Polygraph & Forensic Credibility Assessment: A Journal of Sciencie and Field Practice

VOLUME 51 2022 NUMBER 2

Page Contents

68 **A Field Assessment of Automated Presentations of Polygraph Test**
Donald J. Krapohl, Donnie W. Dutton and Dani Pruett

74 **Survey of Computerized Polygraph Scoring Algorithms Using Kircher Features**
Raymond Nelson

88 **Effects of Comparison Question Type and Between Test Stimulation on the Validity of Comparison Question Test**
Charles R. Honts and Racheal Reavy

134 **A Discussion of Countermeasure Strategies and the Comparison Question Test**
Thomas Kuczek and Raymond Nelson

140 **Terminology Reference for the Science of Psychohysiological Detection of Deception**
Donald J. Krapohl, Mark Handler and Michael Lynch

Published Semi-annually **© American Polygraph Association, 2022**

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

Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice

Editor-in-Chief: Mark Handler
E-mail: polygrapheditor@gmail.com

Managing Editor: Nayeli Hernandez
E-mail: polygraph.managing.editor@gmail.com

******

Associate Editors: Rejean Belley, Ben Blalock, Tyler Blondi, John Galianos, Don Grubin, Maria Hartwig, Charles Honts, Matt Hicks, Scott Hoffman, Don Krapohl, Thomas Kuczek, Mike Lynch, Ray Nelson, Adam Park, David Raskin, Stuart Senter, Joseph R. Stainback IV and Cholan V.

APA Officers for 2022 - 2023

President & Chair - Don Krapohl
E-mail: apakrapohl@gmail.com

President Elect - Donnie Dutton
E-mail: ddutton443@gmail.com

Director 1 - Michael C. Gougler
E-mail: gougler.national@gmail.com

Director 2 - Raymond I. Nelson
E-mail: raymond.nelson@gmail.com

Director 3 - James B. McCloughan
E-mail:mccllougj@gmail.com

Director 4 - Chip Morgan
E-mail: lieguy@gmail.com

Director 5 - Erika Thiel
E-mail:eethiel589@gmail.com

Director 6 - Desiree Hurd
E-mail: deshurd89@yahoo.com

Director 7 - Lisa Ribacoff
E-mail: lisa@iigpi.com

Director 8 - Darryl Starks
E-mail: dstarks1961@gmail.com

Treasurer - Chad Russell
E-mail: lvmpd@me.com

General Counsel - Gordon L. Vaughan
E-mail: gvaughan@vaughandemuro.com

Seminar Chair - Pam Shaw
E-mail: shawpolygraph@gmail.com

Education Accreditation Committee (EAC)
Manager - Barry Cushman
E-mail: eac@apapolygraph.org

National Officer Manager - Lisa Jacocks
Phone: 800-APA-8037; (423)892-3992
E-mail: manager@apapolygraph.org

National Office Assistant - Jennifer Crawley
Email: officeassist@apapolygraph.org

Subscription information: *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice* is published semi-annually by the American Polygraph Association. Editorial Address is Editor@polygraph.org. Subscription rates for 2022: One year $150.00. Change of address: APA National Office, P.O. Box 8037 Chattanooga, TN 37414-0037. THE PUBLICATION OF AN ARTICLE IN *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice* DOES NOT CONSTITUTE AN OFFICIAL ENDORSEMENT BY THE AMERICAN POLYGRAPH ASSOCIATION.

©American Polygraph Association, 2022

Instructions to Authors

Submission deadline for next publication: January 31, 2023

Scope

The journal *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice* publishes articles about the psychophysiological detection of deception, and related areas. Authors are invited to submit manuscripts of original research, literature reviews, legal briefs, theoretical papers, instructional pieces, case histories, book reviews, short reports, and similar works. Special topics will be considered on an individual basis. A minimum standard for acceptance is that the paper be of general interest to practitioners, instructors and researchers of polygraphy. From time to time there will be a call for papers on specific topics.

Manuscript Submission

Manuscripts must be in English, and may be submitted, along with a cover letter, on electronic media (MS Word). The cover letter should include a telephone number, and e-mail address. All manuscripts will be subject to a formal peer-review. Authors may submit their manuscripts as an e-mail attachment with the cover letter included in the body of the e-mail to:

polygrapheditor@gmail.com

As a condition of publication, authors agree that all text, figures, or other content in the submitted manuscript is correctly cited, and that the work, all or in part, is not under consideration for publication elsewhere. Authors also agree to give reasonable access to their data to APA members upon written request.

Manuscript Organization and Style

All manuscripts must be complete, balanced, and accurate. Authors should follow guidelines in the *Publications Manual of the American Psychological Association*. The manual can be found in most public and university libraries, or it can be ordered from: American Psychological Association Publications, 1200 17th Street, N.W., Washington, DC 20036, USA. Writers may exercise some freedom of style, but they will be held to a standard of clarity, organization, and accuracy. Authors are responsible for assuring their work includes correct citations. Consistent with the ethical standards of the discipline, the American Polygraph Association considers quotation of another's work without proper citation a grievous offense. The standard for nomenclature shall be the *Terminology Reference for the Science of Psycho-physiological Detection of Deception* (2012) which is available from the national office of the American Polygraph Association. Legal case citations should follow the *West* system.

Manuscript Review

An Associate Editor will handle papers, and the author may, at the discretion of the Associate Editor, communicate directly with him or her. For all submissions, every effort will be made to provide the author a review within 4 weeks of receipt of manuscript. Articles submitted for publication are evaluated according to several criteria including significance of the contribution to the polygraph field, clarity, accuracy, and consistency.

Copyright

Authors submitting a paper to the American Polygraph Association (APA) do so with the understanding that the copyright for the paper will be assigned to the American Polygraph Association if the paper is accepted for publication. The APA, however, will not put any limitation on the personal freedom of the author(s) to use material contained in the paper in other works, and request for republication will be granted if the senior author approves.

A Field Assessment of Automated Presentations of Polygraph Test

Donald J. Krapohl[1]

Donnie W. Dutton[2]

Dani Pruett[3]

Abstract

Characteristics of polygraph examinations from a large polygraph program were coded to help determine whether there were any effects on polygraph decisions that might be attributable to the use of automation to present test questions during the testing phase of a polygraph examination. Among the 415 cases in this six-month exhaustive sample, a small effect was found for the number of test charts the examiners recorded between the automated and human conditions. No significant differences were found among the proportions of polygraph decisions when comparing examinations in which the examiner read the test questions to examinations in which the computer presented them. The study found no adverse effects for the use of the digital voice in testing within the constraints of the variables tested. Given the advantages that automated presentation of test questions offers for standardization and a more useful allocation of examiner attention, its use in field polygraph examinations warrants consideration.

[1]Director for Educational Services, Capital Center for Credibility Assessment, and regular contributor to this journal.

[2]Mr. Dutton is an APA Past President and the Vice President of the Capital Center for Credibility Assessment.

[3]Ms Pruett is one of the first four UK police polygraph examiners. Now employed by Behavioural Measures UK (BMUK).

The authors appreciate the helpful comments and suggestions provided by the editor and the anonymous reviewer to an earlier version of this paper.

This article is one in a series titled Best Practices. The views expressed are solely those of the authors and do not necessarily represent those of their employers or the American Polygraph Association. Comments can be sent to the first author at APAkrapohl@gmail.com.

Since their introduction more than 30 years ago, digital polygraphs have come to replace pen-and-ink analog instruments that were the mainstay for most of the history of the polygraph. As early as 1962 Dr. Joseph Kubis[4] was conducting feasibility studies on computerization in polygraph exams for the US Air Force, and the first computer-assisted polygraph was developed in the late 1980s (Raskin & Kircher, 1990). New capabilities brought about by the transition to computer polygraphs are many: decision-support algorithms, post hoc data processing, feature extraction, electronic file sharing, and automating portions of the examination process. It is this last capability that is the focus of the present project.

All the major suppliers of computer polygraphs have software capable of presenting the test questions during testing using an automated voice. Replacing the examiner's voice with the automated voice for presentation of test questions has certain advantages. Automated voices are consistent across all questions, alleviating a concern regarding unconscious emphasis on certain test questions or accusations of such emphasis. Timestamping of question onsets is more reliable with automated question presentations because the software initiates the question presentation precisely at the point the examiner presses the key to signal question onset. The reliability of timestamping of question onsets based on the human voice requires coordination between an examiner's voice and the examiner's ability to simultaneously initiate the event marker, a skill that is likely to vary among different practitioners. The use of automated timestamping with the digital voice offers more confidence in the displayed latencies between question onset and response onset, a factor that is considered when assessing whether a response is associated with a test question. Also, allowing the computer to present the test question reduces attention demands on examiners, thus freeing them to visually monitor examinees more closely, perhaps affording more opportunities to detect countermeasures. For programs that have many polygraph examiners, the automated voice reduces variability among examiners, providing a more uniform testing experience for all examinees. Automated voices are also easily presented through headphones, with the additional advantage of removing extraneous sounds that may induce physiological responses. Finally, there is tentative evidence that automated presentation of test questions may improve polygraph decision accuracy (Honts & Amato, 2007).

There are unknowns regarding whether there are differential effects between human and automated question presentations. It could be possible examinees process questions differently if they are asked by a human compared to the same questions being presented by a machine. If, for example, examinees are more comfortable lying to a computer than they are lying to a human, or the reverse, a differential effect may appear in test outcomes. We found no reports that speak to whether the automated option has an adverse effect on the proportion of inconclusive polygraph results, or the number of test charts necessary to avoid inconclusive results. Similarly, the literature is silent as to whether a digitized voice corresponds with higher rates of decisions of deceptiveness or decisions of truthfulness.

In this six-month project we recorded polygraph test outcomes for examiners who used the digital voice and another sample of examiners from the same organizations who used their own voices during testing. While ground truth was not available for most of these cases, we did search for differential effects on test results and the number of test charts.

Method

Testing Examiners

There were 47 examiners in this project. All were trained by the same polygraph education

[4] Perhaps of historical interest, Professor Kubis assumed Father Walter Summers' position at Fordham University when the latter prematurely died in 1938. Reverend Summers, as all students of polygraph history will recognize, was a researcher known for conducting deception tests using a recording electrodermal device, and for his use of "emotional standards," technical questions that approximate what today would be called "comparison questions."

program and participated in the same quality control oversight program. Among the 47 testing examiners, 25 used digitized voice and 22 presented the questions with their own voices. Examiners were free to choose which voice to use during their examinations. The examiners who used digitized voice (DV) had an average of 2.7 years of experience at the beginning of this project, with a range of two months to eight years. Examiners using their own human voices (HV) to present test questions had an average of 4.0 years of experience, with a same range as those who used digitized voices. The difference in experience between the groups fell short of statistical significance, $t(45) = 1.90$, $p = .06$, ns.

Cases

The period of data collection was from January 1 through June 30, 2022. A total of 415 cases were submitted for quality control review from a large offender management program and all were included in this study. The polygraph results decided by the quality control reviewers were used in place of those of the testing examiner if there were differences in results between them. This occurred in 44 cases, or 10.5% of the sample. The Empirical Scoring System provided the basis for all results (Blalock, Cushman & Nelson, 2009; Handler, Nelson, Goodson & Hicks, 2010; Nelson, Krapohl & Handler, 2008).

The four possible test results were No Significant Responses (NSR), Significant Responses (SR), Inconclusive (INC), and No Opinion (NO). An NSR result indicated the data were interpretable and signified a conclusion the examinee was likely truthful to all the relevant questions. An SR result also indicated interpretable data but that the conclusion was the examinee was probably deceptive to at least one relevant question. An INC outcome meant that the data were scorable but that the scores fell short of the thresholds for either an NSR or SR decision. An NO would be rendered if the data were not scorable (e.g., highly erratic tracings, suspected countermeasures, the session was terminated early, etc.). A minimum of three test charts were required. If after three charts the results would be INC, examiners recorded a fourth or fifth chart in an attempt to garner sufficient scores for an NSR or SR. No more than five scorable charts were permitted.

The instruments on which the cases were conducted were all produced by the Lafayette Instrument Company, either models LX5000 or LX6. The DV group submitted 244 of the cases with the remaining 171 cases coming from the HV examiners. Polygraph techniques included the mixed-issue screening Air Force MGQT with either two or three relevant questions, and the British One-issue Screening Test. Being field cases, ground truth was largely unavailable. Given the very restricted and potentially biased ground truth confirmation information, there was no attempt to compare ground truth against the decisions made by the two types of voice.

Table 1 is a cross tabulation of techniques for examiners using either their own voices or the digital voice. Tests of proportions (Bruning & Kintz, 1997) found those who used their own voices tested with the AFMGQT technique significantly more often than those who used the digital voice ($z = 2.24$, $p < 0.05$) and tested less often with the BOST ($z = 3.87$, $p < 0.05$). Differences in the proportion of sessions that were terminated did not reach significance ($z = 1.92$, ns)

Table 1. Cross tabulation of cases in which the examiner tested with her own voice or the digitized voice for the testing techniques of the AFMGQT and BOST, and when the session was terminated before testing was completed.

	AFMGQT	BOST	Session Terminated
Human Voice	142	29	3
Digital Voice	174	71	0

Procedure

As a routine business practice the test results, testing technique and number of test charts for all cases submitted for quality control review are recorded in an Excel spreadsheet, along with an identifier to indicate which type of voice the examiners were using during the testing phase of the examination. The present analyses used only those records. The cases were sorted and tallied for the present investigation.

As an archival study, there were no manipulation of variables or changes in testing procedures. Though of potential interest, neither examiner nor examinee opinions regarding their attitudes about the voices used in testing could be captured.

Results

Regarding a possible relationship between the testing voice and the test results, tests of proportion found no significant differences for any of the four test outcomes. See Table 2.

Table 2. Number and (proportions) of NSR, Inconclusive, No Opinion and SR results for polygraph cases in which either the examiner or the computer presented the test questions during testing. No significant differences were found between the type of voice for any of the examiner decisions.

	NSR	INC	NO	SR
Human Voice	77 (0.45)	22 (0.13)	12 (0.07)	60 (0.35)
Digital Voice	125 (0.51)	37 (0.15)	10 (0.04)	72 (0.30)

We also tried to determine whether the voice used during testing was related to the examiner's decision to record more data to reach a conclusion. Table 3 shows the frequency and (proportions) of cases that used 2, 3, 4 and 5 charts for the HV and DV methods. We compared each of the vertical columns using tests of proportions (Bruning & Kintz, 1997). Only one significant difference was found. It was the number of cases in which the session concluded before a minimum of three charts were recorded. The two-chart cases (early termination of the examination) occurred more often when the human voice was used than when the digital voice had been employed. See Table 3.

Table 3. Number and (proportions) of cases with 2, 3, 4 or 5 charts in which either the examiner or the computer presented the test questions during testing. * indicates a statistical difference at p < 0.05 between HV and DV.

	Number of Charts			
	2	3	4	5
Human Voice	3 (0.02)*	114 (0.67)	21 (0.12)	33 (0.19)
Digital Voice	0 (0.00)*	162 (0.66)	38 (0.16)	44 (0.18)

Discussion

Most polygraph schools teach examiners to present the test questions themselves rather than to rely on automation to manage that task. Most field examiners likewise prefer to use their own voices to present the test questions. Given the advantages discussed earli-

er in this article it remains unclear why this simple task has not been more widely given to automation.

Our evaluation of field data found the type of voice, human or automated, had no meaningful effect on the proportions of NSR, SR, INC or NO results. These null findings are unsurprising and may give reassurance to examiners who are considering implementing the automated voice in their own testing practices. Similarly, the number of test charts deemed necessary by the testing examiners was largely unaffected by the type of voice they used. The sole difference was found for the proportion of cases where the session was terminated after only two test charts, and for which the human voice accounted for more than did the digital voice. While of statistical significance, the difference consisted only of three cases in the HV condition and none in the DV condition. Also, there were eight separate comparisons made between the digital and human voice outcomes. Finding a significant result was more likely simply because so many comparisons were made. Further studies are encouraged to resolve whether the present findings are robust. In addition, it would be of interest and of practical value to solicit examiner and examinee attitudes about the voices used during testing.

Limitation

The source of the data in this study did not permit analyses as to whether the human or digital voice produced higher accuracy.

References

Blalock, B., Cushman, B., and Nelson, R. (2009). A replication and validation study on an empirically based manual scoring system. *Polygraph, 38*(4), 281 – 288.

Bruning, J.L., and Kintz, B.L. (1997). *Computational Handbook of Statistics* (4th ed.). New York, NY: Addison-Wesley (pp. 285 – 288).

Handler, M., Nelson, R., Goodson, W, and Hicks, M. (2010). Empirical Scoring System: A cross-cultural replication and extension study of manual scoring and decision policies. P*olygraph 39*(4), 200 – 215.

Honts, C.R., and Amato, S. (2007). Automation of a screening polygraph test increases accuracy. *Psychology, Crime & Law, 13*(2), 187 – 199.

Kubis, J.F. (1962). *Studies in Lie Detection: Computer Feasibility Considerations*. Project No. 5534, Task No. 553401, Prepared for Rome Air Development Center, Air Force Systems Command.

Nelson, R., Krapohl, D. J. & Handler, M. (2008). Brute force comparison: A Monte Carlo study of the Objective Scoring System version 3 (OSS-3) and human polygraph scorers. *Polygraph, 37*, 185 – 215.

Raskin, D.C., and Kircher, J.C. (1990, May 2). *Development of a Computerized Polygraph System and Physiological Measures for Detection of Deception and Countermeasures: A Pilot Study.* Preliminary Report. Contract 88-L655300-000. Scientific Assessment Technologies, Inc. Salt Lake City, UT.

Survey of Computerized Polygraph Scoring Algorithms Using Kircher Features

Raymond Nelson

Abstract

Published literature was surveyed for automated statistical classifiers that make use of a common set of physiological response features. These features were first introduced and described by researchers at the University of Utah during the 1980s, and have subsequently come to be described with the polygraph profession as *Kircher features*. These include the amplitude of increase for electrodermal and cardiovascular activity, along with a reduction of respiration activity. Constriction or reduction of vasomotor pulse amplitude can also be included. An interesting characteristic of these features, in addition to their statistical correlation with deception and truth-telling, they can be extracted from recorded time-series data both visually and via automated computer methods. Statistical classifiers based on these features include the following: Probability Analysis, a Rank Order Scoring System, an Objective Permutation method, a Bootstrap analysis method, the Empirical Scoring System/Multinomial, the Objective Scoring System (version 1 and 2), and the Objective Scoring System – version 3. Design characteristics of these analysis methods are summarized in the appendices.

Introduction

Polygraph scoring algorithms, like all data analysis algorithms, consist of several fundamental operations or common functions. These include: feature extraction, numerical transformation and data reduction, use of some form of likelihood function, and structured rules or methods for interpretation and classification of output or results. Because all data analysis begins with feature extraction – the identification of useful and informative variation within the available data – development of knowledge about useful response features is an area of knowledge that be foundational to the development of varied methodological approaches when developing solution for the other, subsequent, functions within a data analysis method.

An example of this is a feature set that has come to be referred to as the *Kircher features*, (Kircher, 1981, 1983; Kircher & Raskin, 1988) as first described using this moniker by Krapohl and McManus (1999). In brief, these features consist of the primary signal for each of the traditional recording sensors: electrodermal phasic response amplitude, phasic increase in relative blood pressure, and reduction of respiration activity. Vasomotor activity can also be included in this feature set. Other researchers (Harris, Horner & McQuarrie, 2000; Kircher, Kristjansson, Gardner & Webb, 2004; Podlesny & Truslow, 1993; Rovner, 1986) have also shown the effectiveness of these response features.

Development of a data analysis method is a process of first specifying the desired output information –such as a statistical classification of deception or truth-telling – and then deconstructing the process of achieving that goal into a coherent set of assumptions and a reproducible series of functions. All algorithms consist of four essential functions, including: feature extraction, numerical transformation and data reduction, some form of likelihood function, and a set of rules or procedures used to interpret the test result. Of course, other process descriptions are also possible. However, these four basic functions can be generalized to nearly all data analysis methods, whether manual or automated, and whether based in traditional statistical classification and prediction methods or machine-learning/ artificial-intelligence.

Working backwards from through these functions, interpretation, in this usage refers to the process of translating a numerical and statistical result into conceptual information that may be useful or informative to persons not intimate with, or not involved in, the testing process or data analysis process. In other words, interpretation serves to answer the question: what does the test result actually mean? Categorical test results (i.e., positive or negative, and other allegorical terms) are the most simplistic form of interpretation and provide the smallest amount of detail about the scientific meaning of a test result – but often provide the most practical or actionable form of interpretation. In polygraph field practice, categorical results are the result of a procedural decision rule [1]. However, categorical results can be a source of misunderstanding and confusion when they are naively expected to be infallible. As a rule, scientific tests are not expected to be infallible. Although subject to dichotomous interpretation, all scientific test results are fundamentally probabilistic and therefore subject to inherent uncertainty. One of the main goals of any scientific test is to quantify and/or reduce the degree of uncertainty associated with a conclusion.

