY 14221-6915, or to his e-mail address, JamesAllanMatte@mattepolygraph.com.

[1] The definition of non-current exclusive comparison questions, current exclusive comparison question, and nonexclusive comparison questions can be found in Chapter 16, *Forensic Psychophysiology Using The Polygraph* (1996), J. A. Matte. J. A. M. Publications, Williamsville, New York.

[2] The Reid Technique does not employ a numerical scoring system with increasing threshold in the analysis of the physiological data, and in the field its decision making process includes behavior assessment and factual analysis.

Research Method Used by Amsel

In response to the first author's request for additional data, Amsel provided the following information, which is quoted below:

> "When I started the research I picked out from my business archive 230 different cases. The decision regarding which file to pull out was strictly based on the question whether the result of that specific test was verified or not. This way I had 138 NDI cases and 92 DI cases. When I first examined the charts I found out that most had 3 charts (excluding the stim test) while some had 4 and 5 charts (excluding the stim test). In order to create uniformity I decided that for the research purposes I would use only the scoring of the first 3 charts in those tests that had more than 3 charts (excluding stim test). By doing so I created a new situation in where the scoring of the tests that had 4 - 5 charts now totaled a figure that moved the outcome from a conclusive result to an inconclusive result.
>
> "So originally the 230 cases distribution was the following:

	NDI	DI	Total
Nonexclusive	81	62	143
Exclusive	57	30	87
Total	138	92	230

> "As I mentioned before in the research, for uniformity purposes, after the fourth and fifth charts were omitted, the distribution is the following:

	NDI	INC	DI	Total
Nonexclusive	81	7	55	143
Exclusive	57	9	21	87
Total	138	16	76	230

(Amsel, personal communication, April 27, 2000)

Amsel's research methodology did not attempt to determine the error rate, and selected only the first three charts for evaluation, which may have created an artificial inconclusive rate. Without an error rate for each type of test, it is impossible to determine which technique or comparison question used within the technique is superior. It could be argued that the technique that employed the exclusive comparison question with an inconclusive rate of 10% had no errors while the technique that employed the nonexclusive comparison question with an inconclusive rate of only 5% could have had an error rate of 15%.

It must be recognized that the psychological structure of the test format could have a direct impact on the effectiveness of the comparison questions. Furthermore, Amsel's claim that the nonexclusive comparison questions had significantly larger mean numerical scores than the exclusive comparison questions was based on his use of the 3-position scale which does not evaluate the degrees of response as in the 7-position scale. Amsel's report that the nonexclusive comparison questions had scores significantly more in the correct direction than those produced by exclusive comparison questions is based solely on his analysis of only the first three charts, to the exclusion of additional charts that might have resulted in correct decisions in all cases.

Discussion

The field study conducted by Amsel employed current exclusive comparison questions that excluded only the specific

instant crime, but not other crimes committed during that same period. The noncurrent exclusive comparison questions exclude not only the specific instant crime but also other crimes possibly committed during the period preceding the instant offense by several years. In the Backster and Matte Zone Comparison Techniques, different time bars covering different periods are used for each of the comparison questions to distinguish and differentiate each comparison question, in order to avoid or delay habituation.

Amsel's study employed a 3-position scale of numerical scoring which did not differentiate between a subtle reaction and a dramatic reaction. In the Backster and Matte Zone Comparison Techniques,[3] a 7-position scale is employed which does differentiate between subtle and dramatic reactions by evaluating the degrees of response in each of the three parameters at each spot location. Blackwell's recent (1999) field study found "the PDD examiners mean level of accuracy was 75.7% and 66.3% for the 7- and 3- position scoring scales, respectively." Blackwell stated that "[w]ithout exception, the overall level of accuracy generated by the examiners when using the 7-position scoring scale was higher than when using the 3-position scoring scale. The same was true when looking at the overall percentages for either the innocent examinations or the guilty examinations." Krapohl (1998) found that the 3-position scale with a cutoff (threshold) of +/-4 was statistically equivalent to the widely accepted 7-position scale with the +/-6 cutoff score (threshold). However, Krapohl also found that "the highly experienced raters in this study rarely used the full range of available values in the 7-position scale, employing the narrower range of the 3-position scale for about 90% of the question comparisons." Capps and Ansley (1992a) and Van Herk (1991), like Krapohl, found that the accuracy of the 7- and 3-position scales depended on the threshold used. The Backster and Matte Zone Comparison Techniques use an increasing threshold, whereas other Zone Comparison Technique modifications (DoDPI, Utah) employ a fixed threshold.

Amsel's study employed a sacrifice relevant question (SRQ) that violated the Backster concept and purpose of the SRQ in that it covered both the comparison and relevant test questions. The Backster SRQ is designed to identify with preciseness the specific issue covered by all of the relevant questions included in its single-issue test, and those relevant questions must cover only one and the same act. Hence the examinee, whether guilty or innocent of the instant offense, will perceive the SRQ as the first relevant test question dealing with the specific issue under investigation. The SRQ used by Amsel does not act as the first relevant question dealing with the specific issue under investigation, hence the innocent examinee is only afforded the first relevant question to vent his or her possible anxiety regarding the instant offense. Furthermore, the Backster SRQ also acts as a preparatory question for the introduction of the relevant questions, to direct the guilty examinee's psychological set onto the relevant questions.

The test structure used by Amsel in his field study used a comparison question as the last question in its test sequence, whereas the Backster Zone Comparison Technique uses a symptomatic question. The danger of employing a test question that is used for comparison as the last test question is that the examinee may relieve on the last test question regardless of its nature, if he or she is aware that it is the last question in the sequence. This could have the effect of degrading the effectiveness of that comparison question. Also, the innocent examinee may have a reaction to this last comparison question where its recovery or relief may be affected or distorted by the announcement of the end of the test.

Amsel employed a significantly greater number of nonexclusive comparison questions (62.2%) than current exclusive comparison questions (37.8%) yet makes a visual comparison (Figures 1 & 2) of the score ranges between the two types of comparison questions. Amsel acknowledged, however, that "it may be argued that this result is due

[3] The Department of Defense Polygraph Institute (DoDPI) also uses the 7-position scale in the scoring of the physiological data collected from examinees in PV examinations.

to the larger number of examinations conducted using nonexclusive comparison questions."

Amsel criticized the first author for apparently failing "to adequately review the research literature addressing this issue before writing his text.[4] In 1991 Palmatier reported his research findings after replicating Horvath's (1991) study." However, Palmatier's laboratory study was an unpublished Master of Science degree thesis, not listed in any of the scientific literature.

The studies conducted by Horvath (1991) and Palmatier (1991) were both laboratory studies, in which the examinees may lack the fear of detection, and the lack of the milder, yet important fear relating to character aspects on the part of the innocent examinees. Also lacking was the fear of error by the innocent examinees, and the lack of anger, which can also cause an autonomic response. All are potentially present in field cases. In addition to the problems inherent in Horvath's laboratory study, some of which were identified in Amsel's study, Palmatier made several significant departures from standard field practice in his laboratory study. Palmatier dropped the last two relevant questions used with the Modified General Question Test (MGQT), attempting to convert an already administered MGQT examination to a Zone Comparison Test by evaluating only the first three relevant questions on that test. Hence, we are left to accept that the dropped relevant questions had no effect on the first three relevant questions nor their neighboring comparison questions nor the overall examination. There were other changes to the Backster protocol which may have influenced the findings. Palmatier used the same age category for all exclusive comparison questions. Palmatier used a relevant question as the last test question, as did Amsel. Finally, Palmatier used a stimulation test as the second test. It is the contention of the authors that a second-chart stimulation test may create the potential for an inconclusive or false positive.

The cuff pressure used in Palmatier's laboratory study was between 40 and 55 mm/Hg which is considered inadequate by some, technique-wise, for cardiograph recording, and counterproductive for the pneumograph recording. Furthermore, cuff pressure of 70 mm/Hg or more may divert the examinee's attention from his or her breathing to the cuff pressure. The redirected attention away from one's breathing could produce potentially truer, uncontrolled respiratory patterns. In many field studies, respiration was shown to have equal diagnostic value, and in some field studies greater diagnostic value, than its neighboring parameters (Buckley & Senese, 1991; Elaad, 1985; Elaad & Kleiner, 1990; Matte, Reuss 1992; Nakayama & Yamamura, 1990; Slowick & Buckley, 1975). An experimental scoring technique proposed and tested by Jayne (1990) also supported the pneumograph as providing the most diagnostic information. Furthermore, a study by Elaad, Bonwitt, Eisenberg, Meytes in 1982 revealed that respiration was the only one of the three parameters not affected by beta blockers. Elaad, et al, concluded that "respiration seemed to improve the overall detection rate especially because skin resistance responses have the quality of rapid habituation." Barland (1984) reported that a cuff pressure at 90 mm/Hg, and a mean arterial blood pressure of 100 mm/Hg before reaction which increases to 120 mm/Hg during reaction will show a difference in pulse amplitude of 200%, whereas a cuff pressure at 60 mm/Hg and a mean arterial blood pressure of 100 mm/Hg before reaction which increases to 120 mm/Hg during reaction will show a difference in pulse amplitude of only 50%. In the final analysis, typical of virtually all laboratory studies, Horvath and Palmatier did not produce the emotional and psychological elements experienced by examinees suspected of real-life crimes, and their departures from the Backster protocol for the Zone Comparison Technique limited the generalization of the findings.

In view of the above analysis, it can be stated that questions remain regarding Amsel's conclusion that nonexclusive

[4] Matte's text was published in 1996, not 1997 as reported in the Amsel article.

comparison questions are better than exclusive comparison questions, specifically regarding their generalization to the Backster and Matte techniques. Future research should consider that the non-current exclusive comparison question and the nonexclusive comparison question were designed to be used within different test formats. We would contend that the type of comparison question is inseparable from the testing technique, and that mixing and matching among them leads to faulty conclusions regarding their efficacy.