A *likelihood function* is a device used to obtain a reproducible statistical value for the observed test data. Likelihood functions can take many forms, including both empirical and mathematical distributions. A likelihood function can also be thought of as the parameters and formulae used to calculate a reference distribution. A simple and practical example of a likelihood function is a published table of values for a reference distribution. The most basic and simple form of likelihood function is a numerical cutscore at which a

[1]Polygraph decision rules make use of aggregated numerical information, such as grand total scores and subtotal scores. [See Nelson (2018) for a description of various decision rules including the grand total rule (GTR), two-stage rule (TSR), subtotal score rule (SSR), and others.]

categorical test result is selected. Numerical cutscores can be thought of as associated with some statistical likelihood that a test result is correct or incorrect.

Before any statistical value can be calculated, recorded data, if not already numerical, must be subject to *numerical transformation and data reduction*. Procedures for obtaining these numerical values can vary widely, and can include the use of physical measurement where applicable, likert scales, rank-ordering, use of ratios, z-scores or other mathematical or statistical values. Data reduction methods can also vary, and can include summation, averaging, weighted averaging, discriminate functions, log functions, resampling, and other methods. The functional objective of data transformation and numerical reduction is to transform recorded test data to a set of useful numerical values, and to reduce those values to a small set of numbers for which a statistical value can be obtained.

Before any numerical or statistical results can be calculated, useful response features must be extracted from the recorded data. *Feature extraction* is the beginning of any algorithmic or procedural method for data analysis. All data are a combination of signal and noise. Feature extraction is the process of identifying and isolating the response information of interest, including the identification of response onset and response end. Under ideal circumstances the ratio of signal to noise is very high, and it is very easy to isolate useful signal information from useless noise. Feature extraction research is foundational to the development of solutions for all subsequent analytic functions[2].

Method

Published literature was surveyed for descriptions of statistical classifiers based on physiological response features described by researchers at the University of Utah (Kircher & Raskin, 1988). Design and development characteristics were enumerated for these methods. Information was sought for the type of decision method and type of statistical classifier, along with methods for numerical transformation and data reduction – including the selection of selecting relevant-comparison question pairs where applicable. General methods for the development of a statistical likelihood function are described for each analysis algorithm, Finally, procedural rules for interpretation or classification of deception and truth-telling are described for each analysis method, as applicable to single issue and multiple issue polygraph examinations.

Results

Seven different statistical classifiers were found in the published literature. These include: Probability Analysis (PA; Kircher & Raskin, 1988), a Rank Order Scoring System (ROSS; Honts & Driscoll, 1987; 1988), a Bootstrap Analysis Method (BAM; Honts & Devitt, 1992), the Objective Scoring System – versions 1-2 (OSS 1-2: Krapohl, 2002; Krapohl & McManus, 1999), an Objective Permutation Scoring method (OPS; MacLaren & Krapohl 2003), the Objective Scoring System version 3 (OSS-3; Nelson, Krapohl & Handler, 2008), and the Empirical Scoring System – Multinomial (ESS/ESS-M; Nelson, 2017a; Nelson, Krapohl & Handler, 2008).

Information of interest to this survey, in addition to the use of Kircher features and associated recording sensor, included the following: decision model, statistical classifier, numerical transformation, method or procedure for selecting RQ and CQ analysis pairs, data reduction method, type of likelihood function, and procedural decision rules. Type of decision model refers to the overall method by which a classification is achieved; this can include the use of a z-test, gaussian-gaussian signal discrimination, simple Bayes, or other method. A number of types of statistical classifiers were observed, including, p-values, posterior odds, point and cutscore comparisons that are mapped to TN and TP rates, and other methods. Polygraph field examiners who desire to

[2]For simplicity, this discussion ignores the series of important functions prior to data analysis, including sensor development, stimulus development, test administration, signal processing, and data recording.

better understand differences between manual and automated scoring methods may be interested in the variety of methods employed for the selection and comparison of RQ and CQ value pairs. Methods for numerical transformation included rank transformations, difference scores, z-scores, and other methods. Strategies for reducing sensor and question subtotals to grand total scores included summation, averaging and weighted averaging. A variety of types of likelihood functions were observed, including empirical distributions, multinomial distributions, bootstrap distributions, permutations, and other methods. Decision rules included the use subtotal scores for multiple issue exams, grand total scores for single issue exams and other procedural solutions. Design and development characteristics for each of these analysis methods is shown in Appendices A-G.

Conclusion

Kircher features have been a useful and effective solution to the challenges of polygraphic feature extraction since the 1980s. The existence of an easily identified and easily used feature set has facilitated the study and development of a variety of types of statistical classifiers. These features were first described in the development and of the Probability Analysis algorithm (Kircher & Raskin, 1988). At this time, seven different computer algorithms can be found in the published literature that make use of this common polygraph feature set. Many of these methods are available in commercial and professional products in use by polygraph field practitioners.

Because all of the surveyed analysis methods are based on Kircher features, they all included similar recording sensors. Some differences in signal processing may exist for different polygraph instruments. However, signal processing differences are beyond the scope of this project. A number of different types of statistical classifiers are included in the surveyed analysis methods. These include maximum likelihood estimation, linear discriminate analysis, gaussian-gaussian signal detection/discrimination, permutation and bootstrapping methods, and rank transformations. Some methods make use of simple Bayesian classifiers, for which the posterior result may be thought of pragmatically as a probability of deception or truth-telling.

Vasomotor response data, although it can be thought of as one of the Kircher features, is not included in most these analysis methods. Although information on vasomotor activity can be found in the published literature – including publications by some developers of available scoring algorithms – it has not been included in the structural model of the published and available scoring methods. Reasons for this have not been completely discussed. However, it can be assumed that vasomotor data would more likely have been included if it had improved the effect sized of the published structural models. It can therefore be hypothesized that vasomotor data, though perhaps correlated with the criterion of interest, may not have improved the structural models described in in the published literature. Addition of vasomotor data would require a sufficient basis of data with which to re-develop the various likelihood functions and study the resulting effect sizes. One analysis method, the ESS-M (Nelson, 2017a) does include a likelihood function that can include vasomotor data. However, published information does not show any difference in effect size when including the vasomotor information (Nelson, 2017b). A more complete understanding of the potential vasomotor response data and effect sizes for automated scoring methods will require replication and extension of these algorithm methods, in addition to the recalculation or redevelopment of associated likelihood functions.

Limitations of this project are several, and include the fact that this project is intended only to provide a descriptive summary of design characteristics of these different methods. No mathematical or procedural description of the identified scoring methods is included in this report. Another limitation of this survey is that it does not include other computerized analytic methods that make use of other scoring features. Other analysis methods may exist in publication, including methods that rely on proprietary and boutique feature extraction methods, and response features that are less familiar or intuitive for field polygraph practitioners – for example, spectral response features. This project is limited to methods

that exist in open publication, and does not include algorithms that are subject to proprietary or intellectual property restriction. Other research should address the need for information on those methods. Finally, no information on effect sizes is described for the algorithms included in this report. Other research should address the need for information on effect sizes – including sensitivity, specificity, false-positive, and false-negative rates, and associated errors of measurement. Future research should further investigate potential advantages to the various solutions to the series of challenges inherent to automated statical data analysis and classification.

This project is a brief description of conceptual, albeit non-technical, information that may be useful to readers who want an introduction to the topic of polygraph computer scoring algorithms, along with an introduction to the breadth of activity in this area throughout the past 35 or more years. In addition to the fact that the availability of a useful set of known response features has enabled a variety of researchers to study the application different statistical methods to classification of deception and truth-telling, Kircher features have the advantage that they can provide some intuitive understanding to field polygraph examiners who desire to understand what details of the recorded physiology is included in the analysis. Indeed Kircher feature can be extracted either manually or via automation. It is hoped that this information is useful or informative to those interested in polygraph data analysis algorithm development, and to field polygraph professionals who wish to more fully understand differences between various analysis methods and traditional manual scoring procedures.

References

Harris, J. Horner A. & McQuarrie, A. D. (2000). *An Evaluation of the Criteria Taught by the Department of Defense Polygraph Institute for Interpreting Polygraph Examinations*. Department of Defense Polygraph Institute report No DoDPI00-R-0007

Honts, C. R. & Devitt, M. K. (1992). *Bootstrap decision making for polygraph examinations*. Report number DoDPI92-R-0002. Department of Defense Polygraph Institute report No DoDPI92-R-0002. Reprinted in Polygraph, 3, (1), 1-47.

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

Honts, C. R. & Driscoll, L. N. (1988). A field validity study of rank order scoring system (ROSS) in multiple issue control question tests. *Polygraph, 17*, 1-15.

Kircher, J. C. (1981). *Computerized chart evaluation in the detection of deception*. University of Utah.

Kircher, J. C. (1983). *Computerized decision making and patterns of activation in the detection of deception*. Doctoral dissertation, University of Utah, Salt Lake City. Dissertation Abstracts International, 44, 345.

Kircher, J. C., Kristjansson , S. D., Gardner, M. D. & Webb, A. (2004). *Human and computer decision-making in the psychophysiological detection of deception*. University of Utah.

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

Krapohl, D. J. (2002). Short report: Update for the objective scoring system. *Polygraph, 31*, 298-302.

Krapohl, D. & McManus, B. (1999). An objective method for manually scoring polygraph data. *Polygraph, 28*, 209-222.

MacLaren, V. V. & Krapohl, D. J. (2003). Objective Assessment of Comparison Question Polygraphy. *Polygraph, 32*, 107-126.

Nelson, R., Krapohl, D. & Handler, M. (2008). Brute force comparison: A Monte Carlo study of the Objective Scoring System version 3 (OSS-3) and human polygraph scorers. *Polygraph, 37*, 185-215.

Nelson, R. (2017a). Multinomial reference distributions for the Empirical Scoring System. *Polygraph & Forensic Credibility Assessment, 46* (2), 81-115.

Nelson, R. (2017b). Updated numerical distributions for the Empirical Scoring System. An accuracy demonstration with archival datasets with and without the Vasomotor Sensor. P*olygraph & Forensic Credibility Assessment, 46* (2), 116-131.

Nelson, R. (2018). Practical polygraph: a survey and description of decision rules. *APA Magazine, 51*(2), 127-133.

Nelson, R. (2019). Literature survey of structural weighting of polygraph signals: why double the EDA? *Polygraph & Forensic Credibility Assessment, 48* (2), 105-112.

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

Rovner L. I. (1986). The accuracy of physiological detection of deception for subjects with prior knowledge. *Polygraph, 15*(1), 1–39

Appendix A. Probability Analysis

Probability Analysis (Kircher & Raskin, 1988)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
Decision model	The algorithm computes a discriminate function that serves as a statistical classifier – a statistical value value for which decision cutpoints can make categorical classifications of deception or truth-telling.
Statistical classifier	Bayesian analysis using a likelihood function obtained via discriminant analysis. Results can be thought of as a posterior probability of deception or truth-telling.
CQ Selection	For each sensor, between chart mean for all RQs is compared to the between chart mean for all CQs.
Numerical transformation	Numerical values transformed to z-scores for each sensor using combined RQs and CQs.
Data reduction	Z-scores are averaged between-charts for the individual sensors for all RQs, and for all CQs. Sensor z-scores are then combined using a structural weighting function that was obtained using linear discriminate analysis.
Likelihood function	Two likelihood formulas are used to calculate complimentary likelihood values for deception and truth-telling.
Decision rules – single issue	GTR
Decision rules – multiple issue	none
Comments	Structural coefficients are available from the developers and also from replication studies. Publication describe the application of PA to single issue exam formats. Application of the PA algorithm to multiple issue exams may require a change from the aggregation of RQs and CQs both within and between charts. Advantages of separate within-chart and between-chart transformation schemes have not been fully described in publication, however subsequent algorithm have shown the application of these to multiple issue exams. For this reason it may be possible to adapt PA to multiple issue test formats.

Appendix B. Rank Order Scoring System

ROSS (Honts & Driscoll, 1987, 1988)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular, Vasomotor
Decision model	ROSS decision model is similar to a gaussian-gaussian signal discrimination method, using empirically derived summed rank distributions for guilty and innocent cases.
Statistical classifier	Statistical classifiers are empirically derived TP, TN, FP, and FN rates.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
CQ Selection	CQs and RQs are not paired for analysis as in tradition polygraph scoring. Instead, rank order analysis begins with the assignment of integer rank scores to all test stimuli, including all CQs and all RQs together, within each recorded chart
Numerical transformation	Integer scores are assigned to rank order variance of RQs and CQs for each sensor, within each chart. Rank scores are assigned in reverse order, wherein the response with the greatest change in physiology is assigned a rank value equal to the total number of RQs and CQs, and the smallest response receives a rank value of 1.
Data reduction	Rank values are summed for all RQs for all charts, and also for all CQs for all charts. A CQtotal – RQtotal = RankDifference score is then calculated.
Likelihood function	Empirical distributions can be calculated for RankDifference scores, and numerical cutscores can be selected to achieve desired effect sizes.
Decision rules – single issue	GTR
Decision rules – multiple issue	SSR
Comments	Rank order transformations are a common non-parametric solution, and can sometimes optimize robustness with messy and difficult data with some potential cost due to the granularity of rank values. Each rank value is obtained by comparing each response to all other response (all other RQs and CQs), which may complicate assumptions of independent RQ variance for multiple issue exams.

Appendix C. Bootstrap Scoring System

Bootstrap Scoring System (Honts & Devitt, 1992)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular, Vasomotor
Decision model	Z-test of an observed CQ-RQ difference using a bootstrap null distribution.
Statistical classifier	P-value, indicating the likelihood of obtaining a score equal to or more extreme than the observed score under the null hypothesis that there is no difference between CQ and RQ value.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
CQ Selection	Each RQ is paired with the preceding CQs to calculate CQ – RQ difference values after transforming all RQ and CQ values to z-Scores.
Numerical transformation	For each sensor, presentations of all RQs and CQs, for all combined charts, are transformed to z-Scores. In this way all scores for all sensors and all charts have a common scale value that can easily subject to bootstrapping.
Data reduction	CQ – RQ = Z-difference scores were calculated for the z-Scores. Z-Difference scores are hypothesized to be loaded at greater than zero values innocent subjects and less and zero for guilty subjects. Z-difference scores are aggregated via summation to obtain a single Z-difference score for an exam.
Likelihood function	A null distribution is calculated for each exam by combining all RQ and CQ z-Scores, for all charts and all sensors, into a single vector, and then bootstrapping a null distribution (random sampling with replacement) while arbitrarily assigning values as CQ or RQ.
Decision rules – single issue	GTR
Decision rules – multiple issue	NA
Comments	This method was described using single issue exams with an equal number of RQs and CQs, but could be adapted to multiple issue test formats and test formats with unequal numbers of RQs and CQs.

Appendix D. Objective Scoring System (OSS/OSS-2)

OSS/OSS-2 (Krapohl & McManus, 1999; Krapohl, 2002)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular,
Decision model	Gaussian-Gaussian signal-discrimination model (signal detection classifier applied to a signal discrimination task).
Statistical classifier	P-value, indicative of the likelihood of the observed test statistic under the distribution represented by the training data confirmed as opposite of the selected (deceptive or truthful) classification.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
CQ Selection	Each RQ is compared to the preceding CQ
Numerical transformation	R/C ratios are transformed to integer scores using a distribution of uniform septile bins. OSS-2 7-position scores [-3, -2, -2, 0, +1, +2, +3] differ from tradition polygraph 7 position scores in that the range of OSS-2 scores can occur with equal likelihood, whereas traditional 7-position scores are loaded near 0 with scores further from 0 occurring less frequently.
Data reduction	Integer scores are aggregated via summation for each RQ, for all sensors and all charts. RQ subtotals are then summed for a grand total score. Because integer scores are aggregated via summation it makes no difference whether sensor scores are summed first between charts or within-charts.
Likelihood function	OSS-2 reference tables are empirically derived
Decision rules – single issue	GTR
Decision rules – multiple issue	none
Comments	OSS-2 likelihood functions (reference tables) available for single issue polygraph exams with 3 RQs and 3 charts. The summative design means that the likelihood function may be less robust with missing and artifacted data, and may be overloaded when more than three charts are used, and may become biased with test formats with unequal numbers of RQs and CQs. OSS likelihood functions have not been published for multiple issue exams, OSS-1 and OSS-2 began as manual scoring protocols, for which the structure and procedures were sufficiently structured and unambiguous that they led easily to automation. A result of this is that OSS and OSS-2 are now defacto automated analysis methods.

Appendix E. Permutation Scoring System

Permutation Scoring System (MacLaren & Krapohl, 2003)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular, Vasomotor
Decision model	A simple bayesian classifier using odds form of Bayes' theorem, where p/(1-p) values obtained are from a permutation of uniform 7 position integer scores.
Statistical classifier	Use of Bayes' theorem means that results can be thought of in practical terms as a posterior probability of deception or truth-telling.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
CQ Selection	RQs are paired with preceding CQs.
Numerical transformation	A ratio is calculated for each RQ/CQ pair, after which two sets of integer scores are assigned using two distributions of uniform septile bins that were calculated from confirmed guilty and innocent cases.
Data reduction	Two sets of integer scores are summed to obtain two grand total scores (guiltyTotal and innocentTotal) which are then compared to a PSS likelihood function.
Likelihood function	PSS likelihood function is the permutation of all possible 7 position scores if they are not systematically associated with guilt or innocence. The exact distribution includes 6.57 x 10^22 possible combinations. It can be calculated using a combinatoric formula, and can be easily approximated via simulation.
Decision rules – single issue	GTR
Decision rules – multiple issue	none
Comments	PSS was developed with examination consisting of three presentations of a question sequence that includes three RQs and 3 CQs. Adapting the PSS method to multiple issue exams, and to test formats with two or four RQs, and with four or five presentations requires available confirmed case data to calculate the uniform septile distributions, in addition to recalculation of the permutation likelihood function.

Appendix F. Objective Scoring System – version 3

OSS-3 (Nelson, Krapohl & Handler, 2008)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular
Decision model	Gaussian-Gaussian signal-discrimination model (signal detection classifier applied to a signal discrimination task).
Statistical classifier	P-value, indicative of the likelihood of the observed test statistic under the distribution represented by the training data confirmed as opposite of the selected (deceptive or truthful) classification.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure.
CQ Selection	Mean CQ is compared to each RQ.
Numerical transformation	Log R/C ratios standardized to the training data. Sensor scores are standard scores (mean = 0, standard deviation = 1), from -3 to +3. Standardized log R/C ratios indicated the number of standard deviations an observed response is above or below the mean of the training data when guilty and innocent cases are combined. These scores are intuitively similar to the notion of 7 position polygraph scores, but with decimals.
Data reduction	Grand mean is the mean of between-chart RQ scores. Between-chart RQ scores are the means of within-chart weighted mean sensor scores. Sensor scores are standardized log R/C ratios. Sensor weighting coefficients obtained through linear discriminate analysis, and can also be calculated via logistic regression, bootstrapping and other methods with little difference in the resulting weighting function.
Likelihood function	OSS-3 reference tables are empirically derived for confirmed guilty and innocent cases.
Decision rule – single issue	GTR, TSR
Decision rule – multiple issue	SSR, OSS-3 Screening rule uses the K-W ANOVA method to evaluate differences and similarities between RQs to reduce the occurrence of inconclusive with multiple-issue exams (based on an assumption that truth-telling to all RQs will result in no significant differences in RQ scores).
Comments	OSS-3 was intended extend available knowledge from OSS-1 and OSS-2 to a wide variety of examination formats, including single and multiple issue exams with 2, 3, and 4 RQs, and 3, 4, or 5 charts. OSS-3 was designed to be robust with some missing and artifacted data, and to use the 2nd of any repeated questions within a chart. Data reduction via averaging means that OSS-3 likelihood functions are more readily applicable to exams with 3, 4 or 5 presentations of the question sequence, and the log(RQ/CQmean) transform is applicable to test formats with unequal numbers of RQs and CQs. The algorithm also includes capabilities to mark artifacted and unusable segments for exclusion from analysis. Artifacted segments can be analyzed, using a test of proportions, to make inferences about their cause, whether systematic or random.

Appendix G. Empirical Scoring System (ESS/ESS-M)

ESS/ESS-M (Nelson, 2017a; Nelson, Handler & Krapohl, 2008)	
Sensors	Respiration (thoracic and abdominal), Electrodermal, Cardiovascular, Vasomotor
Decision model	ESS relies on a gaussian-gaussian signal discrimination model. ESS- is a simple bayesian classifier. ESS can also be studied and used with traditional/Federal cutscores to achieve empirically studied effect sizes.
Statistical classifier	ESS uses a p-value, obtained from an empirical distribution, that describes the likelihood of the observed data under a specified hypothesis. Cutscores for ESS were selected empirically, to constrain FN and FP errors to desired alpha levels. ESS-M results can be expressed as the posterior odds of deception or truth-telling. Use of Bayes' theorem means that results can be thought of in practical terms as a posterior probability of deception or truth-telling. ESS-M also provides the lower-limit odds of the 1-alpha posterior credible interval for deception or truth-telling – indicative of the likelihood of obtaining a similar categorical result upon repetition of the test procedure.
Response Features	Reduction of respiration activity, electrodermal amplitude of phasic response, cardiovascular phasic increase in relative blood pressure. Can optionally include vasomotor reduction of pulse amplitude.
CQ Selection	RQs are paired with CQs according to traditional procedures used by field examiners for each different polygraph test format. In general, RQs are compared to the preceding or subsequent CQ with the greater change in physiological activity whenever possible, and with the preceding CQ when two CQs are not available.
Numerical transformation	Integer scores are assigned by comparing differences in response magnitude for RQ and CQ pairs. Question pairs can be used naively or subject to optimization coefficients to reduce scores that may occur due to spurious or random noise. EDA integer scores are doubled prior to summation, so that the structural contribution of EDA data is greater than for other sensors. [See Nelson, 2019.]
Data reduction	Integer scores are summed for each RQ, for all sensors, between charts. RQ subtotals are then summed to obtain a grand total score. Summation via this process means that values are available for single issue and multiple issue test formats.
Likelihood function	ESS likelihood functions were calculated empirically using only respiration, EDA and cardiovascular sensors. ESS-Multinomial likelihood functions are calculated mathematically using the analytic theory of the CQT, and available for single issue and multiple issue polygraph exams with 2, 3, and 4 RQs with 3, 4 or 5 charts, including respiration, EDA, and cardiovascular sensors, in addition to the optional vasomotor sensor. Likelihoods for traditional cutscores are the empirical TP, TN, FP and FN rates.
Decision rules – single issue	GTR, TSR, FZR
Decision rules – multiple issue	SSR
Comments	ESS and ESS-M were introduced as manual scoring methods. Research publications have made use of fully automated ESS and ESS-M models.

Effects of Comparison Question Type and Between Test Stimulation on the Validity of Comparison Question Test

Charles R. Honts

Racheal Reavy

Editorial note regarding Honts & Reavy (2009) from Mark Handler – Editor in Chief

We reprinted the Honts & Reavy Final Report dated 2009 because not only was it one of the most comprehensive and well-designed studies to date in polygraph, it addressed a number of important issues. Also, interested readers should read the associated peer-reviewed publication for a discussion of "man versus machine" algorithmic scoring versus human:

Honts, C. R., & Reavy, R., (2015). The comparison question polygraph test: A contrast of methods and scoring. Physiology and Behavior, 143, 15-26. Published online 24 February 2015,

doi:10.1016/j.physbeh.2015.02.028.

A copy of the full text is available from this hyperlink:

https://www.academia.edu/11510782/The_comparison_question_polygraph_test_A_contrast_of_methods_and_scoring

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRD-ARL-RO-SS-SI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

☐ Reprint (Orig + 2 copies)
☐ Technical Report (Orig + 2 copies)
☐ Manuscript (1 copy)
■ Final Progress Report (Orig + 2 copies)
☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: W911NF-07-1-0670

REPORT TITLE: Effects of Comparison Question Type and Between Test Stimulation on the Validity of Comparison Question Test

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Charles R. Honts, Ph. D.
Psychology Department
Boise State University
1910 University Drive, MS-1715
Boise ID 83725-1715

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	8 September 2009	Final 28 Sept 2007 - 8 Sept 2009

4. TITLE AND SUBTITLE

Effects of Comparison Question Type and Between Test Stimulation on the Validity of Comparison Question Test

5. FUNDING NUMBERS

W911NF-07-1-0670

6. AUTHOR(S)

Charles R. Honts and Racheal Reavy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Boise State University, 1910 University Drive, Boise ID 83725

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This study examined the validity of two approaches to the comparison question test. Probable-lie and directed-lie comparison questions were evaluated in a mock crime experiment with 250 participants. Review of questions between charts was also manipulated Participants took a DACA style single issue polygraph examination. Resultant polygraph data were evaluated with the Objective Scoring System. Analyses found no evidence for significant differences between the validity of the probable-lie and directed-lie approaches. However, there was a significant effect of between repetition stimulation on decisions, reflecting an increased number of true positive outcomes when stimulation was employed and a higher number of false positive errors without stimulation. Although there were no significant effects of the approach to comparison questions, there is much to recommend the directed-lie approach: It is simpler and far more standardized. It is easy to teach and to use. A very small number of directed-lie questions can be used for essentially all examinations. These and other positive factors suggest that the directed-lie should be considered for wider application in field settings. The significant positive effect for between repetition stimulation recommends this practice for adoption in the field.

14. SUBJECT TERMS

Polygraph, Deception Detection, Comparison Question Test, Directed Lie Comparison, Probable Lie Comparison

15. NUMBER OF PAGES

46

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	**UL**

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

B O I S E S T A T E U N I V E R S I T Y

EFFECTS OF COMPARISON QUESTION TYPE AND BETWEEN TEST STIMULATION ON THE VALIDITY OF COMPARISON QUESTION TEST

FINAL PROGRESS REPORT ON

CONTRACT NO. W911NF-07-1-0670

DEFENSE ACADEMY OF CREDIBILITY ASSESSMENT (DACA)

CHARLES R. HONTS, PRINCIPAL INVESTIGATOR

RACHEAL REAVY, RESEARCH ASSOCIATE

PSYCHOLOGY AND LAW LABORATORY

DEPARTMENT OF PSYCHOLOGY

BOISE STATE UNIVERSITY

SUBMITTED

8 SEPTEMBER 2009

Executive Summary

This study examined the validity of two approaches to the comparison question test in a mock crime setting. To assess credibility the comparison question test contrasts subject responses to two types of critical questions. Relevant questions are syntacticly simple questions that directly address the issue being assessed by the examination (Did you shoot John Doe?). It is expected that guilty persons attempting deception to the relevant questions will produce consistent physiological responses to the relevant questions as they provoke memory of the event and concerns about deception detection. However, innocent individuals who are being truthful to the relevant questions might also respond to them because they recognize their importance in the examination. Comparison questions are designed to provoke consistent responses from the innocent. In the probable-lie approach to comparison questions the subject is maneuvered by the examiner into an answer that is probably a lie (Prior to 2008, did you ever do anything that was dishonest or illegal? Answered No). In the probable-lie approach the subject is told that the comparison questions are important because they reveal information about the subject's character that would be useful in assessing their credibility about the issue being assessed. In the directed-lie approach questions similar to those used in the probable-lie approach are used, but the subject is instructed to lie to the questions. Subjects are told that the directed-lies are important for the outcome of the test because they provide an index that the subject is responding appropriate when he or she lies, and that without that index of continued appropriate responding, the test outcome will be inconclusive. The present study manipulated the use of the probable-lie and directed-lie approaches in a 2 (Guilty, Innocent) X Question Type (Probable-Lie, Directed-Lie) factorial experiment. This study also manipulated the stimulation of questions between question repetitions so that the full experiment was a 2 X 2 X 2 fully factorial design.

Two hundred and fifty persons were tested by an experienced polygraph examiner and three research assistants trained at the laboratory to run examinations for this study. Objective Scoring System scores were analyzed as were decisions based upon those scores. The analyses found no evidence for significant differences between the validity of the probable-lie and directed-lie approaches to comparison questions. However, there was a significant effect of between repetition stimulation of the questions on decisions. Analyses revealed that the effect of stimulation was due to an increased number of true positive outcomes when stimulation was employed and a higher number of false positive errors when stimulation was not employed.

Although there were no significant effects of the approach to comparison questions, there is much to recommend the directed-lie approach:

- The directed-lie approach is simpler and far more standardized than is the probable-lie approach. It is easy to teach and to use.

- A very small number of directed-lie questions can be used for essentially all examinations.
- Examiner judgments about which probable-lies balance the relevant issues are eliminated thus improving reliability in test administration.
- Examiner skill and experience required for properly presenting probable-lie comparison questions is not needed for presenting directed-lie which can be presented by script. This too should improve reliability in test administration.
- Assumptions and examiner judgments about what is and is not a probable-lie for individual subjects are eliminated. This should further improve the reliability of test administration.
- The directed-lie does not intrude into the subjects private life, nor does it go beyond the parameters of a forensic investigation. It should thus be perceived as less intrusive and objectionable, even by sensitive subjects.
- The directed-lie has face validity for subjects and for lay persons. It should be much easier to explain to subjects, policy makers, and legal professionals.

These positive factors suggest that the directed-lie should be considered for wider application in field settings. The significant positive effect for between repetition stimulation recommends this practice for adoption in the field.

Introduction

Comparison Question Tests (CQT) are the most commonly used type of psychophysiological deception detection (PDD) test in law enforcement, forensic practice, and national security screening settings (Honts, Raskin & Kircher, 2008; Raskin & Honts, 2002; Vrij, 2008). Such tests play an important role in the United States Government's national security and law enforcement programs. World-wide, the interest in and use of PDD is growing rapidly as was evidenced by the recent European Meeting on Polygraph Testing (Merckelbach & van Koppen, 2006). However, many aspects of the polygraph testing procedure, as it is used in practice, lack strong empirical validation and, in some cases lack any empirical validation. Two aspects of comparison test administration are currently the topic of controversy in polygraph literature. Those areas of controversy concern the type of comparison question used and between chart stimulation of questions.

Types Of Comparison Questions.

Comparison questions are designed to provide the innocent suspect an opportunity to become more concerned about questions other than the relevant questions, thereby causing the innocent suspect to react more strongly to the comparison questions than to the relevant questions. There are two types of comparison questions currently used in field practice. The more common and older form of the comparison question is the probable-lie. Probable-lie comparison questions deal with acts that are similar to the issue of the investigation. However, they are more general in nature, deliberately vague, and cover long periods of time in the life history of the subject. Virtually every subject has difficulty in unequivocally answering them with a simple and truthful "No." An example of a probable-lie question in an examination regarding a robbery is "Prior to 2008, did you ever take something that did not belong to you?" Probable-lie comparison questions are reviewed with the subject after the relevant questions are discussed and reviewed, and they are presented in a manner designed to encourage the subject to answer them with a denial.

A newer form of the comparison question is the directed-lie. With directed-lie comparison questions the subject is instructed to answer certain questions with an obvious lie. A typical directed-lie question is "Prior to 2008, did you ever tell even one lie?" All subjects are told that they must show appropriate responses when lying to the directed-lie questions, or the test will result in an inconclusive outcome. The rationale for using directed-lie comparison questions is similar to the rationale for probable-lie comparison questions. It is assumed that the subject's concern will be focused on the questions that pose the greatest risk of an undesirable test outcome. For guilty subjects, the focus will be on the relevant questions that are answered deceptively. Thus, guilty subjects should show stronger reactions to the relevant questions as compared to comparison questions. It is reasoned that innocent subjects will focus on showing they are suitable subjects, and on clearly demonstrating that their reactions when lying are dif-

ferent from when they are truthful. This focus of concern is designed to enhance the reactions of truthful subjects to the directed-lie questions, making them stronger than the relevant questions. Thus, subjects who are truthful in response to the relevant questions should be most concerned about their reactions to the directed-lie questions and should give larger physiological responses to the directed-lie comparison questions as a result.

The probable-lie version of the comparison question test (CQT) has several inherent problems, and some suggest that the directed-lie is a remedy for most of these problems (Fuse, 1982; Honts, 1994; Honts & Raskin, 1988; Honts et al., 2008; Horowitz, Kircher, Honts, & Raskin, 1997). Probable-lie comparison questions can be difficult to administer in field settings and require psychological sensitivity, sophistication, and skill on the part of the examiner to obtain an accurate outcome (see the review by Raskin & Honts, 2002). Unfortunately, many polygraph examiners lack adequate training in psychological methods and do not understand the basic concepts and requirements of using a standardized psychological test in a field setting. These problems are exacerbated when the examiner formulates and introduces probable-lie comparison questions to the subject, because it is difficult to standardize the wording and discussion of probable-lie comparison questions across different field settings. Clearly, the validity of a probable-lie comparison question test depends on how the subject perceives and responds to the probable-lie questions when they are introduced and discussed during the pretest interview.

The difficulties with probable-lie comparison questions may be compounded by problems related to the characteristics of examinees (Raskin & Honts, 2002). Examinees can be very anxious about the subject matter of the probable-lie questions, making it difficult for the examiner to establish effective comparison questions. These questions may be personally intrusive and offensive to some subjects. For others, the probable-lie questions may encompass prior criminal behavior of a serious nature that poses problems for the subjects, some of whom may refuse to answer the questions. If a person is administered more than one test or tested on multiple occasions, it may become difficult to formulate new probable-lie questions that continue to be effective for the subject. Moreover, it is difficult to explain the function of probable-lie questions and their role in interpreting the outcome of the test to those who use the results of polygraph tests (e.g., investigators, lawyers, judges, and juries) and to laypersons. They often do not understand the rationale of the probable-lie and may interpret strong physiological reactions to probable-lie questions as an indication that the subject is dishonest and guilty. For all of these reasons, the directed-lie test was developed, and, on its face, appears to be a preferable approach.

The stimulation of comparison questions between charts. In a typical CQT, the question series is repeated between three and five times. These repetitions are usually referred to as charts, in reference to the time when all polygraph data were collected on paper charts. One area of marked divergence in field practice concerns what is said to subjects between those question repetitions. The Department of Defense

approach is to not discuss (stimulate) any of the test questions between question repetitions. The University of Utah approach is to discuss both the relevant and comparison questions between charts. In the Utah system, after each presentation of the question sequence, the examiner asks the subject if there were any problems and discusses any concerns that the subject expressed. The examiner then reviews the relevant and comparison questions in order to ensure that the relevant questions are clear and straightforward and the comparison questions remain salient. If the subject makes an admission to a probable-lie question or provides additional information that changes the meaning of a relevant question, this is discussed and appropriate adjustments are made in the affected questions.

Both the type of comparison question and the stimulation of questions between repetitions represent important divergences in field practice. Questions about the effects of these practice differences represent problems that currently cannot be answered definitively with empirical evidence. The meta-analysis by Honts (1999) suggests that between chart stimulation offers a positive improvement in CQT accuracy. Moreover a recent study by Offe and Offe (2007) reported results that between chart discussion had a positive effect on accuracy, when there was minimal explanation of comparison question in the pretest. When there was a normal pretest discussion of the comparison questions between chart stimulation produced non-significantly higher accuracy. However, the Offe and Offe study had relatively few subjects and thus had relatively low statistical power to find small effects.

The literature contrasting the probable-lie and the directed-lie is more equivocal, but even if the accuracy rates associated with the probable-lie and the directed-lie are roughly equivalent, the directed-lie comparison question offers substantial advantages in standardization, face validity to lay audiences, and decreased intrusiveness. Resolution of these two questions through a well-conducted experiment would provide a substantial increase in our knowledge about the best practices to take in the field. Should the evidence support between chart stimulation and the directed-lie comparison question, those techniques could be added to field practice quickly and with minimal cost in retraining. Finally, it may be that the stimulation of questions between repetitions has differential effects on probable-lie and directed-lie comparison questions. Thus it makes the most sense to study these two variables in a factorial design where their possible interaction can be examined directly.

Relation To Personnel Security Issues

The research described here was designed to address two areas where current field practices diverge. The current U. S. Government standard is to use probable-lie comparison questions (except for certain screening tests, notably the Test for Espionage and Sabotage) and to not stimulate questions between charts. However, the current scientific literature provides some support for the notions that the stimulation of questions between repetitions and (to a lesser extent) the use of directed-lie comparison questions may increase CQT accuracy (for a summary see Raskin & Honts, 2002). If either of these inno-

vations offers even a modest improvement (of even 1 or 2 percent) in overall CQT accuracy this could have a major impact given the large number of individuals administered polygraph tests each year by law enforcement and in national security settings. The present study was designed to provide statistically powerful evidence regarding these potentially important variables.

Literature Review

The present first author has recently published two major co-authored reviews of CQT accuracy and practice (Honts et al., 2008; Raskin & Honts, 2002). Both publications provide extensive reviews of the scientific literature concerning the administration and validity of the two types of comparison questions, as well as the scientific literature concerning the stimulation of questions between repetitions. The following review closely follows Raskin & Honts (2002).

Validity Of The Probable-lie Approach

The validity of comparison question polygraph tests is the subject of intense debate among scientists (for example see the most recent exchange between: Honts et al., 2008; Iacono & Lykken, 2008). Although the majority of psychophysiologists and psychologists who work in the area of Psychology and Law express generally positive attitudes concerning the usefulness of polygraph tests for assessment of credibility (Amato & Honts, 1994; Gallup, 1984; Honts, Thurber, Cvencek, & Alloway, 2002), the American Psychological Association expressed serious concerns about their scientific basis and some specific applications (see Raskin, 1986, p 73). In the last 30 years, there has been a great deal of research, development, and experience with various techniques that employ physiological measures for assessing credibility regarding specific facts, events, or knowledge (Honts et al., 2008).

The debate about the accuracy of comparison question centers on two general sources of data from which the accuracy of such tests are estimated. Data are obtained either from laboratory simulations or studies of actual cases that include testing of one or more suspects in a criminal investigation. There are advantages and disadvantages to each type of study, and both types are needed to provide an overall picture of test accuracy.

Laboratory Studies. Laboratory research is an attractive alternative because the scientist can control the environment. Moreover, with regard to credibility assessment studies, the scientist can know with certainty who is telling the truth and who is lying by randomly assigning subjects to conditions. Laboratory research on credibility assessment typically makes subjects deceivers by having them commit a mock crime (e.g. "steal" a watch from an office), and then instructing them to lie about it during a subsequent test. From a scientific viewpoint, random assignment to conditions is highly desirable because it controls the influence of extraneous variables that might confound the results of the experiment (Shadish, Cook &

Campbell, 2002). The most accepted type of laboratory study realistically simulates a crime in which some subjects commit an overt transgression, such as a theft (Kircher, Horowitz, & Raskin, 1988). While the guilty subjects enact a realistic crime, the innocent subjects are merely told about the nature of the crime but do not participate in it. All subjects are motivated to produce a truthful outcome, usually by a cash bonus for passing the test. For example, one such study used prison inmates who were offered a bonus equal to one month's wages if they could produce a truthful outcome (Raskin & Hare, 1978).

The advantages of careful laboratory simulations include total control over the issues that are investigated and the types of tests that are used, consistency in test administration and interpretation, specification of the subject populations that are studied, experimental control over the skill and training of the examiners, and absolute verification of the accuracy of test results. Carefully designed and conducted studies that closely approximate the methods and conditions characteristic of high quality practice by polygraph professionals and that use subjects similar to the target population, such as convicted felons or a cross-section of the general community, provide the most generalizable results (Kircher et al., 1988). Some laboratory research, including some credibility assessment studies, can be criticized for a lack of realism. This lack of realism may (but not necessarily) limit the ability of the scientist to apply the results of the laboratory to real-world settings. However, a recent study reported by Anderson, Lindsay, and Bushman (1999) examined a broad range of laboratory-based psychological research. They concluded the following, "Correspondence between lab- and field-based effect sizes of conceptually similar independent and dependent variables was considerable. In brief, the psychological laboratory has generally produced truths, rather than trivialities." (p. 3). It thus seems reasonable to conclude that high quality studies of the CQT are similar to other areas of psychological research and that those studies produce important information about the validity of such tests and not trivial information as some of the critics have claimed (e.g., Iacono & Lykken, 2008). Moreover, the majority of scientists in the Society for Psychophysiological Research and the American Psychology Law Society opine that judges and other policy makers should give weight to the results of laboratory studies of the polygraph (Honts et al., 2008.) A Committee of Concerned Social Scientists filed a Brief for Amicus Curiae (Honts & Peterson, 1997) with the Supreme Court of the United States in the case of United States v. Scheffer (1998). They found eight high quality laboratory studies of probable-lie versions of the CQT. Honts and Schweinle (2009) updated this work including three more recent studies that meet the criteria. The results of those laboratory studies are illustrated in Table 1. The high quality laboratory studies indicate that the CQT is a very accurate discriminator of truth tellers and deceivers. Over the studies, when inconclusive outcomes are ignored, the CQT correctly classified about 91% of the subjects and produced approximately equal numbers of false positive and false negative errors.

Table 1. Outcome For High Quality Laboratory Studies

	Guilty				Innocent			
Study	n	% Correct	% Wrong	% Inc	n	% Correct	% Wrong	% Inc
Driscoll et al. (1987)[b]	20	90	0	10	20	90	0	10
Ginton et al. (1984)	2	100	0	0	13	85	15	0
Honts et al. (1994)[a]	20	70	20	10	20	75	10	15
Honts et al. (2004)[b]	24	92	0	8	24	92	8	0
Horowitz et al. (1997)[c]	15	53	20	27	15	80	13	7
Kircher & Raskin (1988)	50	88	6	6	50	86	6	8
Offe & Offe (2007)[d]	18	89	11	0	15	93	7	0
Patrick & Iacono (1989)	24	92	8	0	24	64	36	0
Podlesny & Raskin (1978)	20	70	15	15	20	90	5	5
Podlesny & Truslow (1993)	72	69	13	18	24	75	4	21
Raskin & Hare (1978)	24	88	0	12	24	88	8	4
Rovner (1986)[a]	24	88	0	12	24	88	8	4
Means	26.08	82.42	7.75	9.83	22.75	83.83	10.00	6.17

aCountermeasure subjects excluded.
[b]Conditions representing standard field practice.
[c]Traditional comparison question subjects only.
[d]Standard pretest subjects only.

Field studies. The major disadvantage of laboratory simulations is the difficulty of completely simulating the real-life situation in which a person suspected of a crime is administered a polygraph test. To verify test accuracy under field conditions, it is necessary to use tests conducted on actual criminal suspects. However, field studies of criminal suspects also have inherent problems. The major problem is obtaining verification of the accuracy of the test outcomes, which can be very difficult in real cases. Two

general methods have been used to develop a criterion of guilt and innocence against which to assess the accuracy of field polygraph tests (Raskin, 1986). The best and most common method utilizes confessions to verify the guilt and innocence of the examinees. Law enforcement cases that involve polygraph tests produce rates of confessions in the range of 30% to 80%, but it is not known how these cases compare to those that did not produce confessions (Raskin, 1986).