References

Amsel, T. T. (1999). Exclusive or nonexclusive comparison questions: A comparative field study. *Polygraph*, 28(4), 273-283.

Amsel, T. T. (April 4, 2000). Request for data. Personal communication to first author.

Amsel, T. T. (April 27, 2000). Request for data. Personal communication to first author.

Amsel, T. T. (January 19, 2000). Letter from first author requesting data regarding Amsel's article entitled "Exclusive or Nonexclusive Comparison Questions: A Comparative Field Study" published in *Polygraph*, 28(4), 1999.

Amsel, T. T. (February 9, 2000). Letter to first author providing information regarding test questions used in Amsel's study published in *Polygraph*, 28(4), 1999.

Backster, C. (1969). *Standardized polygraph notepack and technique guide.* New York: Backster Research Foundation.

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

Barland, G. H. (June 1-2, 1984). The Cardio Channel: A Primer: Paper presented at the Florida Polygraph Association Seminar. 15 pages.

Blackwell, N. J. (1999). PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examinations from actual criminal investigations. *Polygraph*, 28(2), 149-175.

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

Capps, M. H. & Ansley, N. (1992a). Comparison of two scoring scales. *Polygraph*, 21(1), 39-43.

Capps, M. H. & Ansley, N. (1992b). Analysis of federal polygraph charts by spot and chart total. *Polygraph*, 21(2), 110-131.

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

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

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

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

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

Horvath, F. (1991). The utility of control questions and the effects of two control question types in field polygraph techniques. *Polygraph*, 20(1), 8-25.

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

Krapohl, D. J. (1998). A comparison of 3- and 7-position scoring scales with laboratory data. *Polygraph*, 27(3), 210-218.

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

Matte, J. A. (1999). Numerical scoring systems in the triad of Matte polygraph techniques. *Polygraph*, 28(1), 46-55.

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

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

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

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

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

Swinford, J. (1999). Manually scoring polygraph charts utilizing the seven-position numerical analysis scale at the Department of Defense Polygraph Institute. *Polygraph*, 28(1), 10-27.

Van Herk, M. (1991). Numerical evaluations: Seven point scale +/-6 and possible alternatives; A discussion. *Polygraph*, 20(2), 70-79.

Addendum to 1990 Field Study of the Friendly Polygrapher Hypothesis

James Allan Matte and Thomas E. Armitage

Abstract

A field study of the Friendly Polygrapher Hypothesis (Orne, 1975) was conducted by Matte and Reuss (1990), which showed that the mean chart score for defense attorney cases under the privilege communication umbrella was –9.38, compared with police cases which showed a mean chart score of –9.10. These comparable values suggested similar states of differential autonomic arousal. The number of confirmed deceptive cases for defense attorneys was 34 versus 13 for the police cases. In view of the small number of police cases in the 1990 study, a search of the files of the Buffalo Police Department for verified deceptive cases since 1990 was conducted for the purpose of acquiring a larger sample for comparison with the aforesaid data from defense attorney cases. Thirty-two confirmed deceptive cases were found, and they proved to have a mean chart score of –8.6. The difference in the mean scores of the defense cases and the police cases were not statistically significant, adding to the growing body of evidence against the Friendly Polygrapher Hypothesis.

Key word: Friendly Polygrapher Hypothesis

The prior field study of the Friendly Polygrapher Hypothesis (FPH) by Matte and Reuss (1990) had been conducted to determine the validity of Dr. Martin Orne's theory. Several courts had uncritically adopted the FPH, including the California Court of Appeals in *People v. Adams* (1975). The Matte et al study used 34 verified deceptive defense attorney cases and 13 verified deceptive police cases for comparison of their mean chart scores to determine the degree of autonomic arousal for each group of examinees by their chart scores. The results of the 1990 field study revealed that the mean chart scores were very similar; -9.10 for the police cases versus –9.38 for the defense cases under the umbrella of privileged communication. However, the small sample for the police cases in the 1990 study warranted a search for a greater sample from the same police department which employs the identical technique as that used in the sample provided by the private firm in the initial study, the Matte Quadri-Track Zone Comparison Technique.

Method

Thirty-two verified deceptive psycho-physiological veracity (PV) examinations were collected from the files of the Buffalo Police Department for review by the second author for the period from 1996 through 2000. The table on the following page depicts the case number, the number of charts, the total score and the mean score. The mean scores were added for a total score of 270.8, which was divided by the number of cases (32), resulted in a mean chart score of –8.5 per chart for police cases. The polygraph used at the Buffalo Police Department was a Stoelting electronic four-pen, double-pneumograph, UltraScribe, which is the same type of instrument used in the collection of the physiological data for the defense attorney cases.