The other method of case verification relies on a panel of legal experts who review the case facts to provide judgments concerning the guilt or innocence of individual suspects. The use of panel decisions produces different problems than a criterion based on confessions (Raskin, 1986). Since the panel criterion for guilt and innocence is an educated guess, it is subject to unreliability and error. However, Honts (1996) examined the reliability of a panel decision as compared to the a confession criterion and found that such panel decisions could be reliable, and were not significantly different from cases confirmed by a confession criterion. Additional research is needed to refine the panel criterion approach. The other major problems with field studies concern the representativeness of the cases selected, the training and skill of the polygraph examiners who conducted the tests, and the adequacy of the test methods and diagnostic procedures employed. To estimate the accuracy of polygraph tests on criminal suspects, it is necessary to select cases in which the subjects were suspects, not victims or witnesses. Although it is generally recognized that polygraph tests are most likely to produce false positive errors on victims of serious crimes (Raskin, 1986), at least one major field study (Horvath, 1977) used a large number of tests in which verified innocent victims had been tested (see Raskin, 1986).

It is important that field studies select cases according to scientifically acceptable sampling procedures, using only cases in which properly trained polygraph examiners employed standard field methods for conducting the tests and interpreting their outcomes. Some of the frequently cited studies (e.g., Kleinmuntz & Szucko, 1984) failed to adhere to these principles. As a result, such studies provide limited information concerning the accuracy of properly conducted and interpreted polygraph tests. The Kleinmuntz and Szucko study stands out because it embodied all the serious methodological errors. They used only cases in which persons suspected of theft were ordered by their employers to take tests from a commercial polygraph firm, and they did not describe how they selected cases from the files of the commercial polygraph firm. In addition, they based the results on interpretations made by students in a commercial polygraph training course who were not trained in systematic methods of test interpretation, and they required the student examiners to make definite judgments of guilt or innocence on the basis of reactions to a single relevant question. It is not surprising that their study produced low rates of accuracy. Gross violations of acceptable scientific methodology and polygraph procedures render that study totally meaningless for estimating the accuracy of standard field polygraph examinations conducted by competent examiners under appropriate conditions.

In summary, it now seems to be generally agreed by persons doing field research in this area (Honts et al., 2008) that useful field studies of the psychophysiological credibility assessment tests should have all of the following characteristics:

- Subjects should be sampled from the actual population of subjects in which the researcher is interested. If the researcher wants to make inferences about tests conducted on criminal suspects, then criminal suspects should be the subjects who are studied.
- Subjects should be sampled by some random process. Cases must be accepted into the study without reference to either the accuracy of the original outcome or to the quality of the physiological recordings.
- The resulting physiological data must be evaluated by persons trained and experienced in the field scoring techniques about which inferential statements are to be made. Independent evaluations by persons who have access to only the physiological data are useful for evaluating the information content of those data. However, the decisions rendered by the original examiners probably provide a better estimate of the accuracy of polygraph techniques as they are actually employed in the field.
- The credibility of the subject must be determined by information that is independent of the specific test. Confessions documented by physical evidence are presently the best criterion available.

Unfortunately, there are few field studies from which we can estimate the accuracy of properly conducted comparison question tests. In 1983, the Office of Technology Assessment of the United States Congress selected 10 field studies that it felt had at least some degree of scientific merit. The overall accuracy of the polygraph decisions was 90% on criterion-guilty suspects and 80% on criterion-innocent suspects. In spite of the inclusion of many studies with serious methodological problems, accuracy in field cases was higher than is claimed by some of the most vocal critics (Iacono & Lykken, 2008).

A more recent survey of the available field studies was performed by the Committee of Concerned Social Scientists (Honts & Peterson, 1997). Four field studies were found that met the criteria, listed above, for meaningful field studies of psychophysiological credibility assessment tests. The results of the independent evaluations for those studies are illustrated in Table 2. Overall, the independent evaluations of the field studies produce results that are quite similar to the results of the high quality laboratory studies. The average accuracy of field decisions for the CQT was 90.5 percent. However, with the field studies nearly all of the errors made by the CQT were false positive errors.

Table 2. Outcomes Of High Quality Field Studies.

	Guilty				Innocent			
Study	n	% Correct	% Wrong	% Inc	n	% Correct	% Wrong	% Inc
Honts (1996)[a]	7	100	0	0	6	83	0	17
Honts & Raskin (1988)[b]	12	92	0	8	13	62	15	23
Mangan et al., (2008)	15	100	0	0	15	93	7	0
Patrick & Iacono (1991)[c]	52	92	2	6	37	30	24	46
Raskin et al. (1988)[d]	37	73	0	27	26	61	8	31
Means	24.6	91.4	0.4	8.2	19.4	65.8	10.8	23.4

aSub-group of subjects confirmed by confession and evidence.
[b]Decision based only on comparisons to traditional comparison questions.
[c]Results from the mean blind rescoring of the cases "verified with maximum certainty" (p.235)
[d]These results are from an independent evaluation of the "pure verification" cases.

The scientific data concerning the validity of the polygraph can be summarized as follows: High quality scientific research from the laboratory and the field converge on the conclusion that, when properly conducted, the CQT is a highly accurate discriminator of truth tellers and deceivers. The research results converge on an accuracy estimate that exceeds 90 percent. Moreover, original examiners, who are most likely to offer testimony, produce even higher estimates of accuracy. There may be a tendency for the CQT to produce more false positive than false negative errors, but this trend in the current literature is not particularly strong. Moreover, no tendency toward false positive errors is seen in the decisions of the original examiners.

Validity Of The Directed-lie Approach

Since the directed-lie is relatively new, there are fewer studies of its validity. As with the probable-lie there are some laboratory validity (Department of Defense Polygraph Institute [DoDPI], 1995; 1997; 1998; Honts & Alloway, 2007; Horowitz et al., 1997; Reed, 1994) and one field validity (Honts & Raskin, 1988) studies. The Horowitz et al. study used a mock crime that closely approximated the field situation. Horowitz et al. compared the effectiveness of the directed-lie with the probable-lie and a relevant/irrelevant (RI) approach. Different groups received one of two types of directed-lies: the personally relevant

directed-lies using the procedures described, or simple directed-lies to three of the trivial neutral questions that were used in the RI. The results of the Horowitz et al. (1997) study indicate that compared to the other three conditions, the personal directed-lie produced the highest accuracy, except for the RI with guilty subjects. The outcomes for the four types of tests are presented in Table 3. Among all question structures, the personal directed-lie produced the highest number of correct decisions on innocent subjects and among the three tests that employed comparison questions, it produced the highest number of correct decisions on guilty subjects.

Table 3. Test Outcomes On The Horowitz Et Al. (1997) Study.

	Test Outcomes (%)			
Experimental Groups	Correct	Wrong	Inconclusive	% Correct Decisions
Guilty				
Relevant-irrelevant	100	0	0	100
Trivial Directed-Lie	53	20	27	73
Personal Directed-Lie	73	14	13	84
Probable-Lie Comparison	53	20	27	73
Innocent				
Relevant-irrelevant	20	73	7	22
Trivial Directed-Lie	67	13	20	84
Personal Directed-Lie	87	13	0	87
Probable-Lie Comparison	80	13	7	86

n = 15 for each of the experimental groups.
The percentage of correct decisions was calculated by excluding inconclusive outcomes.

The U. S. Department of Defense reported three sets of studies concerning the validity of the directed-lie comparison question in national security settings. Barland (1981) examined the validity of the Military Intelligence version of the directed-lie in a mock screening setting with 26 truthful subjects and 30 subjects who attempted deception. All subjects were tested with the directed-lie comparison; no other techniques were examined. Excluding inconclusive outcomes, Barland's evaluators correctly classified 79% of the subjects. Although this might be considered modest performance in comparison to that obtained in the University of Utah studies, it must be remembered that the Barland (1981) study was a screening study. When compared with other mock-screening studies, which have often produced near chance performance with probable-lie tests (e.g. Barland, Honts, & Barger, 1989; Honts, 1992), the performance of the directed-lie in Barland (1981) was actually quite strong. Abrams (1991) reported the only other laboratory

study of the directed-lie. Unfortunately that study was so poorly designed and methodologically flawed that the data it generated are meaningless, by the author's own admission (see the Abrams quotes in Honts & Gordon, 1998, p.248; and in Honts, Raskin, Amato, Gordon, & Devitt, 2000, p.158.) Abrams and Matte (Abrams, 1999; Matte, 1998; 2000; Matte & Reuss, 1999) are outspoken critics of the directed-lie approach, but their criticisms are easily shown to be without merit and essentially all of their attacks are baseless. Interested readers are referred to the research and commentary by Honts and his colleagues (Honts, 1999; 2000; Honts & Gordon, 1998; Honts et al., 2000).

The other studies on the directed-lie concern a newer test, the Test of Espionage and Sabotage (TES) developed by DODPI for use in national security screening tests. Reed (1994; also published as DODPI Research Staff, 1997), reported three laboratory mock screening studies. Following a series of studies that indicated that the national security screening tests of the time were making an unacceptably high number of false negative errors (Barland et al., 1989; Honts, 1991; 1992; 1994) the DODPI attempted to develop a more accurate screening test. It should be noted that the primary concern in conducting national security screening tests is a desire to avoid false negative errors. Following a series of studies that are not publicly available, Reed described the product of the DODPI's efforts. In the first study reported in Reed, the TES, a test format with only directed-lie comparison questions, was tested against two versions of the Counter-intelligence Scope Polygraph (CSP) test. One version of the CSP used probable-lie comparison questions while the other used directed-lie comparison questions. The TES outperformed both of the CSP formats in terms of correctly identifying guilty subjects. The CSP with directed-lie comparisons was slightly, but not significantly, better at identifying guilty subjects than was the CSP with probable-lie comparisons. The second Reed study reported even higher accuracy for the TES, a directed-lie comparison test format. Little information is provided about the third study, but it also appears to show considerable accuracy for the directed-lie based TES.

DODPI Research Staff (1998) reported a mock espionage/sabotage study that involved 82 subjects. All subjects were tested with the TES. Excluding one inconclusive outcome, the examiners correctly identified 98% of the innocent subjects and 83.3% of the guilty subjects. This study also indicates that the directed-lie comparison-based TES is extremely successful in discriminating between innocent and guilty subjects. Honts and Alloway (2007) report a study of the susceptibility of the TES to information as a countermeasure. They failed to find any effects of exposing subjects to Masche and Scalabrini's (2000) online book about beating the lie detector. Honts and Alloway did not report a direct comparison of directed-lie to probable-lie. They reported 72.5% overall accuracy for their subject, but noted that the study was not designed to be an estimate of the field validity of the TES, rather it was designed as a test of the effectiveness of information as a countermeasure.

To date, Honts and Raskin (1988) have reported the only field study of the DLT. They conducted polygraph tests of criminal suspects over a four-year period and obtained 25 confirmed tests in which one personal directed-lie was included along with probable-lie comparison questions. Each of the investigators then performed blind interpretations of the charts obtained by the other investigator, scoring them with and without the use of the directed-lie question. The results of the Honts and Raskin study indicated that inclusion of the directed-lie question in the numerical evaluation of the charts had a noticeable effect on the confirmed innocent suspects, reducing the false positive rate from 20% to 0%. For the confirmed guilty suspects, it had the slight effect of changing one inconclusive outcome to a false negative. The effects of the directed-lie question on the numerical scores were more dramatic. Inclusion of the directed-lie comparisons almost doubled the size of the total numerical scores for the confirmed innocent suspects, raising the mean score from +4.7 to +9.0. It had a lesser effect on the scores of the confirmed guilty suspects, lowering them from -13.8 to -11.5. Thus, the directed-lie question had the effect of raising the mean score for innocent suspects from the inconclusive range into the definite truthful area, while leaving the mean score for guilty suspects clearly in the deceptive area. The main impact of the directed-lie question was a reduction in false positive errors.

Matte & Reuss (1999) and Matte (1998; 2000) have claimed that the directed-lie approach lacks construct validity. However, analyses by Honts and his colleagues (Honts, 2000; Honts & Alloway, 2007; Honts & Gordon, 1998; Honts et al., 2000) have clearly demonstrated that Matte's conceptual arguments are without merit. Moreover, the only empirical data to support their position was based on a simulation study where individuals were asked to imagine that they were taking a polygraph examination. Such data must clearly be questioned on the basis of external, face and construct validity, especially in the face of the positive validity data from subjects actually given polygraph examinations.

In summary, the results from the laboratory and the field are consistent with the proposition that the directed-lie test represents substantial conceptual and practical advantages over the probable-lie comparison question test. It is more standardized in its structure; it is easier to administer; it requires less manipulation of the subject and creates fewer problems for the subject; it is more readily explained to layperson, lawyers, judges, and juries. Most importantly, some evidence suggests that the directed-lie produces fewer errors as compared with probable-lie tests.

Between Repetition Stimulation Of Test Questions

As was noted above, one area of marked divergence in field practice concerns what is said to subjects between question repetitions (charts.) The Department of Defense approach is to not discuss (stimulate) any of the test questions between question repetitions. The University of Utah approach is to discuss both the relevant and comparison questions between charts. In the Utah system, after each presentation of the question sequence, the examiner asks the subject if there were any problems and discusses any concerns

that the subject expressed. The examiner then reviews the relevant and comparison questions in order to ensure that the relevant questions are clear and straightforward and the comparison questions remain salient. If the subject makes an admission to a probable-lie question or provides additional information that changes the meaning of a relevant question, this is discussed and appropriate adjustments are made in the affected questions.

Abrams (1999) and Matte (2000) claim that the between-charts discussion and review of questions places undue emphasis on the comparison questions and increases the risk of a false negative error. However, Honts (1999) analyzed data from 19 studies that involved 1092 polygraph tests. The results of the Honts analysis strongly suggest that between-charts discussion, even when limited only to the comparison questions, decreases the risk of error (see additional discussion of this issue in: Honts, 2000 and Honts et al., 2000.) However, none of the studies cited by Honts and his colleagues was designed to directly test the question of the effects of between repetition question stimulation. The present study directly addressed this issue in an experimental design.

Research Questioned Addressed in the Project

- Are there significant differences in accuracy between CQT tests conducted with probable-lie comparison questions or directed-lie comparison questions?
- Does a University of Utah style stimulation of questions between charts significantly effect CQT test accuracy?
- Does a University of Utah style stimulation of questions between charts interact with the type of comparison questions used in the CQT?

Method

Participants. Two hundred and fifty individuals (126 female, 124 male) were recruited via help-wanted ads on craigslist.com (see Appendix A), which stipulated an hourly wage of $15 for approximately 2 1/2 hours of participation in a polygraph research study. Individuals who were currently pregnant, taking prescription medication for high blood pressure, a heart condition, or to treat a psychological disorder, or had previously taken a polygraph examination, were deemed ineligible for participation in the study (see Appendix B). Those who met the selection criteria were randomly assigned to one of eight experimental conditions. Participants ranged in age from 18 to 65 (Mode = 20, M = 30, SD = 10.5).

Examiners. An experienced polygraph examiner (the principal investigator, 32 years of field polygraph experience at project onset) used reference materials provided by the Department of Defense Polygraph Institute (now DACA) to train three individuals, none of whom was a practicing polygraph

examiner, to conduct polygraph examinations. Two of the examiners were undergraduate research assistants, the third was a recent graduate with a B.A. in psychology and who worked as the paid Research Assistant for the project. The latter examiner had run polygraph examinations as part of a previous research project in our laboratory. The goal of the training was that the examinations should follow field procedures as closely as possible. The polygraph examiner and the assistants who greeted the participants were unaware, at all times, of the participants' guilt or innocence. The experienced polygraph examiner conducted 92 examination in the project. The Research Assistant conducted 84 examination. The female undergraduate examiner conducted 38 examinations while the male undergraduate examiner conducted 35 examinations.

Apparatus. Physiological data were collected with CPSII field polygraph instruments. The following physiological responses were monitored: Thoracic and abdominal respiration were monitored with Pneumotrace strain sensors placed around the chest and abdomen; electrodermal response was measured from disposable Vermed GSR-13 electrodes placed on the plam in the area of the thenar and hypothenar eminences; relative blood pressure was monitored from a cuff placed on the subjects upper left arm; and peripheral blood flow was monitored with a photoelectric plethysmograph placed on the distal surface of the subject's right thumb. A Stoelting movement sensor was placed in the seat of the subject's chair. Instrumentation filtering and sampling was modeled after field instrumentation procedures as closely as possible.

Design. The design of the study was a 2 (Guilty, Innocent) X 2 (Probable-Lie, Directed-Lie) X 2 (Between Chart, Not) between subjects factorial. Subjects were randomly assigned to eight conditions with the constraint that each condition would be considered to be complete when 24 subjects had been run in that condition.

Procedures. The design was implemented using of a variation of the mock crime paradigm developed at the University of Utah (e.g., Podlesny & Raskin, 1978). Upon arriving at the Applied Cognition Research Institute, participants were directed to a room in which they privately watched a video (the script of which was also presented in typewritten form; see Appendix C). This script/video described that their participation in the study may involve stealing some money and that they, regardless of their assigned condition, would be taking a polygraph examination during which they were to try to convince a polygraph examiner that they were giving truthful responses to the questions. If they agreed to the described conditions of the study, participants signed an informed consent sheet (see Appendix D). After their consent was obtained, participants selected an unmarked sealed envelope from a box of unmarked envelopes. That envelope contained instructions for watching another video that would describe their condition assignment and instructions for carrying out their task(s).

Some participants (Innocent) were shown a video informing them that they were assigned to the innocent condition and thus they were not going to be stealing any money during the study. These participants were told that they would be paid a $20 bonus if they successfully convinced the polygraph examiner that they were innocent of stealing $20 from the Education Building (see Appendix E). These participants were instructed to leave the laboratory building and go to the Education Building (that houses the Psychology Department), where they were to deliver an envelope to the door of Dr. Anooshian's office and return to the laboratory 20 minutes later to take a polygraph examination.

Other participants watched a video informing them that they were assigned to the guilty condition and thus they were going to be stealing money during the study. These participants were also informed that if they were successful in passing the polygraph examination, by producing a truthful outcome concerning the theft of $20 from the Education Building, they would be paid a $20 bonus (see Appendix F). These participants were instructed to leave the laboratory building and go to the Education Building. They were asked to find Dr. Anooshian's office and steal an envelope addressed to Sam Stone that was taped to the door. They were then asked to open the envelope and hide its contents (a $20 bill) on their person. They were asked to return to the laboratory 20 minutes later to take a polygraph examination.

Upon returning to the laboratory, an assistant introduced the participants to the polygraph examiner. Participants were reminded by the examiner that their polygraph examination would be videotaped and that the purpose of the examination was to identify the person who had stolen an envelope containing $20 from the door of Dr. Anooshian's office in the Education Building earlier that day. Examination sessions began with the examiner collecting some general information from the participant concerning things such as the participant's general health, how well they had slept the night before, whether he/she had ever taken a polygraph exam. This was done using the built-in biographical forms in the CPS II software. Participants were then told that they were a suspect in the theft of $20 from the Education Building and were asked if they had, in fact, stolen the envelope containing the money. After participants denied the accusation, the examiner asked them to explain where they had been and what they had been doing for approximately the last two hours.

At this point, the examiner briefly discussed the nature of the autonomic nervous system. In essence, that although individuals are largely able to control their motor behavior, many functions of the body, such as temperature regulation, heart rate, and breathing are largely uncontrollable and vary automatically in response to physical and psychological stressors, such as lying.

Next, the function of each sensor was described to participants, and participants were told to expect that, due to the pressure applied from the blood pressure cuff, they might experience a tingling sensation in and/or some discoloration of the arm on which the blood pressure cuff was placed. At this point, participants were asked to sign another informed consent sheet (see Appendix G). Next, participants were

told that a practice test was going to be conducted before the actual polygraph examination concerning the theft. The practice test was introduced under the guise of being necessary for establishing participants' unique physiological reactions to lying. Participants were told to pick a number between 2 and 6 and inform the examiner of the number that was chosen. It was explained that after the sensors were attached to the participant a series of questions would be posed, beginning with "Concerning the number that you chose, was it the number 1?" and continuing through to number 7. Participants were instructed to answer "no" to each of the seven questions, so that during the asking of the question regarding the number that was selected (and hence their deception was known) their unique physiological responses to lying could be identified.

Participants were asked to wash their hands (so that the best possible recordings from the sensors could be obtained). At this point, the sensors were attached, and the practice test was conducted. All participants were told that the polygraph revealed a highly distinct change in their physiological responses on the question to which they lied. Hence, the participant was a suitable subject for a polygraph examination. Next, each of the questions that would be asked during the polygraph examination concerning the theft of $20 was reviewed with the participants. As the examiner read each question, the participant was instructed to answer with a "yes" or "no" just as they would during the actual examination. All participants were asked 3 relevant questions, 3 control questions, 2 neutral questions and three other questions (see Appendix H). After all the questions were reviewed and responded to by the participants, a comparison question test was conducted according to standard procedures used by the U. S. Federal polygraph examiners.

After the examination was completed, participants received a thorough debriefing by the examiner, during which they were told about the outcome of their examination (i.e., whether their responses were scored as truthful or deceptive) and the various conditions that were being compared as part of the study. Finally, participants were thanked and paid for their participation.

The resulting physiological data were edited independently by an experienced polygraph examiner who was not informed about subject assignment to conditions. Following editing, the data were analyzed with the Objective Scoring System module that is part of the CPSII software.

Results

OSS total scores were analyzed with a 2 (Guilt) X 2 (Test Type) X 2 (Stimulation) ANOVA. Means for all of the Cells of that ANOVA are shown in Table 3. The ANOVA revealed only one significant effect, a

Table 3. Cell Means For The 2 X 2 X 2 Design.

Test Type	Stimulation	Guilt	Mean	Std. Deviation	N
Probable-Lie	Yes	Guilty	-24.30	27.749	30
		Innocent	16.03	25.350	31
		Total	-3.80	33.269	61
	No	Guilty	-26.91	20.847	32
		Innocent	14.33	26.914	30
		Total	-6.95	31.574	62
	Total	Guilty	-25.65	24.262	62
		Innocent	15.20	25.926	61
		Total	-5.39	32.331	123
Directed-Lie	Yes	Guilty	-21.30	28.444	30
		Innocent	20.06	25.259	34
		Total	.67	33.754	64
	No	Guilty	-21.27	25.266	33
		Innocent	12.45	25.727	29
		Total	-5.50	30.438	62
	Total	Guilty	-21.29	26.607	63
		Innocent	16.56	25.556	63
		Total	-2.37	32.186	126
Total	Yes	Guilty	-22.80	27.900	60
		Innocent	18.14	25.185	65
		Total	-1.51	33.458	125
	No	Guilty	-24.05	23.190	65
		Innocent	13.41	26.127	59
		Total	-6.23	30.893	124
	Total	Guilty	-23.45	25.462	125
		Innocent	15.89	25.643	124
		Total	-3.86	32.228	249

large main effect for Guilt, F(1, 241) = 143.82, p < .001. The complete ANOVA table for the analysis is provided in Table 4. Review of the effect sizes and significance levels associated with the non-significant effects reveals that none of them approached significance or accounted for any appreciable amount of variance in the data. Virtually all the systematic variance in these data was accounted for by the Guilt variable.

Table 4. Complete ANOVA Table For The Primary Analysis.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	98021.527[a]	7	14003.075	21.149	.000	.381
Intercept	3706.549	1	3706.549	5.598	.019	.023
Test Type	450.519	1	450.519	.680	.410	.003
Stimulation	548.437	1	548.437	.828	.364	.003
Guilt	95223.943	1	95223.943	143.821	.000	.374
Test Type X Stimulation	41.699	1	41.699	.063	.802	.000
Test Type X Guilt	163.544	1	163.544	.247	.620	.001
Stimulation X Guilt	175.780	1	175.780	.265	.607	.001
Test Type X Stimulation X Guilt	283.345	1	283.345	.428	.514	.002
Error	159566.553	241	662.102			
Total	261297.000	249				
Corrected Total	257588.080	248				

a. R Squared = .381 (Adjusted R Squared = .363)

To provide some prospective on the effect of the independent variables on decisions, the OSS scores were turned into decisions using the simple +/- 6 rule. That is, examinations with OSS total scores of +6 or greater were classified as truthful. Examinations with OSS total scores of -6 or less were classified as deceptive and examinations with totals scores between -6 and +6 were classified as inconclusive. The resulting decisions were coded as deceptive = -1, inconclusive = 0 and truthful = 1. This coding scheme retains the ordinal characteristics of the underlying interval scaling of the OSS values. This data vector was submitted to a a 2 (Guilt) X 2 (Test Type) X 2 (Stimulation) ANOVA. Although the scaling of these data may violate the assumptions of ANOVA they clearly are ordinal and one could argue that the coding method here is a simple transformation of the original interval scale that retains characteristics of an interval scale, albeit, a truncated one. In any event we wanted to provide the most powerful test possible of the effects of the independent variables and currently there is no non-parametric test of interactions. The resultant ANOVA table is provided here as Table 5. Similar to the ANOVA of the underlying OSS scores,

Table 5. ANOVA Table From The Analysis Of The Coded Decision Data.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	84.262[a]	7	12.037	19.123	.000	.356
Intercept	10.711	1	10.711	17.015	.000	.066
Test Type	.406	1	.406	.644	.423	.003
Stimulation	2.592	1	2.592	4.117	.044	.017
Guilty	78.303	1	78.303	124.393	.000	.340
Test Type X Stimulation	.003	1	.003	.005	.942	.000
Test Type X Guilt	.513	1	.513	.814	.368	.003
Stimulation X Guilt	1.146	1	1.146	1.821	.178	.007
Test Type X Stimulation X Guilt	.059	1	.059	.093	.761	.000
Error	152.334	242	.629			
Total	247.000	250				
Corrected Total	236.596	249				

a. R Squared = .356 (Adjusted R Squared = .338)

this analysis revealed a large main effect of Guilt, $F(1, 242) = 124.39$, $p < .001$. However, this analysis also revealed a significant, but small effect of Stimulation between charts, $F(1, 242) = 4.12$, $p = .04$. Exploring this effect revealed that there were more cases were classified as deceptive when no between chart stimulation was employed. We explored that effect by creating cross-tabulations of Stimulation X Decision layered by Guilt. The resulting cross tabulation is presented here as Table 6. Examination of Table 6 revels that the effect of stimulation is due to an increased number of true positive outcomes when stimulation was employed and a higher number of false positive errors when stimulation was not employed, *Kendal's tau-c* = -.173, $p = .038$, for the Innocent table, all tests with the Guilty table were not significant. These results indicate that between chart stimulation of the comparison questions produced a positive effect on decisions with Innocent subjects, but no effect on the accuracy of decisions with Guilty Subjects.

Table 6. Cross Tabulation Of OSS Decisions And Stimulation Between Charts Layered By Guilt.

Guilt				OSS Decision			
				Deceptive	Inconclusive	Truthful	Total
Guilty	Stimulation	Yes	Count	52	0	8	60
				86.7%	.0%	13.3%	100.0%
		No	Count	58	1	6	65
				89.2%	1.5%	9.2%	100.0%
		Total	Count	110	1	14	125
				88.0%	.8%	11.2%	100.0%
Innocent	Stimulation	Yes	Count	15	1	49	65
				23.1%	1.5%	75.4%	100.0%
		No	Count	24	1	35	60
				40.0%	1.7%	58.3%	100.0%
		Total	Count	39	2	84	125
				31.2%	1.6%	67.2%	100.0%

Examiners. To see if there was an effect of examiner experience on objective scores an Examiners (4) X Guilt (2) X Test Type (2) X Stimulation (2) ANOVA was conducted. That analysis failed to reveal any significant effects involving the Examiner variable. Moreover, none of the effects involving Examiners approached significance.

Countermeasures. After being tested subjects were debriefed concerning countermeasures use. Forty-eight percent of all subjects reported spontaneously attempting a countermeasure. One criticism often raised against the directed-lie approach is that because of its clear face validity it will invite more spontaneous countermeasures attempts from guilty subjects. Table 7 provides a breakdown of countermeasure attempts by Guilt and Test Type. Table 5 reveals no suggestion that the directed-lie approach invites more countermeasure attempts than the probable-lie. In fact, although not significant, *Chi-Square* (1) = 2.73, p = .09, the trend in the data is in the opposite direction. With probable-lie examinations 83.9% of the guilty subjects report countermeasure attempts while with the directed-lie approach 71.4% of the guilty subjects report countermeasure attempts.

Reported countermeasure use was then treated as an independent variable and was used to explore the effects of countermeasure use on the objective scores. A Countermeasure (2) X Guilty (2) X Test Type

(2) X Stimulation (2) ANOVA was performed on the objective scores. No effect involving Countermeasures reached significance.

Table 7. Frequency Of Countermeasure Attempts By Guilty And Test Type.

				Countermeasures		
Test Type				No	Yes	Total
Probable-lie	Guilt	Guilty	Count	10	52	62
				16.1%	83.9%	100.0%
		Innocent	Count	51	10	61
				83.6%	16.4%	100.0%
		Total	Count	61	62	123
				49.6%	50.4%	100.0%
Directed-Lie	Guilt	Guilty	Count	18	45	63
				28.6%	71.4%	100.0%
		Innocent	Count	51	13	64
				79.7%	20.3%	100.0%
		Total	Count	69	58	127
				54.3%	45.7%	100.0%

Discussion

Research Question 1: *Are There Significant Differences In Accuracy Between CQT Tests Conducted With Probable-lie Comparison Questions Or Directed-lie Comparison Questions?*

The research reported here failed to find any statistically significant differences between comparison question tests conducted with probable-lie and directed-lie comparison questions. Moreover, examination of Tables 4 and 5 indicate that the manipulation of comparison questions type (Test Type) no only was not statistically not significant, it accounted for almost no systematic variability in either the objective scores or decisions based upon them. Given the large number of subjects and the statistical power of the tests used, this is a strong indication that there are, in fact, no differences in the results of comparison question tests based upon the type of comparison question used.

Research Question 2: *Does A University Of Utah Style Stimulation Of Questions Between Charts Significantly Effect CQT Test Accuracy?*

The research reported here failed to find any significant effects of between chart stimulation in the objective scores, see Table 4. Moreover the Stimulation independent variable accounted for little variance in the design. However, there was a statistically significant effect of stimulation in the decision data, see Table 5. Further analyses indicated that this effect was expressed as a positive effect of increased accuracy of decisions with Innocent subjects, but no effect on the accuracy of decisions with Guilty Subjects. It should be noted that this effect was of relatively small magnitude, *partial eta squared* = .017, indicating that approximately 2% of the variance in the decisions was due to the effects of the Stimulation independent variable.

Research Question 3: *Does A University Of Utah Style Stimulation Of Questions Between Charts Interact With The Type Of Comparison Questions Used In The CQT?*

The research reported here failed to find any significant interactions between Stimulation and Test Type in either the objective scores or the decision (see Tables 4 and 5). Not only were none of the interactions statistically significant, but the effect sizes associated with the interaction effects were all very small, the largest *partial eta squared* being .007. Given the large number of subjects and the high statistical power of the tests, these results strongly suggest that comparison question type and between chart stimulation do not interact.

General Discussion

As we noted in the introduction of this report, in the circumstance where the directed-lie and the probable-lie produce equivalent results there are important reasons to prefer the directed-lie.

- The directed-lie approach is simpler and far more standardized than is the probable-lie approach. It is easy to teach and to use.
- A very small number of directed-lie questions can be used for essentially all examinations.
- Examiner judgments about which probable-lies balance the relevant issues are eliminated thus improving reliability in test administration.
- Examiner skill and experience required for properly presenting probable-lie comparison questions is not needed for presenting directed-lie which can be presented by script. This too should improve reliability in test administration.
- Assumptions and examiner judgments about what is and is not a probable-lie for individual subjects are eliminated. This should improve the reliability of test administration.
- The directed-lie does not intrude into the subjects private life, nor does it go beyond the parameters of a forensic investigation. It should thus be perceived as less intrusive and objectionable, even by sensitive subjects.
- The directed-lie has face validity for subjects and for lay persons. It should be much easier to explain to subjects, policy makers, and legal professionals.

The combination of these factors with a determination of no differences in accuracy between the directed-lie and the probable-lie approaches makes an argument for widespread adoption of the directed-lie on the basis of practical reasons and because of likely gains in the reliability of test administration. In this highly controlled laboratory situation reliability of test administration was experimentally controlled for both techniques. However, in the field this is not the case. There appears to he a high degree of variability in field practice due to varying examiner skill, experience and training. Much of that unreliability could be ameliorated by the widespread adoption of a simpler and more standardized approach like that of the directed-lie.

Despite concerns by critics of the directed-lie approach, no evidence was found in this study that the directed-lie approach was either more inviting to countermeasure use or that it would be easier to beat with countermeasures than the probable-lie. Moreover, spontaneous countermeasure in this study were not effective in producing significant effects, despite their widespread use. This is particularly telling in that this was a sample of subjects recruited primarily from an online resource (Craig's List). A number of

subjects reported learning about countermeasures from public media sources, including the INTERNET. With regard to spontaneous countermeasures these results are completely consistent with a 20 year literature on spontaneous countermeasure use (see the review by Honts & Amato, 2002) that indicates that spontaneous attempts of countermeasures are ineffective in altering test scoring. However, one current finding does go against the trend of the reported spontaneous countermeasure data. In this study only 20% of the Innocent subjects reported countermeasure use while in the most recent study in the literature (Honts, Amato & Gordon, 2004) 45.8% of the Innocent subjects reported spontaneous countermeasure attempts. The lack of countermeasure effects associated with the directed-lie adds additional support for widespread adoption.

References

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

Abrams, S. (1999). A response to Honts on the issue of the discussion of questions between charts. *Polygraph*, *28*, 223-229.

Amato, S. L., & Honts, C. R. (1994). What do psychophysiologists think about polygraph tests? A survey of the membership of SPR. *Psychophysiology, 31*, S22. (Abstract).

Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality?, *Current Directions In Psychological Science*, *8*, 3-9.

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

Barland, G. H., Honts, C. R., & Barger, S. D. (1989). *Studies of the Accuracy of Security Screening Polygraph Examinations*. Department of Defense Polygraph Institute, Fort McClellan, Alabama. Available online: http://truth.boisestate.edu/raredocuments/bhb.html

Driscoll, L. N., Honts, C. R., & Jones D. (1987). The validity of the positive control physiological detection of deception technique. *Journal of Police Science and Administration*, *15*, 46-50.

Department of Defense Polygraph Institute Research Staff (1995|1997). *Report DODPI94-R-0008: A comparison of psychophysiological detection of deception accuracy rates obtained using the Counterintelligence Scope Polygraph and the Test for Espionage and Sabotage question formats*. Available from the Defense Technical Information Center as report #ADA319333; also reprinted *Polygraph*, *26* (1997), pp. 79-106.

Department of Defense Polygraph Institute Research Division Staff (1998). Psychophysiological detection of deception accuracy rates obtained using the Test for Espionage and Sabotage (TES). *Polygraph*, *27*, 68-73.

Fuse, L. S. (1982). *Directed-Lie control testing technique*. Unpublished manuscript.

The Gallup Organization, (1984). Survey of the members of the Society for Psychophysiological Research concerning their opinions of polygraph test interpretations, *Polygraph*, *12*, 153-165.

Ginton, A., Netzer, D., Elaad, E., & Ben-Shakhar, G. (1982). A method for evaluating the use of the polygraph in a real-live situations. *Journal of Applied Psychology, 67*, 131-137.

Honts, C. R., (1991). The emperor's new clothes: Application of polygraph tests in the American workplace. *Forensic Reports*, *4*, 91-116.

Honts, C. R., (1992). Counterintelligence scope polygraph (CSP) test found to be a poor discriminator. *Forensic Reports*, *5*, 215-218.

Honts, C. R. (1994). The psychophysiological detection of deception. *Current Directions in Psychological Science*, *3*, 77-82.

Honts, C. R. (1996). Criterion development and validity of the control question test in field application. *The Journal of General Psychology*, *123*, 309-324.

Honts, C. R. (1999). The discussion of comparison questions between list repetitions (charts) is associated with increased test accuracy. *Polygraph*, *28*, 117-123.

Honts, C. R. (2000). A brief note on the misleading and the inaccurate: A rejoinder to Matte (2000) with critical comments on Matte and Reuss (1999). *Polygraph*, *29*, 321-325.

Honts, C. R. & Alloway, W. (2007). Information does not affect the validity of a comparison question test. *Legal And Criminological Psychology*, *12*, 311-312. (Available online in 2006)

Honts, C. R., & Amato, S. (2002). Countermeasures, in M. Kleiner (Ed.), *Handbook of polygraph testing*. London: Academic (251-264).

Honts, C. R., Amato, S., & Gordon, A. (2004). Effects of outside issues on the Control Question Test. The *Journal of General Psychology, 151*, 53-74.

Honts, C. R., & Gordon, A. (1998). A critical analysis of Matte's analysis of the directed-lie. *Polygraph*, *27*, 241-252.

Honts, C. R., & Peterson, C. F. (1997). *Brief of the Committee of Concerned Social Scientists as Amicus Curiae*. United States v. Scheffer, in the Supreme Court of the United States. Available from the first author.

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

Honts, C. R., Raskin, D. C., Amato, S. L., Gordon, A., & Devitt, M. K. (2000). The hybrid directed-lie test, the overemphasized comparison question, chimeras and other inventions: A rejoinder to Abrams (1999). *Polygraph*, *29*, 156-168.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. Journal of Applied Psychology, 79, 252-259.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (2008). Scientific Status: The case for polygraph tests. In, D. L. Faigman, M. J. Saks, J. Sanders, and E. Cheng (Eds.) *Modern scientific evidence: The law and science of expert testimony (Volume 5): 2008-2009 Edition*. Thompson West: Eagan, Minnesota.

Honts, C. R., & Schweinle, (2009). Information gain of psychophysiological detection of deception in forensic and screening settings. *Applied Psychophysiology and Biofeedback*. (Published online July 2009)

Honts, C. R., Thurber, S., Cvencek, D., & Alloway, W. (2002, March). *General acceptance of the polygraph by the scientific community: Two surveys of professional attitudes.* Paper presented at the American Psychology-Law Society biennial meeting, Austin, Texas.

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

Horvath, F. S., (1977). The effect of selected variables on interpretation of polygraph records. *Journal of Applied Psychology, 62*, 127-136.

Iacono, W., & Lykken, D. T., (2008). Scientific Status: The case against polygraph tests. In, D. L. Faigman, M. J. Saks, J. Sanders, and E. Cheng (Eds.) *Modern scientific evidence: The law and science of expert testimony (Volume 5): 2008-2009 Edition*. Thompson West: Eagan, Minnesota.

Kircher, J. C., Horowitz, S. W., & Raskin, D. C. (1988). Meta-analysis of mock crime studies of the control question polygraph technique. *Law and Human Behavior*, *12*, 79-90.

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.

Kleinmuntz, B., & Szucko, J. (1984). A field study of the fallibility of polygraphic lie detection. *Nature*, *308*, 449-450.

Mangan, D. J., Armitage, T. E., & Adams, G. C., (2008). A field study on the validity of the Quadri-Track Zone Comparison Technique. *Physiology & Behavior, 95*, 17-23.

Maschke, G. W., & Scalabrini, G. J. (2000). *The lie behind the lie detector*. Internet: AntiPolygraph.org, Retrieved 2001 from http://antipolygraph.org

Matte, J. A. (1998). An analysis of the psychodynamics of the directed-lie control questions in the control question technique. *Polygraph*, *27*, 56-67.

Matte, J. A. (2000). A critical analysis of Honts' study: The discussion (stimulation) of Comparison Questions. *Polygraph*, *29*, 146-149.

Matte, J. A., & Reuss, R. M. (1999). Validation of potential response elements in the directed-lie control question. *Polygraph*, *28*, 124-142.

Merckelbach, H., & van Koppen, P., (2006). The European Expert Meeting on Polygraph Testing, Maastrich, The Netherlands.

Offe, H., & Offe, S., (2007). The comparison question test: Does it work and if so how? *Law and Human Behavior, 31*, 291-303.

Patrick, C. J., & Iacono, W. G. (1989). Psychopathy, threat, and polygraph test accuracy. *Journal of Applied Psychology, 74*, 347-355.

Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology, 76*, 229-238.

Podlesny, J. A., & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology, 15*, 344-358.

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

Raskin, D. C. (1986). The polygraph in 1986: Scientific, professional, and legal Issues surrounding applications and acceptance of polygraph evidence. *Utah Law Review*, *1986*, 29-74.

Raskin, D. C., & Hare, R. D. (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology*, *15*, 121-136.

Raskin, D. C., & Honts, C. R. (2002). The comparison question test. In M. Kleiner (Ed.), *Handbook of polygraph testing*. London: Academic (1-49).

Raskin, D. C., Kircher, J. C., Honts, C. R., & Horowitz, S. W. (1988). *A study of the validity of polygraph examinations in criminal investigations: Final report to the National Institute of Justice, Grant Number 85-IJ-CX-0400,* Department of Psychology, Salt Lake City, University of Utah.

Reed, S., (1994). A new psychophysiological detection of deception examination for security screening. *Psychophysiology*, *31*, S80. (Abstract)

Rovner, L. I. (1986). The accuracy of physiological detection of deception for subjects with prior knowledge. *Polygraph, 15*, 1-39.

Shadish, W. R., Cook, T. D., & Campbell, D., (2002). *Experimental and quasi experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

United States v. Scheffer, 523 U. S. 303 (1998).

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities, Second Edition*. Hoboken, NJ: Wiley.

Appendix B. ***Eligibility Criteria***

Criteria:	Response:		Answers Required to Participate:
Are you 18 years or older?	Y	N	Yes
Do you have heart problems or take medication for high blood pressure?	Y	N	No
Are you currently under the care of a psychiatrist?	Y	N	No
Have you ever taken a polygraph before?	Y	N	No
Is English your primary spoken and written language?	Y	N	Yes
If female: Is there any chance you are pregnant right now.	Y	N	No

Subject Does Not Meet Criteria:	Subject Meets Above Criteria:
I'm sorry; you do not meet the necessary requirements for the purpose of this study. Thanks for your time and interest.	You meet all of the criteria; we would like to schedule an appointment with you.