Dr. Matte is a regular contributor to this journal. Reprint requests should be sent to Dr. James Matte, Matte Polygraph Services, 43 Brookside Drive, Williamsville, NY 14221-6915, or to his e-mail address, JamesAllanMatte@mattepolygraph.com.

Table 1
Cases from Buffalo Police Department, Buffalo, New York
Verified Deceptive PV Examination Cases.
Technique: Matte Quadri-Track Zone Comparison Technique
Forensic Psychophysiologist: Thomas E. Armitage

Case Number	Number of Charts	Total Score	Mean Score
018-00	2	-20	-10.0
012-00A	3	-16	- 5.3
012-00B	2	-13	- 6.5
011-00	2	-13	- 6.5
284-99	2	-16	- 8.0
285-99	2	-20	-10.0
231-99	3	-24	- 8.0
234-99	2	-24	-12.0
151-99	3	-25	- 8.3
127.99	2	-22	-11.0
139.99	2	-23	-11.5
140.99	3	-30	-10.0
144-99A	3	-18	- 6.0
144-99B	2	-15	- 7.5
090-99	3	-32	-10.7
018.99	2	-21	-10.5
161-98	2	-15	- 7.5
081-98	2	-19	- 9.5
084-98	3	-18	- 6.0
083-98	3	-38	-12.7
301-97	3	-29	- 9.7
244-98A	2	-11	- 5.5
244-98B	2	-15	- 7.5
251-98A	3	-23	- 7.7
251-98B	2	-14	- 7.0
224-97	2	-15	- 7.5
227-97	2	-11	- 5.5
230-97	2	-13	- 6.5
232-97A	2	-16	- 8.0
232-97B	2	-21	-10.5
188-96	2	-22	-11.0
173-96	2	-14	- 7.0
n= 32	Mean Score Per Chart:		-8.5

Conclusion

The data of this current field study clearly show that the mean scores for all of these guilty cases are similar, and well beyond the required threshold for making the deceptive decisions. The minimum score required in the Matte Quadri-Track Zone Comparison Technique for a determination of deception must average –5 per chart, with a minimum of two charts conducted. Therefore, two charts require a minimum score of –10 and three charts require a minimum score of –15 before a determination of deception can be rendered. The data in this field study show that the mean score for the confirmed deceptive police examinees was –8.6 per chart. The mean score for defense attorney examinees under the umbrella of privileged communication reported in the 1990 study that employed the same type of polygraph and the same identical polygraph technique was –9.38 per chart, which was not statistically different from the present data (z=1.64, p=.06,

ns). The combined data reported in this study fail to support the concept that the degree of autonomic arousal is different for cases conducted in these two settings; defense attorney or police.

References

Matte, J. A. & Reuss, R. M. (1990). A field study of the "Friendly Polygraphist" concept. *Polygraph*, 19(1), 1-8.

Orne, M. T. (1975). Implications of laboratory research for the detection of deception. In N. Ansley, (ed.) *Legal Admissibility of the Polygraph*. Springfield, Illinois: Charles C. Thomas: 114-116.

People v. Adams (1975) 53 Cal.App.3d 109 [125 Cal.Rptr. 518].

The Published Works of Norman Ansley

Genni Arledge

Journal Articles

Ansley, N. (1953). Testing the physically handicapped. *Board of Polygraph Examiners Bulletin*, 3 p.

Ansley, N. (1959). A comparison of arm-cuff and wrist-cuff blood pressure patterns in polygraph charts. *Journal of Criminal Law, Criminology, and Police Science*, 50(2, July-August), 192-194.

Ansley, N. (1972). Acceptance test for instruments. *Polygraph*, 1(1), 32-35.

Ansley, N. (1972). The Keeler Polygraph Institute. *Polygraph*, 1(1), 27-31.

Ansley, N. (1973). Israeli police polygraph school. *Polygraph*, 2(2), 146-147.

Ansley, N. (1973). Vermont State Police polygraph service. *Polygraph*, 2(3), 239-244.

Ansley, N. (1974). The education and training of Maryland polygraph examiners. *Maryland Polygraph Review*, (October), 1-5.

Ansley, N. (1974). A Maryland licensing bill. *Polygraph Review*, 1(1), 11-12.

Ansley, N. (1974). U.S. District Court supports police polygraph screening. *Polygraph*, 3(4), 412-416.

Ansley, N. (1975). The 1975 hearings before the District of Columbia Council. *Polygraph Review*, 2(Fall), 1-7.

Ansley, N. (1975). The autonomic nervous system: A film review. *Polygraph*, 4(3), 289.

Ansley, N. (1975). Police screening - The results of 413 examinations. *Maryland Polygraph Review*, (April), 2-7.

Ansley, N. (1975). The polygraph profession: An opportunity for psychologists. *Maryland Psychological Association Newsletter*, (November), 4-6.

Ansley, N. (1975). Stipulated admissibility reversed by Georgia Appellate Court. *Polygraph*, 4(4), 340.