Appendix C. *Video Script, Informed Consent*

USERNAME: xxx
PASSWORD: yyy

Hello! I want to welcome you to the laboratory research study being conducted through the Psychology & Law laboratory at BSU.

The purpose of this video is to inform you about your role in this study, your rights as a participant, and to give you the opportunity to consent to participation in this study.

You have been given an INFORMED CONSENT FORM that we will ask you to read carefully and then sign if you agree to participate in this study. Please listen carefully to this video and then carefully read the form. If you are then willing to participate in this study please sign the form when the assistant returns.

When this video is completed, the assistant will return to the room. If you have any questions or concerns, please raise them with the assistant at that time.

The goal of this study is to determine if the computerized polygraph (lie detector) can detect deception. The results of this study will have important implications for future studies of credibility assessment.

Procedure: You will be asked to complete a simple task or series of tasks. The task or tasks that you will complete depend upon which condition of the study you are assigned to. All tasks will be explained to you on separate videos. It will be randomly determined which condition you are in, but please be aware that one of the conditions requires you to commit a mock theft of money. Upon completion of the tasks, you will be given a polygraph examination.

The polygraph examination will focus on your answers to the questions regarding the theft of money. The polygraph examiner will not know if your answers are true or false, and your goal during the examination will be to convince the polygraph examiner that you had nothing to do with the theft. **If you are successful in beating the polygraph, you will receive a bonus of $20. To win the bonus, you must appear truthful to all of the questions on the polygraph tests regarding the theft of money.** After completing the polygraph examination, you will then be debriefed by a research assistant and awarded the bonus if earned.

After signing the consent form, you will be asked to watch another brief video. It will explain the condition to which you were randomly assigned. The task or tasks that you need to complete will be fully explained in the video. After you've watched the video, you may replay the video if you want. After you complete the tasks, the research assistant will show you where to go for the polygraph examination.

During the polygraph test portion of this experiment, several sensors will be attached to your left hand, a blood pressure cuff will be placed on your right arm, and two elastic straps will be attached around your chest and abdomen (on top of your clothes) to measure breathing. All sensors are attached with adhesive collars or Velcro.

None of these sensors will hurt you or harm you in any way, although the blood pressure cuff will squeeze your arm and **may** cause your arm to feel as if it has gone to sleep. Because of the way we measure blood pressure, the cuff pressure will be much less than that used by a physician. This is not harmful or dangerous in any way, and is a standard part of all polygraph tests. The cuff will be inflated for as little time as possible to conduct the test.

Prior to their placement on your body, the sensors will all be shown to you and their use will be explained. Please feel free to ask questions.

After the sensors are attached, the polygraph examiner will read the test questions while making the polygraph recordings. You will have to answer each question with just a "Yes" or "No", but please remember that your goal is to appear truthful and as though you were not involved in the theft.

Duration and Benefits:
The entire experiment can take up to 3 hours to complete. You will be compensated for your time, at $15 per hour.

You will receive an additional bonus of $20 if a standard computer analysis of your physiological responses during your polygraph tests reveals that you are truthful to all of the test questions about the theft of the money.

This determination of truthfulness is calculated by comparing your patterns of physiological responding to the individual polygraph test items. In theory, an individual displays different patterns of responding when being truthful or deceptive. The computer analysis will examine your patterns of responding and assign a value indicating the probability of your truthfulness. If the analysis determines that your patterns of physiological responding are similar to patterns displayed by a truthful individual, the probability of your being truthful will be higher than the probability of your being deceptive (greater than .50).

If this analysis determines you are truthful to all of the test questions, you will earn the bonus.

If you decide to withdraw from the study, you will receive compensation for the time you have given, but will not be eligible for the bonus.

Video recording:
The polygraph examinations in this study will be video recorded. When you sign the consent form, you are giving us permission to video record your polygraph examination. These videos are being made to document the procedures of this study and to be used in presentations, teaching and training in professional settings. By signing the consent form you will be giving us permission to use your image on the video for professional presentations. If you are uncomfortable with your image being recorded or being used for any of these purposes, you should decline to participate at this time. However, if we decide to use your image for such professional purposes, your name and any identifying information will be edited from the recording.

Confidentially:

A list of the names of participants will be maintained for 3 years after the study ends, at which time it will be destroyed.

A code number will be used to organize the physiological data but will not be associated with your name or any other personally identifying information.

Your name or any other personally identifying information will not be included in any publication or reports of this research or in any presentations on this study.

Risks:
There are no known risks to you physically or mentally for participating in this study. If anything about this study makes you feel bad or uncomfortable, we can arrange for a consultation for you at the Boise State counseling center on your request.

Withdrawal: Your participation in this study is voluntary. If at any time you wish to withdraw from the study you may do so and receive payment for the time you have spent to that point.

Concerns: If you have any concerns about how this study was conducted or about protection of your confidentiality you should contact the principal investigator, Dr. Charles Honts, (426-3695) in the Psychology Department at Boise State University.

Please wait for the return of the research assistant. Do not sign the Informed Consent Form until the assistant returns.

Thank you again for your interest and time in assisting the Department of Psychology in this research.

Appendix D. *Informed Consent*

INFORMED CONSENT FORM

The purpose of this document is to inform you about your role in this study, your rights as a participant, and to give you the opportunity to consent to participating in this study.

Purpose: The goal of this study is to determine if people can defeat the polygraph (lie detector), and to validate the accuracy of polygraph tests. This project is funded by the Department of Defense.

Procedure: During the first stage of the study, you will be assigned a condition (i.e., innocent/deceptive). You will then be asked to complete a simple task or series of tasks. Upon completion of the tasks, you will be given a polygraph examination. The polygraph examination will focus on your answers to the items regarding the task that you were asked to complete. The polygraph examiner will not know if your answers are true or false, and your objective during the examination will be to convince the polygraph examiner that your answers to all of the questions are true. If you are successful in beating the polygraph, you will be paid a bonus of $20. To attain the bonus, you must appear truthful to all of the questions on the polygraph tests. After completing the polygraph examination, you will then be debriefed by a research assistant and awarded your earned compensation.

After completing the required tasks, you will be taken to another room and introduced to the polygraph examiner. The polygraph examiner will review each of the questions to be used on the tests and will describe the testing procedure to you. To monitor your physiological reactions to test questions, several sensors will be attached to your right hand, a blood pressure cuff will be placed on your left arm, and two elastic straps will be attached around your chest and abdomen (on top of your clothes) to measure breathing. All sensors are attached with adhesive collars or Velcro. None of these sensors will hurt you in any manner. Prior to their placement on your body, the sensors will all be shown to you and their use will be explained. In your polygraph test the polygraph examiner will read the test questions to you while making the polygraph recordings. Regardless of the testing format, you will answer each question with only a "Yes" or "No."

Duration and Benefits: You will be paid $15 per hour for approximately three hours of your time. You will be paid an additional bonus of $20 if a standard computer analysis of your physiological responses during your polygraph tests reveals that you are truthful to all of the test questions. This determination of truthfulness is calculated by comparing your patterns of physiological responding to the individual polygraph test items. In theory, an individual displays different patterns of responding when being truthful and deceptive. The computer analysis will examine your patterns of responding and assign a value indicating the probability of your truthfulness. If the analysis determines that your patterns of physiological responding are similar to patterns displayed by a truthful individual, the probability of your being truthful will be higher than the probability of your being deceptive (greater than .50). If this analysis determines you are truthful to all of the test questions, you will earn the bonus. No deception is involved in this study. But if for any reason you decide to withdraw from the study, you will be paid $15 per hour of participation, but will not be eligible for the bonus.

Videotaping: Some of the polygraph examinations in this study will be videotaped. When you sign the consent form, you are giving us permission to videotape your polygraph examination. These videotapes are being made to document the procedures of this study and to be used in presentations, teaching and training in professional settings. By signing the consent form, you will be giving us permission to use your image on the videotape for professional presentations. However, if we decide to use your image for such professional purposes, your name and any identifying information will be edited from the tape.

Confidentially: A list of the names of participants will be maintained until three years after the end of the study, at which time it will be destroyed. A code number will be used to organize the physiological data but will not be associated with your name or any other personally identifying information. Your name or any other personally revealing information will not be included in any publication or reports of this research.

Risks: There are no known risks to you physically or mentally for participating in this study.

Withdrawal: Your participation in this study is voluntary. If at any time you wish to withdraw from the study you may do so and receive payment for the time you have spent to that point.

Concerns: If you have any concerns about how this study was conducted or about protection of your confidentiality you may contact the Project Director, Racheal Reavy (426-3601) in the Psychology Department at Boise State University.

Survey: Following the polygraph examination, you will be asked to fill out a very brief survey regarding your experiences during the appointment.

Consent: I, _____, have read the above description of this study and understand it. I also understand this study is funded by the Department of Defense. I have received a copy of this Informed Consent form. I agree to participate in this research.

_____/_____
(Signed) (Date)

_____/_____
Witness (Date)

This project has been reviewed and approved by the Boise State University Institutional Review Board for the protection of human subjects in research.

Appendix E. ***Video Script, Innocent Condition***

By now you have agreed to be a participant in this study and have signed the Informed Consent.

As stated in the previous videotape, there are different conditions in this study. This video will tell you about your role in the experiment.

Your condition assignment was made on a random basis, and you actually chose it with the envelope you selected.

Neither the research assistant who had you choose the packet, nor the polygraph examiner, knows whether you will be telling the truth on the polygraph examination.

You have been selected to be in the innocent condition.

Please listen to these instructions carefully and make sure that you understand exactly what you are to do. Replay this video if necessary. You may make a few notes to help you remember what to do as you carry out these instructions.

There are writing materials on the desk for you, and a written (text) version of this entire process provided in the packet that you have been given. This is a polygraph, or lie detection, experiment. Half of the subjects in the experiment are instructed to commit a theft. They are instructed to go to a room and steal some money from an envelope. Then they report back for a polygraph examination. If they are found innocent on the test, they are paid a bonus in addition to the amount paid for participating in the experiment.

You are not one of those subjects. You are not to steal anything. Your mission, if you choose to accept it, will be to drop off an envelope (located in this packet) in a file folder outside the door of room E619 in the Education Building.

You are an innocent suspect. But you too can receive the bonus by being found innocent on the polygraph examination. The bonus, in addition to the amount you will be paid for your time participating in this experiment, is $20.00. Therefore, it is in your best interest to be truthful during the test and deny having anything to do with the theft of the money.

Before you leave this room, check the time. You have 30 minutes to complete this task. Do not return early. If you finish early, wait until the 30 minutes are up, and then return to the room you are in now, and wait until an assistant comes for you.

You will then be given a test by a polygraph examiner. The examiner will not know if you are innocent or guilty of the theft, which is why you will be treated as though you are a suspect. This is so that the decision can be made entirely on the results of the polygraph test. Remember, you will be in the Education Building delivering an **envelope** to the office, but you won't see money or steal anything. You could easily give yourself away by accidentally revealing any other details, so please maintain your innocence wisely.

You will receive the bonus only if the examiner finds you innocent. So you must actually convince the examiner of your innocence. If the examiner decides that you are deceptive or cannot determine whether you are deceptive or innocent, you will not receive the bonus.

Those are your instructions. **You must follow those instructions exactly to be eligible for your payment.** If you do not wish to participate in this experiment, please inform anyone in the reception area (front of the building). If you are not entirely sure of what you are to

do, push the back button and start this video over. Then push the play button to hear the instructions again. When you are done, push the stop button.

Take the written script you got from the packet you chose and run it through the shredder before you leave this room.

Once you leave this room, you should return in exactly 30 minutes, not sooner, and not later. That is it. Good luck with the examination.

Appendix F. *Video Script, Guilty Condition*

By now you have agreed to be a participant in this study, and have signed the Informed Consent.

As stated in the previous videotape, there are different conditions in this study. This videotape will tell you about your role in the experiment.

Your condition assignment was made on a random basis, and you actually chose it with the envelope you selected.

Neither the research assistant who had you choose this packet, nor the polygraph examiner, knows whether you will be telling the truth on the polygraph examination.

You have been selected to be in the deceptive condition.

Please listen to these instructions carefully and make sure that you understand exactly what you are to do. Replay this video if necessary. You may make a few notes to help you remember what to do as you carry out these instructions.

There are writing materials on the desk for you, and a written (text) version of this entire process provided in the packet that you have been given. This is a polygraph, or lie detection, experiment. Because you are in the deceptive condition, you will steal an envelope containing money from the Education Building. You will then be given a polygraph examination. If you can beat the polygraph by appearing innocent on that test, you will be paid a bonus of $20.00 in addition to the amount that you will be paid for participating in the experiment.

Your mission, if you choose to accept it, is as follows: You will go to room E619 of the Education Building and remove the envelope from the door. That envelope is addressed to Sam Stone. You will verify its contents. Take the contents out of the envelope and conceal it on your person. You can hide it in your wallet or in any of your pockets, but do not put it in your shoe or in your sock. Tear the envelope up and dispose of it in any trashcan. **If you are found innocent on the polygraph examination, you will be paid an amount equal to that which you stole.** However, you must return the money from the envelope when the polygraph examination is completed. Be careful not to leave any fingerprints, and be sure to dispose of the envelope where it will not be found. It is extremely important that you steal the money without alerting anyone to the theft. For example, since room E619 is a faculty office, be sure to have your alibi ready in case someone asks you what you are doing.

You are not, and I repeat, <u>not</u> to tell anyone that you are participating in an experiment. YOU DO NOT WANT TO GET CAUGHT COMMITTING THIS CRIME so be prepared to do this mission in a discrete fashion. If you do get caught please call Dr. Patt Elison Bowers at 426.4119

Before you leave this room, check the time. You have 30 minutes to complete your theft once you leave. Do not return early. If you finish early, wait until the 30 minutes are up, and then return to the room you are in now, and wait until an assistant comes for you.

You will then be given a test by a polygraph examiner. The examiner will be testing you about the theft of the missing money, and he or she will not know if you are innocent or guilty of the theft because half of the subjects in the experiment have not committed the theft. This is so that the decision can be made entirely on the results of the polygraph test.

Do not make the examiner suspicious when he or she is interviewing you during the initial portion of the test. Your alibi is to tell the examiner you were in the Education Building delivering an **envelope** to the office, but that you never saw money or stole anything. You could easily give yourself away by accidentally revealing any other details, so please maintain your innocence wisely.

So, when the polygraph examiner asks you questions about any other details about the theft, you must not only deny knowing anything other than that, but you must do so sincerely so that he or she does not become suspicious. If at some point during the test you think you blew it, do not give up.

You will receive the bonus only if the examiner finds you innocent. So you must actually convince the examiner that you are innocent. If the examiner decides that you are deceptive or cannot determine whether you are deceptive or innocent, you will not receive the bonus.

Those are your instructions. **You must follow those instructions exactly to be eligible for the bonus payment.** If you do not wish to participate in this experiment, please inform anyone in the reception area (front of the building). If you are not entirely sure of what you are to do, push the back button and start this video over. Then push the play button to hear the instructions again. When you are done, push the stop button.

Take the written script you got from the packet you chose and run it through the shredder before you leave this room

Once you leave this room, you should return in exactly 30 minutes, not sooner, and not later. That is it. Good luck with the examination.

Appendix G. Polygraph Examination Consent

POLYGRAPH EXAMINATION STATEMENT OF CONSENT

Place: Boise State University Date: _____ Time: _____

Statement of Consent of: _____

The polygraph examiner, _____, has explained the nature of the polygraph examination and told me that I cannot be required to take such examination without my consent. I was further advised that the examination room is equipped with a video camera and that the examination will be video-tape recorded. I have been advised that the results of this examination will be kept confidential. Understanding my unqualified right to refuse, I, _____, do hereby, this date, voluntarily and without duress, coercion, or unlawful inducement, consent to a polygraph examination concerning the theft of an envelope containing $20 from an office in the Education Building.

Signature of Examinee _____

Signature of Examiner _____

Appendix H. Polygraph Examination Question List

Neutral Questions:

Are we in the State of Idaho?

Are the lights on in this room?

Comparison Questions:

Prior to 2008, did you ever lie to someone who trusted you?

Prior to 2008, did you ever do anything that was dishonest illegal?

Prior to 2008, did you ever lie to a person in a position of authority?

Relevant Questions:

Did you steal that missing envelope?

Did you steal the envelope from the door of Room 619 in the Education Building?

Do you know where the missing money is now?

Other Questions:

Regarding the envelope that was stolen from the Education Building, did you intend to truthfully answer each question about that?

Do you believe that I will only ask you the questions we reviewed?

Prior to 2008, did you ever lie to a person in a position of authority?

A Discussion of Countermeasure Strategies and the Comparison Question Test

Thomas Kuczek[1]

Raymond Nelson

Abstract

Following is a discussion of how counter measures may cause observable change in the pattern of electrodermal responses to comparison question test stimuli. Here, the notion of pattern refers not to the qualitative or quantitative aspects of the individual responses to individual test stimuli, but to the pattern of loading of greater observable changes in physiological activity. For this discussion, a counter measure technique will increase or dampen physiological responses to a polygraph question, while a counter measure strategy consists of the application of a countermeasure technique to a subset(s) of stimulus questions during a polygraph exam.

[1]Dr. Kuczek is an Emeritus Professor of Statistics, Purdue University.

Introduction

A counter measure technique can be thought of generally as any technique employed by an examinee to increase or dampen a physiological response to a question during a polygraph exam. A counter measure strategy can be thought of as a defined subset of questions to which an examinee plans to alter their physiological responses - increasing, decreasing, or a combination of both. Execution of a counter measure strategy, if undetected, may increase the likelihood of a guilty and deceptive examinee's exam being evaluated incorrectly as truthful. For broader discussion see Honts, (2014).

Following is a discussion of the effect of a counter measure strategy on electrodermal activity (EDA) in the context of a Comparison Question Test (CQT), a widely used polygraph examination method (Raskin, Honts & Kircher, 2014). EDA consists of both tonic and phasic components, and it is the phasic response information – also referred to as electrodermal responses (EDRs) – that are extracted and scored by field polygraph examiners. CQTs consist essentially of comparison question (CQs) and relevant questions (RQs), along with other procedural questions. Test data analysis consists essentially of the analysis of differences in pattern of physiological responses to RQs and CQs. In this usage, the notion of pattern refers not to the qualitative or quantitative aspects of the individual responses to individual test stimuli, but to the pattern of loading of greater observable changes in physiological activity.

How a specific counter measure strategy may cause observable changes in the pattern or loading of EDRs is of particular interest to field polygraph examiners, and also to those who are more generally interested in the validity and accuracy of the CQT. However, the details of actual counter measure techniques are not of particular interest to this discussion. Also, only EDA data is considered here, though readers may begin to surmise the degree to which this discussion may generalize to other recording sensors. Three different counter measure strategies are considered.

Counter measure Strategies and Tonic EDR.

CQT formats consist commonly of three CQs and two to four RQs, in addition to other procedural questions. RQs may be formulated with an assumption of independence, in the context of a multiple issue exam, and may also be formulated around a single investigation target issue. Regardless of whether RQs are developed as single issue, or more multiple issues, the sequence of questions is typically presented three to five times.

Consider the EDR to any question, whether RQ or CQ. It is assumed that the EDA data has been subject to any necessary signal processing and feature extraction so that the response magnitudes of EDRs can be compared for the RQs and CQs. A variety of stimuli evoke a noticeable EDR, this can include such things as stimulus salience, novelty, and intensity. Here, for the purpose of this discussion, we shall assume that these are reasonably controlled for in the administration of the CQT, and that differences in observed EDRs is a function of, or indication of, the degree of salience of the various test stimuli. Scoring is done for each question pair – referred to by field polygraph examiners as an analysis subtotal (a.k.a. "spot") – by comparing the RQ EDR to the CQ EDR.

For this discussion, the following notation will be used: the EDR for the CQ will be denoted EDRComp while the EDR for the RQ will be denoted EDRRel. Differences in the distributions of EDRs are interpreted as an indication of differences in the salience of the test stimuli. The difference between the two will be denoted Diff, and is defined by:

$$Diff = EDR_{Comp} - EDR_{Rel}.$$

1. If the examinee is innocent and truthful, the distribution of EDRs for CQs is expected to be of overall greater magnitude than the distribution of EDRs for RQs.

2. If the examinee is guilty and deceptive, the distribution of EDRs for the RQs is expected to be of greater magnitude than the CQ distribution.

These two points are reflected in the Empirical Scoring System (ESS) (Handler et al., 2010). The ESS converts the EDR values into an integer according to the Diff value. A positive integer score (+2) is given if EDR Diff is positive, while a negative integer score (-2) if the EDR Diff is negative. A score of 0 is assigned if the if the EDR Diff is close enough to zero to be uninformative. Some information on the individual EDR values is lost in this transformation, but the simplicity of ESS integer scores is advantageous when computing norms for ESS scores. A multinomial distribution can be calculated from the (+, 0 , -) integer scores, representing the null hypothesis that there is no association between observed EDRs and the criterion of deception or truth-telling.

After feature extraction and transformation to integer values, ESS scores are then tallied. The resulting aggregated values can be compared to scoring norms. This normative comparison can support conclusions for a number of questions. For example, should the test be classified as deceptive or truthful? Or, what is the likelihood of a correct or incorrect classification? Also, what is the range in which we expect to observe the result of another similar exam?

To understand the range of possible observations for an examine who is employing countermeasures it is worthwhile to examine possible patterns in EDR values. Consider the examinee who is employing no counter measures. If the examinee is innocent and truthful, the EDR of answers to CQs will be of generally larger magnitude than EDRs to RQs. The $Diff = EDR_{Comp} - EDR_{Rel}$ will be a positive number.

Suppose an examinee is guilty or deceitful and not employing a counter measure strategy. The EDR_{Rel} values are expected to be of generally greater magnitude than EDR_{Comp} values. The Diff value will be a negative number. The distance of the EDR_{Comp} - EDR_{Rel} difference from zero can be evaluated statistically to answer empirical questions such the proportion of guilty or innocent persons that produce an equivalent or more extreme score, or the proportion of expected false-negative or false-positive errors.

Now consider an examinee who is employing a countermeasure strategy. If this examinee is dampening their EDR_{Rel} magnitudes, the potential effect will be to bias ESS scores towards a truthful classification. Alternatively, if an examinee is increasing or augmenting their EDR_{Comp} magnitudes, the effect may also be to bias the resulting ESS scores towards a truthful classification.

Examples: Strategies, EDR and Biases.

Example 1.

Consider an examinee who is guilty or deceptive during a polygraph exam and employs a countermeasure strategy in attempt to appear innocent or truthful. A successful strategy would cause the magnitudes of EDR_{Comp} to generally exceed the magnitudes of EDR_{Rel}. The examinee's strategy may be to dampen EDR_{Rel} to make Diff more positive or increase EDR_{Comp} or combine both strategies. All three of these strategies, if undetected, might bias the Diff value in a positive direction, leading to more truthful ESS scores, and increasing the likeli-

hood of a truthful classification.

1. Dampen EDR response to relevant questions: If an examinee is dampening their physiological response to the RQs, then the magnitudes or distribution of EDR_{Rel} values will be biased downward relative to the distribution of EDR_{Comp} values. The effect of this may be to produce more truthful ESS scores, that may begin to mimic that of an innocent and truthful examinee.

2. Increase EDR_{Comp} responses to comparison questions: If the examinee is successful in increasing the magnitude of EDR_{Comp} values for the comparison questions, relative to EDR_{Rel} values, a practical effect of this can be to increase the likelihood of positive scores that can begin to mimic those of truthful or innocent persons.

3. Employ both strategies, dampen EDRs to RQs and increase EDRs to CQs: By increasing EDRComp for the comparison questions and also dampening EDR_{Rel} for the relevant questions, the Diff values may be increased even more than using either technique alone. A rationale for attempting both strategies might be to mitigate the imperfect effects of both strategies or to better disguise observation and detection of the countermeasure strategy. If successful, this would possibly contribute to an even greater increase in the likelihood that a guilty or deceptive person may produce scores that lead to a truthful classification.

Example 2.

Consider an innocent examinee who answers truthfully but is fearful of a false positive error or inconclusive result. This examinee may employ a counter measure strategy similar to that of a deceptive person, in attempt to bias the ESS scores towards a truthful classification. Again, consider how each of three counter measure strategies, if undetected, may affect the distribution of EDRs.

1. Dampen EDR_{Rel} magnitudes:

In this situation, the EDR_{Comp} magnitudes are presumed to be unaffected, and the Diff values will be biased toward a positive direction. Corresponding ESS scores may be biased in a positive direction, increasing the likelihood of a truthful classification, for which field polygraph examiners use the terms No Deception Indicated or No Significant Reactions.

2. Increase EDRComp magnitudes: If an examinee is successful in increasing EDR_{Comp} magnitudes, leaving EDR_{Rel} magnitudes unaffected, the result might be to bias the resulting Diff values in a positive direction. The effect of making Diff = EDR_{Comp} - EDR_{Rel} more positive could correspondingly bias the ESS scores towards a truthful classification.

3. Employ both strategies to the entire exam: Intuitively, dampening EDR_{Rel} magnitudes and increasing EDR_{Comp} magnitudes may bias the Diff scores and ESS scores in a positive direction, increasing the likelihood of a truthful classification – with a corresponding decrease in the likelihood of an inconclusive or erroneous classification.

Discussion

Three different countermeasure strategies have been described, as applied to the CQT in two different contexts. These strategies include the dampening of EDR_{Rel} values, amplification of EDR_{Comp} values, and the combination of these two strategies. All three strategies may be attempted by guilty or deceptive examinees wishing to evade the likelihood of detection, and also by innocent and truthful examinees wishing to increase the likelihood of a favorable test result. The rationale for these strategies is to alter the loading of the relative magnitude of change in physiological activity, in response to different

types of test stimuli, so that greater changes in physiology are loaded at the CQs. Although the strategies may be similar for guilty and innocent persons, the requirement for successful execution differ substantially. Whereas innocent person seek only to augment the expected pattern of loading, guilty persons will seek to consciously and voluntarily reverse the pattern of loading – while also producing data of normal autonomic quality.

An obvious limitation of this discussion is that it does not address the matter of exactly how an examinee might dampen or amplify their EDRs to individual questions. Nor does this discussion address whether reliable and effective methods exist to achieve these goals. To be maximally effective, a countermeasure would ideally produce EDRs that cannot be differentiated from authentic autonomic activity – in terms of timing, magnitude, and other response characteristics. It remains an ambitious suggestion that this could be easily accomplished via voluntary/somatic activity, whether behavioral or cognitive. Also, although the dampening and amplification of physiological responses is discussed herein as arbitrarily equivalent in their possible effect sizes – there may be inherent practical limitations or differences associated with dampening or amplification strategies such that these may actually be asymmetrical in terms of expected effect sizes or ease of execution.

Another matter that is not addressed in this discussion is the methods – whether via subjective expertise and experience, or objective technology and automation – through which field polygraph professionals may observe and detect the use of a countermeasure strategy. An effective countermeasure strategy will reverse the loading of EDA_{Comp} – EDA_{Rel} = Diff scores while remaining undetectable. A systematic review of countermeasure studies (Handler, Honts, & Goodson, 2015) indicates that both the execution and detection of countermeasures are complex objectives that extend beyond simplistic assertions.

A potential hazard to field examiners is that failure to detect countermeasure strategies may result in an increase in false negative errors. Another hazard is that undetected countermeasures may lead to an increase in inconclusive test results, and these may be vulnerable to misrepresentation. A corresponding hazard is that scrutinizing data for countermeasure activity may contribute to an increase in false-positive or inconclusive classifications if conclusions about countermeasures are reached precipitously.

Hazards to the examinee may include the potential that countermeasures may fail to produce the desired reversal of loading in physiological activity, in addition to the fact that ineffective or poorly executed countermeasure strategies may be more easily observed and detected. An effective countermeasure strategy, in addition to being undetected, must reverse the loading of reverse the loading of EDA_{Comp} – EDA_{Rel} = Diff scores, while appearing indistinguishable from normal autonomic activity. It remains speculative as to whether voluntary strategies can easily accomplish this, and there is a possibility that voluntary activity may alter the normal autonomic data in observable ways.

Because the CQT is a structured and systematic method, an effective countermeasure must be correspondingly systematic. This is because altering the score in an unsystematic, random, or haphazard manner is less likely to alter the systematic loading of EDRs. Whereas more systematic activity patterns may be more easily observed and detected, more random or unsystematic activity may be less likely to be effective. Additional research is needed to increase the available knowledge about potential strategies for dampening and amplification of EDR_{Comp} and EDR_{Rel} magnitudes, toward the differentiation of autonomic can voluntary activity, and towards the automated statistical identification and detection of countermeasure strategies.

References

Handler, M., Honts, C. R., & Goodson, W. (2015). A literature review of polygraph countermeasures and the comparison question technique. *Polygraph, 44*(2), 129-139.

Handler, M., Nelson, R., Goodson, W, and Hicks, M. (2010). Empirical Scoring System: A cross-cultural replication and extension study of manual scoring and decision policies. *Polygraph 39*(4), 200 – 215.

Honts. C. R. (2014) Countermeasures and credibility assessment. In D. Raskin, C. Honts, J. Kircher, Eds.). Credibility assessment. (130-154). Elsevier.

Raskin, D. C., Honts, C. R., & Kircher, J. C. (2014). Credibility Assessment. Elsevier.

Terminology Reference for The Science of Psychophysiological Detection Of Deception

4th Edition, 2022

Donald Krapohl

Mark Handler

Michael Lynch

The information published in the *Terminology Reference for the Science of Psychophysiological Detection of Deception* does not necessarily reflect the position of the American Polygraph Association.

Introduction

Since the first edition of this reference was published 25 years ago much has taken place in the polygraph field, progress which has continued in the decade since the last edition of the *Terminology Reference*. The transition from analog to digital polygraph is now complete, the polygraph profession has accomplished the shift from authority-based practices to those that are evidence-based, concepts and terms adopted by the polygraph field are substantially more mainstream, new credibility assessment technologies have appeared while others have disappeared, and much to the surprise of critics and detractors, predictions of the demise of the polygraph have proven to be premature.

In this fourth edition of the *Terminology Reference for the Science of Psychophysiological Detection of Deception* we strived to capture these changes. We have updated the references, added new terms, removed others, and included images for some terms to help readers understand them better. We hope readers appreciate these updates.

And finally, we sadly report the passing of our friend, colleague and contributor to previous editions of this work, Shirley Sturm in 2020. Shirley was one of the greats in the polygraph field. She was the first woman President of the American Polygraph Association, a teacher, a coveted mentor, and a fiercely independent thinker. We and the rest of the polygraph field will miss Shirley, her sense of humor, her passion for polygraph, her concern for her fellow practitioners, her commitment to the field. There was always only one Shirley Sturm and we were fortunate to have known her. Because of her lasting contribution to our shared field of endeavor we dedicate this edition of the *Terminology Reference* to Shirley Sturm.

3-position scale

Abbreviated form of the *7-position scale* for PDD test data analysis. The major difference is that the range of values for each comparison is from -1 to +1, rather than the range of -3 to +3 in the 7-position scoring system. This process is based on the simple and robust principle that physiological reactions of greater magnitude are caused by stimuli that are more salient to the examinee due to emotional, cognitive, or behaviorally conditioned factors. See: Capps and Ansley (1992); Harwell (2000); Krapohl (1998); Van Herk (1990).

7-position scale

System for assigning values to individual physiological responses in PDD, based on differential responding to relevant and comparison questions. The values in 7-position scoring are whole numbers between -3 and +3. By convention, negative values represent greater responding to relevant questions, while positive values indicate greater responses to comparison questions. A zero usually indicates equal or no reactions to the relevant and comparison questions, or that the spot does not meet minimum standards for interpretation. In the PDD literature, the 7-position scale is sometimes referred to as a semi-objective scoring system and is loosely based on the psychometric scales developed by Rensis Likert. See: Bell, Raskin, Honts, & Kircher (1999); Handler & Nelson (2008); Swinford (1999); Weaver (1985).

acetylcholine (ACh)

Neurotransmitter substance found in the motor nerves to skeletal muscle, in preganglionic autonomic endings (sympathetic and parasympathetic), in postganglionic parasympathetic nerves, postganglionic sympathetic nerves to eccrine sweat glands and muscle vasodilator endings, and many parts of the brain as well as in some cells in the retina. Neurons that release ACh are called *cholinergic neurons*.

acquaintance test

Typically, a test given at the beginning of the test phase of an examination in which the examinee agrees to lie to a number or letter. The test is not scored nor used for veracity testing. The acquaintance test serves several purposes: to familiarize the examinee with the test procedures; to properly set the gains and centerings; to help detect countermeasures; and to assess the range of responsiveness of the examinee. Other forms and terms for this type of test include Numbers Test, Demonstration Test, Stimulation (Stim) Test, Practice Test, Blind Numbers Test, Calibration Verification of Sensitivity Test (CVOS), and True Blue Control Test.

activity monitor

Device attached to or built into PDD testing chairs designed to record peripheral behavioral activity and cooperation during the examination.

Adrenalin

Trademark name for epinephrine (called *adrenaline* in British references), discovered and named by J. Takamine in 1901. See: *epinephrine*.

adrenergic

Those neurons that release the neurotransmitter norepinephrine. Also, substances that mimic norepinephrine in its physiological effects.

afferent nerves

Nerve fibers that carry impulses from the periphery toward the central nervous system. Also called *sensory nerves*.

Air Force Modified General Question Test (AFMGQT)

PDD comparison question testing format with flexible question orderings and numbers of relevant questions. The AFMGQT can be used in single-issue, multiple-facet, and mixed-issue PDD examinations. The AFMGQT uses relevant, probable-lie, sacrifice relevant, and irrelevant questions. Some versions permit the use of directed-lie comparison questions. See: Senter, Waller & Krapohl (2008).

algorithm

System of standardized steps that lead to a solution. A series of decision rules. The term algorithm is frequently used in the context of automated data analysis that produces a decision or result. Algorithms have many applications; Polygraph algorithms analyze the physiologic data from structured examinations and make estimates of the likelihood of deception or to assess which question has elicited the greatest physiologic response. There are several algorithms now available for analysis of polygraph data.
See: Dollins, Krapohl, & Dutton (1999); Nelson, Krapohl & Handler (2008).

alarm reaction

The first stage of general adaptation, which is triggered by the impact of the stressor and is characterized by heightened sympathetic activity.

all-or-none law

A neuron will respond to its greatest ability or not at all. Stimuli that do not meet or exceed the threshold will not be sufficient to cause a response. Skeletal muscles, cardiac muscles and nerve tissues conform to the all-or-none law.

allostasis

A central nervous system mediated, integrated brain-body response geared towards viability or survival. It occurs in regulatory systems which have no fixed set point and is comprised of both physiological and behavioral processes designed to maintain internal parameters within limits essential for life. A modern concept that replaces "homeostasis" as a central concept in physiological responding. See: Handler, Rovner & Nelson (2008).

alpha

A Greek letter (α) used to denote the probability of incorrectly rejecting a null hypothesis in statistical testing. It refers to the probability of making a Type-1 or false-positive error. If the probability of obtaining the sample result is less than alpha (often set a priori less than or equal to 0.05), the results are declared "statistically significant."

American Association of Police Polygraphists (AAPP)

Professional organization dedicated to serving the needs of criminal justice and military PDD examiners. Founded in 1977, AAPP has about 1500 members and is headquartered in Waynesville, Ohio.

American Polygraph Association (APA)

Professional organization made up of PDD professionals from law enforcement, government, and the private sector. Incorporated in 1966 in Washington, D.C., the APA resulted from the merger of the several polygraph associations, including the Academy of Scientific Interrogation, the American Academy of Polygraph Sciences, the National Board of Polygraph Examiners, the International Association of Polygraph Examiners, and the International Association for Polygraph Research. It currently has about 2,800 members and is headquartered in Chattanooga, Tennessee.

Americans with Disabilities Act (ADA)

"An Act to establish a clear and comprehensive prohibition of discrimination on the basis of disability." PDD is used in employee selection and loss investigation in both the public and private sector, and some provisions of the ADA limit the types of inquiries that can be included in the test coverage. Some of the lines of questioning historically taken during pre-employment screening may not be in compliance with ADA rules (e.g. historic alcohol or drug use). Provisions of the Act can be found at the Department of Labor Web site: www.dol.gov/.

anacrotic limb

Ascending portion of an arterial pulse wave.

analog study

Experimental design that attempts to replicate real world activities under controlled circumstances, i.e., a mock crime study. Analog methodologies are frequently used in polygraph validation studies because it is possible to establish ground truth, something extremely difficult to determine independently in field cases. Despite this important feature, mock crime studies in polygraphy have been criticized for not being able to induce the level of emotional involvement or personal risk engendered in real criminal investigations.

analysis spot

The specific location on a polygraph chart where the spot analysis concept is employed and is generally anchored to a relevant question.

anecdote

Personal narrative relating to an issue or event. Anecdotal evidence pertains to non-empirical observations, and it is sometimes used to forward an assertion for which there may be no clear evidence.

anticlimax dampening concept

Based on Cleve Backster's overarching concept for the polygraph that a person's fears, anxieties, and apprehensions will be directed to that situation which holds the greatest threat to his or her well-being or self-preservation at that point in time. In the polygraph examination, a guilty examinee's concern over an intense relevant question may result in a full or partial dampening of responses to other questions. The anti-climax dampening concept remains a proposed explanation for an often-observed phenomenon in which examinees sometimes react only to a single relevant question when they are actually being deceptive to two or more in the same test. The concept is a description, however, and not an explanation. See: Backster (1963a).

anti-countermeasures

Preventative measures used by PDD examiners to block or neutralize the countermeasure efforts of examinees. For example, if the aim were to be to preclude an examinee from pressing his toes against the floor during testing, an anti-countermeasure might include elevating the examinee's feet from the floor so that this strategy cannot be acted upon. Lynn Marcy is credited with making the distinction between anti-countermeasures (proactive) and the counter-countermeasures (reactive).

aorta

Main systemic artery from the heart. The aorta receives blood from the left ventricle through the aortic valve, normally tricuspid and having three leaflets. The upward extending portion is considered the ascending aorta, followed by a downward bend, the arch of the aorta. The portion passing through the chest is the thoracic aorta, from where the blood flows to all parts of the body.

apnea

Temporary cessation of breathing. Apnea is considered the ultimate manifestation of respiratory suppression or slowing. When they are specifically associated with certain questions during a polygraph examination, they are considered significant physiological reactions and strongly diagnostic of deception. True involuntary apneas almost always take place near the bottom of the exhalation cycle.

a posteriori

(L: from what comes later). That which is done after the experiment. This expression is commonly seen in research. An example of a typical a posteriori decision is post hoc statistical analyses of data that had not been anticipated by the experimenters. Opposed to *a priori*.

a priori

(L: from what precedes). Refers to that which is done before the conduct of an experiment. An example is the number and type of subjects to be recruited for the study, or the alpha level for a PDD test result to be considered significant. Opposed to *a posteriori*.

Army Modified General Question Test (Army MGQT)

Test format patterned after the Reid test and developed by the U.S. military. Unlike the Zone formats, it has more relevant questions than comparison questions and does not include symptomatic questions, though some versions employ a sacrifice relevant question. The use of the Army MGQT has declined since research has shown poor validity. See: Blackwell (1998); Krapohl & Norris (2000); Podlesny & Truslow (1993).

arousal

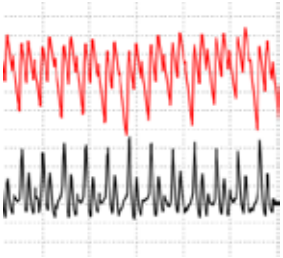
A state of excitement. Lacey proposed that three types of arousal exist 1) behavioral, 2) autonomic, and 3) cortical. Behavioral arousal can be observed in a person's outward responses, autonomic arousal can be measured by psychophysiological changes in the peripheral nervous system, and cortical arousal can be measured as EEG desynchonization and fast waves. See: Lacey (1967).

arousal theory

One of several theories that attempt to explain PDD. It holds that stimuli have different intrinsic cognitive and/or emotional arousal potential for each subject, and discrimination between guilty and innocent subjects is achieved by assessing which type of stimuli has drawn the most responsivity which are mediated via behavioral conditioning.

arrhythmia

Any variation of the heart's normal rhythm is considered an arrhythmia. An example of a regular variation, associated with respiratory activity, is sinus arrhythmia. Examples of irregular variations include premature contractions (PVCs), heart block, flutter, fibrillation and other ectopic beats. Rapid arrhythmias are termed tachycardias while slow arrhythmias are termed bradycardias. The image below shows the effect of PVCs on the cardiograph (top) and photoplethysmograph (bottom) tracings.



PDD Terminology

Arther technique

A probable-lie comparison question technique developed by Richard O. Arther (d. 2007) which he taught at his training school in New York City. It was built upon the methods of the Reid school, of which he was a student. The Arther technique relies heavily on behavioral indicators to assess deception. His question series is unique in that it includes a known truth pseudo-relevant question. Arther taught thousands of students in the U.S. and abroad and focused on the law enforcement sector.

attention

The focusing or concentration of mental activity. Attention is central to the theory of *salience*, currently the most parsimonious explanation for the differential arousal patterns observed with the Comparison Question Technique.

autogenic training

Instruction in one form of self-hypnosis. Subjects can be taught to induce a state of profound relaxation through a series of simple exercises. Because autogenic training can help individuals modify physiological arousal, it is considered by some as a potential countermeasure against PDD. There is little data to support or refute this argument. Because autogenic training focuses on self-regulation of tonic levels of physiological activity, and conventional polygraphy relies on phasic response patterns, its usefulness in defeating modern PDD techniques is not expected to be high.

automatic mode

Depending on the polygraph manufacturer, this mode may also be labeled as Auto, Auto-Center or Auto EDA. It is a setting for the electrodermal activity channel, which uses a combination of filtering that may include; a low-pass filter, a high-pass filter and/or a smoothing filter. It is intended to keep the electrodermal tracing near the center of the display by filtering out low frequency changes that are typically a consequence of shifts in tonic skin conductivity. All modern polygraphs have this feature. While the self-centering action of the automatic mode is attractive to many PDD examiners for practical reasons, it has been criticized for altering data in ways that affect data analysis. Most modern instrument manufactures include both a manual and automatic mode.

autonomic nervous system (ANS)

In vertebrates, the system of nerves that regulates all innervated tissues and organs except striated muscle fibers. The ANS is divided into the sympathetic and parasympathetic portions. The ANS performs the vegetative functions and regulates arousal levels. All conventional PDD methodologies monitor ANS activity. See *sympathetic* and *parasympathetic nervous systems*.