Ansley, N. (1976). Examiner reviews criticism by psychologist. *The Hennepin Lawyer*, 44(5, May-June), 5, 31.

Ansley, N. (1976). The polygraph profession: An opportunity for psychologists. *California State Psychologist*, 11(1), 5-6.

Ansley, N. (1976). Testing the physically handicapped. *Polygraph*, 5(1), 59-67.

Ansley, N. (1976). U.S. v. Oliver, a federal landmark decision. *Security Management*, 20(2), 28-29.

Ansley, N. (1977). Drugs for hypertension and polygraph results - A case example. *Polygraph*, 6(1), 53-57.

Ansley, N. (1977). Free polygraph service for unjustly convicted. *Federal Probation*, 41(3, September), 51-52.

Ansley, N. (1977). Van Gregory Johnson v. Maryland, a paradoxical precedent. *Polygraph Review*, 3(1), 1-6.

Ansley, N. (1979). Polygraph review. *Polygraph*, 8(1), 78-79.

Ansley, N. (1979). Review of "The polygraph and probation", by John O. Knowlton. *Polygraph Law Reporter*, 2(2).

Ansley, N. (1980). General problems. Polygraph Review, 1(2), 1-17.

Ansley, N. (1980). Polygraph review. *Empire State Polygraph Society Newsletter*, 4(2), 6-13.

Ansley, N. (1980). Research on the validity of the relevant-irrelevant technique as used in screening. *Polygraph*, 9(3), 153-161.

Ansley, N. (1980). Some observations on question formulation. *Polygraph Review*, 1(1), 1-21.

Ansley, N. (1981). Habeas corpus relief and exclusion of polygraph results. *Polygraph Review*, 2(2), 1-7.

Ansley, N. (1983). A compendium on polygraph validity. *Polygraph*, 12(2), 53-61.

Ansley, N. (1987). The status of polygraph evidence in federal courts. *Polygraph*, 16(1), 75-78.

Ansley, N. (1988). Grand jury admissibility. *Polygraph*, 17(3), 118-119.

Ansley, N. (1988). PL 100-347: Employee Polygraph Protection Act of 1988. *Polygraph*, 17(4) 138-148.

Ansley, N. (1988). Vittorio Benussi and research: Historical note. *Polygraph*, 17(3), 121-122.

Ansley, N. (1989). A compendium on polygraph validity. *The Narc Officer*, 5(5, May), 43-45.

Ansley, N. (1989). Law notes: Case decisions on admissibility. *Polygraph*, 18(2), 117-121.

Ansley, N. (1989). Law notes: Jury polled. *Polygraph*, 18(1), 46-47.

Ansley, N. (1989). Law notes: Text of H.R. 3451 to amend EPPA, and case abstracts. *Polygraph*, 18(4), 223-229.

Ansley, N. (1989). Law notes: Two jurisdictions and admissibility: Federal and case notes and abstracts. *Polygraph*, 18(3), 177-182.

Ansley, N. (1989). Resource materials on the polygraph and child abuse cases. *Polygraph*, 18(1), 51-53.

Ansley, N. (1990). Law Notes. *Polygraph*, 19(3), 210-220.

Ansley, N. (1990). Law notes: Civil and criminal cases. *Polygraph*, 19(2), 72-102.

Ansley, N. (1990). Quick Reference Guide To Polygraph Admissibility, Licensing Laws, And Limiting Laws. *Polygraph*, 19(2), 140.

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

Ansley, N. (1990). Was He Driving? *Polygraph*, 19(2), 147-149.

Ansley, N. (1990). Zone comparison is the proper name. *Polygraph*, 19(2), 161-163.

Ansley, N. (1992). Historical note: A.R. Luria: Motor reactions and lie detection in the 1920s. *Polygraph*, 21(1), 80-81.

Ansley, N. (1992). The history and accuracy of guilty knowledge and peak of tension tests. *Polygraph*, 21(3), 174-247.

Ansley, N. (1992). A. R. Luria: Motor Reactions And Lie Detection In The 1920s. *Polygraph*, 21(1), 80-81.

Ansley, N. (1992). A review of polygraph case law in 1991. *Polygraph*, 21(1), 58-71.

Ansley, N. (1992). Twenty years of *Polygraph*. *Polygraph*, 21(2), 146-153.

Ansley, N. (1993). Review of polygraph case law for 1992-1993. *Polygraph*, 22(2), 194-226.

Ansley, N. (1994). Appellate Decisions Involving Polygraph Issues. *Polygraph*, 23(1), 95-104.

Ansley, N. (1994). Legal Articles On The Employee Polygraph Protection Act Of 1988. *Polygraph*, 23(1), 112-115.

Ansley, N. (1995). Simulations and polygraph research. *Polygraph*, 24(4), 275-289.

Ansley, N. (1998). The irrelevant question: a descriptive review. *Polygraph*, 27(4), 276-283.

Ansley, N. (1998). Question Formulation. *Polygraph*, 27(3), 181-187.

Ansley, N. (1998). The validity of the modified general question test (MGQT). *Polygraph*, 27(1), 35-44.

Ansley, N. (1998). The zone comparison test. *Polygraph*, 27(2), 108-122.

Ansley, N. (1999). Development of deception criteria prior to 1950. *Polygraph*, 28(1), 28-36.

Ansley, N., & Beaumont, L.S. (1992). Quick reference guide to polygraph admissibility, licensing laws, and limiting laws, 16th edition. *Polygraph*, 21(1), Center pull-out.

Ansley, N., & Furgerson, R.M. (1987). Pioneers in the polygraph: The Federal Bureau of Investigation in the 1930s. *Polygraph*, 16(1), 33-52.

Ansley, N. & Krapohl, D.J. (2000). The frequency of appearance of evaluative criteria in field polygraph charts. *Polygraph*, 29(2).

Ansley, N., & Pumphrey, J.K. (1986). Survey results: Polygraph exams of 100 police and fire applicants. *Journal of Healthcare Protection Management*, 2, 86-93.

Capps, M.H., & Ansley, N. (1992). Analysis of federal polygraph charts by spot and chart total. *Polygraph*, 21(2), 110-131.

Capps, M.H., & Ansley, N. (1992). Analysis of private industry polygraph charts by spot and chart control. *Polygraph*, 21(2), 132-142.

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

Capps, M.H., & Ansley, N. (1992). Comparison of two scoring scales. *Polygraph*, 21(1), 39-43.

Capps, M.H., & Ansley, N. (1992). Numerical scoring of polygraph charts: What examiners really do. *Polygraph*, 21(4), 264-320.

Capps, M.H., & Ansley, N. (1992). Strong control versus weak control. *Polygraph*, 21(4), 341-348.

Cobb, J., Brooks, J., Safire, W., Lykken, D.T., Wilson, G.C., Saxe, L., Edwards, D., Young, C.W., Boyce, C.J., Levchenko, S., Inman, B., Stilwell, R.G., & Ansley, N. (1985). To tell the truth: Do national security concerns warrant the expanded use of polygraphs by federal government? *Common Cause Magazine*, (September/October), 33-37.

Decker, R.E., Stein, A.E., & Ansley, N. (1972). A cardio activity monitor. *Polygraph*, 1(3), 108-124.

McKinnon, M., & Ansley, N. (1980). Selected abstracts of nonverbal deception for polygraph examiners. *Polygraph*, 9(4, December), 197-231.

Olsen, D.E., Ansley, N., Feldberg, I.E., Harris, J.C., & Cristion, J.A. (1991). Recent developments in polygraph technology. *Johns Hopkins APL Technical Digest*, 12(4), 347-357.

Olsen, D.E., Harris, J.C., Capps, M.H., & Ansley, N. (1997). Computerized polygraph scoring system. *Journal of Forensic Sciences*, 42(1), 61-70.

Weir, R.J., & Ansley, N. (1973). Reference works and the polygraph. *Polygraph*, 2(3), 257-262.

Weir, R.J., & Ansley, N. (1954). Interrogation rooms and facilities for polygraph examinations. *Board of Polygraph Examiners Bulletin.*

Weir, R.J., Ansley, N., & Atwood, W.F. (2000). We remember 1952. *Polygraph*, 29(2), 137-145.

Books

Ansley, N. (1953). *Directory of polygraph instruments.* Washington, DC: Board of Polygraph Examiners.

Ansley, N. (1978). *Admissibility of polygraph evidence in criminal and civil cases.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N. (1978). *Laws relating to the polygraph.*. Linthicum Heights, MD: American Polygraph Association.

Ansley, N. (1978). *Quick reference guide to polygraph admissibility.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N. (1980). *Quick reference guide to polygraph admissibility, 6th ed.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N. (1983). *Quick reference guide to polygraph admissibility, licensing laws, and limiting laws.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N. (1983). *A review of the scientific literature on the validity, reliability, and utility of polygraph techniques (3rd ed.).* Fort Meade, MD: National Security Agency.

Ansley, N. (1985). *Quick reference guide to polygraph admissibility, licensing laws, and limiting laws, 10th edition.* Severna Park, Maryland: American Polygraph Association.

Ansley, N. (1987). *Quick reference guide to polygraph admissibility, licensing laws and limiting laws, 11th ed.* Severna Park, MD: American Polygraph Association.

Ansley, N. (1987). *Quick reference guide to polygraph admissibility, licensing laws and limiting laws, 12th edition, revised.* Severna Park, Maryland: American Polygraph Association.

Ansley, N. (1987). *Quick reference guide to state laws on the polygraph.* Severna Park, MD: American Polygraph Association.

Ansley, N. (1988). *Quick reference guide to polygraph admissibility, licensing laws and limiting laws, 13th ed.* APA.

Ansley, N. (1989). *Accuracy and utility of RI screening by student examiners at DoDPI.* Fort Meade, MD: National Security Agency.

Ansley, N. (1989). *Quick Reference Guide to Polygraph Admissibility, Licensing Laws, and Limiting Laws, 14th Edition, Revised.* Severna Park, MD: American Polygraph Association.

Ansley, N., & Horvath, F.S. (1977). *Truth and science: A comprehensive index to international literature on the detection of deception (lie detector) technique.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N., Garwood, M., & Barland, G.H. (1984). *The accuracy and utility of polygraph testing.* Washington, DC: Department of Defense.

Ansley, N., Horvath, F.S., & Barland, G.H. (1983). *Truth and science: A comprehensive index to international literature on the detection of deception and the polygraph (lie detector) technique (2nd edition, revised).* Linthicum Heights, MD: American Polygraph Association.

Ansley, N., & Pumphrey, J.K. (1985). *Justice and the polygraph.* Severna Park, MD: American Polygraph Association.

Ansley, N., & Weir, R.J., Jr. (1956). *Selected papers on the polygraph.* Washington, DC: Board of Polygraph Examiners.

Ansley, N. (1999). *Vergennes, Vermont and the War of 1812: The Battle of Lake Champlain.* Severna Park, Maryland: Brooke Keefer Ltd. Editions.

Book Chapters

Ansley, N. (1956). The board of polygraph examiners. In N. Ansley (Eds.), *Selected papers on the polygraph* (pp. 63-66). Washington, DC: Board of Polygraph Examiners.

Ansley, N. (1975). Capillary responses as a polygraph channel. In N. Ansley (Ed.), *Legal admissibility of the polygraph* (pp. 257-265). Springfield, IL: Charles C. Thomas.

Ansley, N. (1975). Plethysmography - A bibliography. In N. Ansley (Ed.), *Legal admissibility of the polygraph* (pp. 333-338). Springfield, IL: Charles C. Thomas.

Book Reviews

Ansley, N. (1972). Book review of International registry of computer users among psycho-physiologists, by S. David Kahn and Emmett B. Swint, Jr., (Eds.). *Polygraph*, 1(4), 252.

Ansley, N. (1972). Investigation and preparation of criminal cases, federal and state - A book review. Polygraph, 1(3), 180-181.

Ansley, N. (1974). Book review of So who's afraid of a polygraph test?, by S. E. Turner. *Polygraph*, 3(3), 344.

Ansley, N. (1975). Book review of Cardiovascular psychophysiology: Current issues in response mechanisms, biofeedback, and methodology, by Paul A. Obrist, A. H. Black, Eugene R. Bleeker, Jasper Brener and Leo V. DiCara (Eds.). *Polygraph*, 4(2), 180-182.

Ansley, N. (1975). Book review of Electrodermal activity in psychological research, by William F. Prokasy and David C. Raskin. *Polygraph*, 4(4), 377-379.

Ansley, N. (1975). Book review of Scientific and expert evidence in criminal advocacy, by Juris G. Cedarbaums and Selma Arnold (Eds.). *Polygraph*, 4(4), 380-381.

Ansley, N. (1975). Book review of Undercover investigation, by J. Kirk Barefoot. *Polygraph*, 4(3), 288.

Ansley, N. (1979). Book review of Employee theft investigation, by J. Kirk Barefoot. *Polygraph*, 8(4), 384-385.

Ansley, N. (1979). Book review of How to sting the polygraph, by Douglas Williams. *Polygraph*, 8(3), 289.

Ansley, N. (1984). Book review of Ten years of *Polygraph*: A ten-year index, 1972-1982, by Albert D. Snyder and Janet K. Pumphrey, (Eds.). *Polygraph*, 13(3), 279.

Ansley, N. (1997). Book review of Forensic psychophysiology, using the polygraph: Scientific truth verification - lie detection. by James A. Matte. *Polygraph*, 26(1), 65-68.

Horvath, F. S., & Ansley, N. (1990). Book review of The polygraph test: Lies, truth and science, by Anthony Gale (Ed.). *Criminal Justice Review*, 15(1), 125-127.

Bibliographies

Ansley, N. (1973). Plethysmography - A bibliography. *Polygraph*, 2(2), 153-161.

Ansley, N. (1975). Plethysmography - A bibliography. In N. Ansley (Ed.), *Legal admissibility of the polygraph* (pp. 333-338). Springfield, IL: Charles C. Thomas.

Ansley, N. (1976). The pneumograph: A bibliography. *Polygraph*, 5(3), 269-272.

Ansley, N. (1976). The pretest interview: A bibliography. *Polygraph*, 5(1), 92-93.

Ansley, N. (1977). Stress in the voice - A bibliography. *Polygraph Review*, 3, 13-15.

Ansley, N. (1979). Bibliography on paternity and the polygraph. *Polygraph Law Reporter*, 2(3), 3.

Ansley, N. (1979). Peak of tension - A bibliography. *Polygraph*, 8(1), 68-72.

Ansley, N. (1979). Testifying in court - A bibliography. *Polygraph*, 8(3), 239-241.

Ansley, N. (1982). Analysis of polygraph charts: A bibliography. *Polygraph*, 12(1), 28-37.

Ansley, N. (1986). A bibliography of recent scientific studies on the polygraph. *Polygraph*, 15(3), 228-230.

Ansley, N. (1986). Bibliography on stipulated admissibility of polygraph test results. *Polygraph*, 15(1), 77-80.

Ansley, N. (1986). Hypnosis and the polygraph: A bibliography. *Polygraph*, 15(4), 276-278.

Ansley, N. (1989). The eye and lie detection - A bibliography. *Polygraph*, 18(4), 221-222.

Ansley, N. (1989). *Law and the polygraph: a bibliography.* Fort Meade, MD: National Security Agency.

Ansley, N. (1989). Management and supervision of police polygraph operations: A bibliography. *Polygraph*, 18(3), 168-176.

Ansley, N. (1989). *Polygraph evidence and courtroom testimony: a selected bibliography.*

Ansley, N. (1989). Resource materials on the polygraph and child abuse cases. *Polygraph*, 18(1), 51-53.

Ansley, N. (1991). Drugs and the polygraph: A bibliography. *Polygraph*, 20(3), 214-216.

Ansley, N. (1991). Emotion and Respiration: A bibliography. *Polygraph*, 20(1), 46-51.

Ansley, N. (1992). Polygraph and probation: A bibliography. *Polygraph*, 21(2), 143-145.

Ansley, N. (1993). Police Applicant Screening--A bibliography. *Polygraph*, 22(1), 120-124.

Ansley, N., & Horvath, F.S. (1975). Bibliography on screening of police applicants with the polygraph. *Polygraph*, 4(3), 282-284.

Ansley, N., & Horvath, F.S. (1977). *Truth and science: A comprehensive index to international literature on the detection of deception (lie detector) technique.* Linthicum Heights, MD: American Polygraph Association.

Ansley, N., Horvath, F.S., & Barland, G.H. (1983). *Truth and science: A comprehensive index to international literature on the detection of deception and the polygraph (lie detector) technique (2nd edition, revised).* Linthicum Heights, MD: American Polygraph Association.

Ansley, N., & Knill, B.L. (1989). Nonverbal detection of deception: A bibliography. *Polygraph*, 18(2), 90-105.

Ansley, N., & Knill, B.L. (1991). Psychological variables and the detection of deception - An indexed bibliography. *Polygraph*, 20(2), 144-158.

Testimony

Ansley, N. (1983). Testimony of Norman Ansley, Chief, Polygraph Division, Office of Security, National Security Agency. Hearings before the Subcommittee on Civil and Constitutional Rights of the Committee on the Judiciary, House of Representatives, 98th Congress, First and Second Sessions, on Presidential Directive on the use of polygraphs and prepublication review Washington, DC: GPO.

Ansley, N. (1983). Testimony of Norman Ansley. Hearing before the Committee on Government Affairs, United States Senate, 98th Congress, First Session, on National Security Decision Directive 84 (pp. 51-57). Washington, DC: GPO.

Ansley, N. (1984). Testimony of Norman Ansley, Chief, Polygraph Division, Office of Security, National Security Agency. Hearing before the Committee on Armed Services, United States Senate, 98th Congress, Second Session, on Polygraphs for counterintelligence purposes in the Department of Defense (pp. 44-53, 87-94). Washington, DC: GPO.

Ansley, N. (1984). Statement of Norman Ansley, Chief, Polygraph Division, Office of Security, National Security Agency. Hearing before the Investigations Subcommittee of the Committee on Armed Services, House of Representatives, 98th Congress, Second Session, on H.R. 4681, (pp. 45-62). Washington, DC: GPO.

Ansley, N. (1985). Memoranda laid before the Committee. "The Implications for Industrial Relations and Employment of the Introduction of the Polygraph". Employment Committee, House of Commons Session 1984-85 (p. 13, 19) London: Her Majesty's Stationary Office.

Erratum

In *Polygraph*, 29(1) an article appeared entitled “The State of Polygraph Testing on Sex Offenders Under Community Supervision in Texas.” The author, Mr. Brian McKay, would like to express sincere thanks to Margaret Griffin, employed by the American Probation and Parole Association as a sex offender management specialist with the Center for Sex Offender Management. In her former employment as a community supervision officer with Hunt County CSCD, Ms. Griffin conducted the 1996 and 1997 administrations of the Texas sex offender supervision survey. Her pioneering efforts have helped lead Texas probation agencies to identify and promote effective supervision techniques, and she has been an ardent proponent for the use of polygraph testing on sex offenders under criminal justice supervision. The author regrets any oversight which prevented this appreciation from being expressed in his article.