AVATAR

A platform of multiple sensors designed to detect deception at border crossing and airports. AVATAR stands for Automated Virtual Agent for Truth Assessment in Real-time. Field testing has suggested it performs at levels higher than chance. See: Elkins, Golob, Nunamaker, Burgoon and Derrick (2014).

Axciton

An American manufactured computer polygraph, developed and marketed by Bruce White of Houston, Texas.

axon

The long central process of a neuron. A single axon extends from each cell body of the neuron to the synapse or end organ and is responsible for the transmission of the nerve impulse. In humans and other

vertebrates, most peripheral axons are sheathed in a fatty layer called myelin, which acts to insulate the axon from the surrounding tissue. There are regular breaks in the myelin sheathing, called the nodes of Ranvier, that allow the electrical impulses to jump from node to node (saltatory transmission) rather than transit in the axon via the slower chemical depolarization process.

Backster, Cleve

Originator of the Trizone Comparison Test, more often referred to as the Zone Comparison Technique. Backster (d. 2013) also introduced to the polygraph profession the concepts of "psychological set," zones, spots, superdampening, anticlimax dampening, symptomatic (outside issue) questions, exclusionary comparison questions, and 7-position scoring for use in chart analysis. Backster owned a private training facility in San Diego, California, and provided training for thousands of examiners since the late 1940s. He began the CIA polygraph program in 1949.

baseline arousal

Term used in PDD to characterize a marked upward shift in the entire breathing tracing. Baseline arousals do not always occur during deception; however, when they are observed they are reliable indicators of stress. Some baseline arousals are relatively short-lived, lasting only a few breathing cycles, while others may continue much longer.

base rate

Incidence of something in a population, often expressed as a proportion or percentage. Base rates can affect the confidence in a decision. When base rates are relatively high, detection is much easier than in low base rate conditions. As an example, in medicine there is a much lower false positive rate with a diagnostic technique when the base rate of incidence is 50% than 0.1%. Similarly, in PDD it is much easier to find the one guilty individual in a tested population of two (50% base rate) than the 1 guilty person among 1,000 suspects (0.1% base rate). Statements of PDD validity are incomplete unless they also identify the characteristics of the tested population. The base rate problem is not unique to PDD but is also found in all other diagnostic tests. See: Gastwirth (1987); Kircher & Raskin (1987); Murphy (1987).

behavioral analysis

Method for developing extra-polygraphic information that may be useful in the diagnosis of truth or deception. The behaviors of interest may be naturally occurring within the context of the polygraph examination or are the result of questions posed by polygraph examiners during the interview phase of the session. The inclusion of behavioral analysis in the polygraph decision process is controversial.

behavior countermeasures

A class of countermeasures by which an examinee attempts to sway the examiner or influence the conduct of the session. The intent is to influence the decision maker or restrict the ability of the decision maker to conduct a valid examination. By definition, behavioral countermeasures have no influence on the physiologic tracings. The effect is sociopsychological in nature (interpersonal dynamics) rather than psychophysiological. See: Krapohl (1996).

Berkeley psychograph (or Lee Polygraph)

Two-channel polygraph with event marker assembled by C.D. Lee of Berkeley and used in criminal testing. It included a stimulation marker, pneumograph and cardiograph, and recorded these channels simultaneously on moving graph paper. It was designed to be as portable as a suitcase. Lee sold his Berkeley psychographs from the 1930s into the 1950s. They were delivered complete with pens, sensors, paper, and instructions.

PDD Terminology

Benussi, Vittorio

One of the first researchers to examine breathing tracings for detecting deception. He proposed that a ratio created by the time need to inhale divided by the time to exhale produced a ratio that changed during deception. Though Italian, Benussi did most of his work in Austria at the University of Graz. Benussi's approach was very different from what is sometimes taught in polygraph schools, and subsequent researchers did not find the same degree of accuracy that Benussi reported. See: Benussi (1914); Landis & Wiley (1926).

beta blocker

Adrenergic blocking agent affecting responsivity of the cardiovascular system and used to treat specific cardiovascular conditions such as high blood pressure and arrhythmias. Because they appear to dampen cardiovascular phasic responsiveness, beta blockers are of concern to PDD examiners because of their potential as a pharmacological countermeasure. Testing of examinees who take beta blockers is routine practice, as some cardiovascular responsiveness usually persists even with the medication.

bias

In research it is a source of systematic error that can influence the outcome of the experiment. Bias can be introduced into research by factors such as, among others, nonrandom sampling, faulty subject instructions, or expectations of the researcher or the participants. In a PDD study looking at the validity of blind scoring, for example, a researcher using only cases which were verified by the original examiner are likely biasing the study, since cases that the original examiner made the wrong diagnosis could be systematically excluded from the research sample. Researchers attempt to control bias through experimental design.

biphasic response

Physiological reaction that has two phasic responses in opposite directions. Biphasic responses of a diagnostic nature in PDD are found in the skin potential and heart rate recordings.

biofeedback

Use of a device to measure and convey physiological information back to a subject. It has been proposed as a potential countermeasure approach. Biofeedback is not a countermeasure in itself, but rather a means of teaching a subject to influence autonomic responsivity. Research indicates that biofeedback is better suited for self-regulation of tonic activity than phasic activity. As a result, it may be useful to dampen or increase responsiveness in general during a test, but less so within a test to a given question. It has been suggested that this has implications for test formats such as Relevant/Irrelevant and Peak of Tension, but probably not for tests using comparison questions such as the Reid and Zone formats. Biofeedback can also be used to enhance responsiveness in subjects during testing. It has been hypothesized that the immediate biofeedback serves to elicit greater responding to questions when deception is practiced by the examinee. There is some empirical evidence of this effect, but it has not been unequivocally established.

bizone

Incorrect term applied to a Zone Comparison Test containing two relevant questions. See *zone*.

black zone

In the Backster framework, it is a 20- to 35- second block of polygraph chart time initiated by a symptomatic question having a unique psychological focusing appeal to the examinee who is fearful that the examiner may ask an unreviewed question dealing with an outside issue. Because subsequent research

has not been supportive of the symptomatic question, the value of the black zone has been called into question. See: Backster (2001); Krapohl & Ryan (2001).

blind chart analysis

Evaluation of PDD recordings without the benefit of extra-polygraphic information, such as subject behavior, case facts, pretest admissions, base rates of deception, etc. Studies employ various degrees of "blindness." It is a popular research approach to gauge interrater reliability. Assessments of the accuracy of PDD test evaluation techniques also use blind chart analysis.

blind stimulation test

Stimulation test in which the examiner does not know the critical item at the beginning of the test. See *stimulation test*.

blocking

Pattern sometimes seen in the pneumograph tracing in which the examinee discontinues breathing at or near the inspiratory peak. Blocking can be differentiated from a typical apnea in that the latter most often occurs near the end of the expiratory segment of the breathing cycle. Blocking is many times a deliberate attempt on the part of the examinee to influence the physiological activity recorded by the polygraph.

blood pressure

The force blood exerts against the walls of the blood vessels, usually measured in millimeters of mercury, is called blood pressure. PDD examiners evaluate only relative blood volume changes, as current polygraphs are not capable of providing absolute blood pressure measurements. See *systolic blood pressure*, *diastolic blood pressure*, *mean blood pressure*, and *pulse pressure*.

blood volume (BV)

Quantity of blood in an organ or limb, usually recorded as relative increases and decreases in the circumference of the affected area or size of blood vessels. Localized changes in blood volume are mediated by chemical and neural mechanisms, such as the shunting of blood to the major muscle groups during sympathetic nervous system activations.

bogus pipeline (BPL)

Sociopsychological effect whereby a subject will make more candid and sometimes more incriminating revelations about himself when he believes a device attached to him will reveal his true knowledge or attitudes. The BPL was first reported by Jones and Sigall, who convinced subjects that the electromyograph used in their study could measure internal feelings and found that their subjects were more willing to disclose socially-undesirable attitudes. Some critics of polygraphy contend that PDD is merely an elaborate BPL that has only the power to elicit confessions but not assess truthfulness or deception. See: Jones & Sigall (1971).

bootstrapping

A statistical technique. Bootstrapping involves pooling the data from two samples and drawing samples repeatedly from the pool, with replacement, to create a single distribution. The number of samplings taken to create the distribution is typically in the thousands. Then the two original samples are compared using this grand distribution as an estimate of the true population to determine whether they are significantly different from one another. Bootstrapping has been applied both to PDD and brain wave approaches. See: Farwell & Donchin (1988); Honts & Devitt (1992); MacLaren & Taukulis (2000).

PDD Terminology

brachial artery

Major blood vessel located in upper arm. Occlusion blood pressure sensors are frequently placed there, and it is the preferred placement site for the blood pressure cuff in PDD.

bradycardia

Heart rate of under 60 beats per minute. Brachycardia is common among athletes and those with hypothyroidism. Slow heart rate can also indicate the influence of medications.

bradypnea

Very slow and abnormal breathing, longer cycle time. The term does not distinguish between autonomic and deliberate breathing slowing.

brain stem

Includes the adult brain structures, i.e., midbrain, pons and medulla (mesencephalon) region structures including the thalamus, third ventricle and hypothalamus. These structures are essential for the automatic control of respiration and cardiovascular systems.

breakdown (or breakout) test

PDD test in which a single issue is addressed and is always given after a multiple-issue test has indicated that the examinee has consistently responded to that issue. It is generally conducted in one of the validated test formats, such as Zone Comparison Tests or the Modified General Question Test. Question coverage can be single- or multi-faceted. The rationale for this two-stage approach to PDD testing is that, while multiple-issue screening examinations are very useful in identifying which among several issues the examinee is concerned with, they lack the power of these single-issue formats in making correct determinations of deception or truthfulness. This approach is used in many polygraph screening programs to maximize both utility and accuracy of PDD in preemployment screening and other applications.

breathing

One of the standard physiological signals in PDD testing. Respiratory data are generally obtained via a pneumograph transducer placed around the thorax and abdomen of the test subject. PDD examiners historically evaluated breathing movement data through a subjective approach that relied on the presence or absence of various signature patterns indicative of deception. Since breathing is more readily controlled than other activity recorded with the polygraph, it is one of the first areas examiners look for indications of countermeasures. Such indications are paced breathing, holding one's breath, very slow breathing, irregularly shaped waveforms, hyperventilation, and tactical use of deep breaths. In physiology, *respiration* refers to the movement of gases across membranes in the lungs, while *ventilation* is the term used for the expansion and compression of the chest during breathing. See Handler, Reicherter, Nelson, & Fausett (2009).

Brilograf

Device for measuring changes in skin resistance. Built in the 1940s by Jacques Bril, a criminologist, Brilograf was based on similar work he had done with Rev. Walter Summers on the "Pathometer." Not generally used and is only of historical note.

British One-issue Screening Test (BOST)

A variation of the screening version of the Air Force Modified General Question Test (AFMGQT). It was designed to fill a requirement for a screening method that allowed testing of only one issue. The two differences between the standard screening AFMGQT and the BOST is that in the latter the two relevant

questions cover identical behaviors and time periods, and that the decision rules include the grand total score of both questions. See Krapohl, Grubin, Benson & Morris (2020).

Calibration Verification of Sensitivity Test (CVOS)

A testing procedure conducted as the first chart, which is designed to assess an examinee's capability to process information, detect psychological or chemical countermeasures, permit adjustment of gain settings to match the response capability of the examinee, determine whether sufficient professional rapport has been established with the examinee, and reduce excessive anxiety.

card test

One of several types of stimulation tests used in conjunction with the standard PDD examination. In the traditional card test the examinee is invited to select a card from a deck of cards and then is tested on which number, letter, color, or character is on the card. While widely practiced in the early years of the polygraph profession, it has been replaced by other methods. See *stimulation test*.

cardioactivity monitor (CAM)

One of several cardiovascular sensors used in PDD. The CAM sensor is placed on the end of the finger or thumb, and it detects changes in distal blood volume via small strain gauge sensors attached to a metal diaphragm. CAMs require electronically enhanced cardiograph components. While some consider the CAM useful, it is employed less often than the traditional blood pressure cuff.

cardiograph

General term for any recording of heart activity. In PDD the use of a blood pressure cuff to monitor relative arterial blood pressure changes and pulse wave is more precisely described as *sphygmography* (recording of the arterial pulse) or *occlusion plethysmography* (partial blockage of circulation to measure volume changes in a body part). While cardiograph is not incorrect in this context, it lacks precision in denoting the actual phenomenon being recorded in PDD. The term *cardiograph* in the psychophysiological and medical literature most often refers to the electrocardiograph.

cardio-pneumo-psychograph

A two-channel polygraph developed by John Larson in the 1920s and used in criminal cases to uncover deception. See Larson (1923).

cardiosphygmograph

Alternate term for the pulse wave and relative blood pressure tracing used in PDD. While the term *cardiosphygmograph* was common parlance in the 1930s through 1950s, it is used less frequently today even though it is more precise than the current expression *cardiograph* or its abbreviated form, *cardio*.

cardiotachometer

Instrument that measures heart rate. Since heart rate can only be accurately measured over several seconds, real-time displays usually reflect the inter-beat interval that has been converted into the reciprocal to give the cardiac rate.

catacrotic limb

Descending portion of an arterial pulse wave.

category bar

One method of restricting the coverage of the comparison question so that it will not overlap the relevant question. For example, if a relevant question concerned whether the examinee had stolen anything

PDD Terminology

from an employer, a comparison question that excluded the relevant question by covering another category might involve whether the examinee had ever been dishonest with an authority figure. There is a school of thought that examinees may confuse the relevant questions with the comparison questions unless these two types of questions are designed to avoid any degree of overlap. Research has not supported this hypothesis. See: Amsel (1999); Podlesny & Raskin (1978); Horvath (1988); Horvath & Palmatier (2008). Also see *exclusive (exclusionary) comparison question*

central nervous system (CNS)

That portion of the nervous system consisting of the brain and spinal cord. CNS activity, although closely integrated with autonomic nervous system (ANS) activity, is not separately considered in traditional PDD approaches. It has been used with event-related potentials (ERPs) in Concealed Information Tests.

chart

Graphical record of phenomena. In PDD it refers to the polygram on which is recorded the physiological activity during testing. The term *chart* is sometimes used interchangeably with *test*.

chart identification

Information annotated on a PDD chart by the examiner to record identifying data such as date, time, test number, examiner, case number, signatures, fingerprints, or other details required by the polygraph program. Not to be confused with *chart markings*.

chart markings

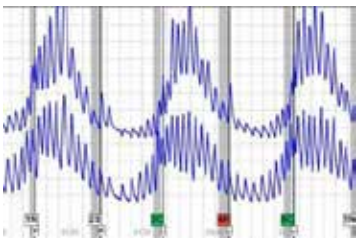
Annotation of the physiologic tracings to denote stimulus (question) onset and offset, examinee's answer, question number, question label, artifacts, and other details important to the interpretation of the physiological data.

chemical countermeasures

See *pharmacological countermeasures*.

Cheyne-Stokes respiration

Periods of cyclical variation in the amplitude of breathing cycles sometimes interspersed with periods of apnea. This pattern of breathing is usually associated with brain damage, congestive heart failure, kidney disease or drug abuse. See image below as an example.



chi-square test (χ^2)

A nonparametric inferential statistical test based on the chi-square distribution. The chi-square test is frequently applied to determining the randomness of deviations between observed and expected values. Generally, the test is used to evaluate hypotheses dealing with the relationship between two categorical variables and "goodness of fit tests." The chi-square test lacks the sensitivity of other available methods and is used primarily when the data can only be tabular in form. It is often reported in PDD literature, principally when comparing groups of test outcomes.

PDD Terminology

cholinergic
Those neurons that release the neurotransmitter acetylcholine.

classical conditioning
Characterized by the establishing of a response to a stimulus that does not normally evoke the response. The Russian physiologist Ivan Pavlov reported producing a conditioned reflex in dogs during the last part of the 19th century. He used the sight of food (unconditioned stimulus) to induce salivation in the dog (unconditioned response) and fostered a mental connection between the food and the sound of the bell, and the bell's ring became the conditioned stimulus. It is theoretically possible to classically condition physiological responses to occur uniquely when a subject is deceptive, and this method is one of the promising areas of future research in lie detection. See: Petty & Cacioppo (1981).

clearing chart
Usually, a Relevant/Irrelevant screening test used after a breakdown test and includes the remaining relevant test questions that were not covered in the breakdown test. In the field it normally involves one chart.

clinical approach
Assessment technique used in PDD that includes the use of extra-polygraphic information to arrive at a conclusion of truthfulness or deception. Also called *global analysis*.

closed-eyes technique
As the name indicates, it is the PDD examination in which the examinee is instructed to keep his eyes closed during testing. This method is used by a minority of PDD examiners, and questions about the effects of open or closed eyes have not been thoroughly investigated.

community safety examinations
A broad category of examinations that serve to detect and deter illegal behaviors that jeopardize the safety of communities. Types of community safety examinations include Post-Conviction Sex Offender Testing (PCSOT), Intoxicated Drivers on Probation (IDOP), and Domestic Violence Offender Testing (DVOT).

comparison question
Type of question used to elicit responses that are compared with the responses to relevant questions. There are two main types: directed lie (DLC) and probable lie (PLC). Subtypes for the DLC are the trivial and the personal. For the PLC they are the exclusive (exclusionary), and non-exclusive (inclusive). Historically called the *control question*, *comparative response question*, and *emotional standard*. See: Waller (2001).

Comparison Question Technique (CQT)
An umbrella term for standard testing formats that use probable-lie or directed-lie comparison questions. Included are the Reid, the MGQT, the Zone Comparison, the Positive Control, the Utah, the Arthur, the Quadri-Track, and the Test for Espionage and Sabotage. None of the following are considered CQTs: Relevant/Irrelevant, Peak of Tension, and Concealed Information Tests.

comparative response question
Name given by John Reid in 1947 for what would later be called the *control question*, and ultimately *comparison question*. See *comparison question*.

Computer Assisted Polygraph System CAPS

System developed by David Raskin and John Kircher of the University of Utah that permitted modified field analog polygraphs to interface with a computer. Signals were extracted from the pen drive motors and routed to an analog-to-digital converter, where they were digitized, stored, edited, and analyzed. Discriminant analysis was used to weigh and combine measurements, and the software produced a statement of the probability of deception. CAPS, which stood for Computer Assisted Polygraph System, was later replaced by CPS, Computerized Polygraph System, which used a computer without the analog polygraph.

Computer Voice Stress Analyzer (CVSA)

National Institute for Truth Verification produces the CVSA, first introduced in 1988. The company advertises that this device was developed from the Psychological Stress Evaluator (PSE) and is widely distributed amongst law enforcement agencies. Examiner certification is required. The CVSA testing format is dissimilar to PDD formats. Like other voice-based deception detection systems, published scientific research has failed to find any accuracy with the CVSA. See P*sychological Stress Evaluator (PSE)* and *voice stress analysis.*

Concealed Information Test (CIT)

Otherwise known as the Guilty Knowledge Test. The CIT is a series of tests, perhaps as many as 10 or more, in which there is only one critical item in each series. The tests are constructed so that the order of the item presentations is random except the first item, which is not a critical item and is used as a buffer. The theoretical operating mechanism of the CIT is there is greater signal value in the critical item for guilty examinees than in the irrelevant items. The CIT is believed to rely on cognitive processes and is therefore not subject to false positives from nervous examinees. CIT tests could be used in a small proportion of all criminal cases where sufficient details were available to construct it, however in most crimes such details are lacking or would be already known to innocent persons via the media or investigating officers. Despite assertions of theoretical superiority of the CIT over the CQT, the CIT has practical limitations that have hindered its broad acceptance among field practitioners. Moreover, the preponderance of independent research suggests that false negatives may be a problem with the CIT. See Lykken (1959); MacLaren (2001); Podlesny (1993).

conditional probabilities

A statistical concept in which the likelihood of an event is predicated on a set of conditions. Conditional probabilities are expressed as p(A | B), which may be read as: what is the probability that event A will occur given event B has occurred? More specifically, a false negative polygraph result can be represented as p(no evidence of DI responses | the person is lying). Conditional probabilities are important when characterizing the accuracy of PDD. The following illustration is one offered by critics of the use of PDD in screening. Suppose that PDD is 90% accurate in detecting both deception and truthfulness. Also, assume that it is used to test 1,000 government employees, only one of whom is involved in the activity of interest, say, treason. There is a 90% chance that the one guilty person will be caught. Of the 999 innocent employees, 899 (90%) will pass the examination, and the remaining 100 will be false positives. The ratio of true positives (1 guilty) versus false positives (100 innocent) is a low payoff if the consequences are employment termination or criminal prosecution. Not used in this example is the influence of repeated testing and other methods that can reduce false positives, but it is clear from this example that PDD validity estimates are not well represented by a single percentage. See *base rates*.

conditioned response theory

One of several theories that attempt to explain the underlying mechanisms of PDD. Conditioned response theory holds that physiologic responding is the consequence of an emotional response triggered by the conditioned stimulus. When a given stimulus is associated with strong emotions, larger responses are expected. There is some support for this theory in that physiologic responding has been established to be positively correlated with the personal significance of the test question. It does not explain why PDD continues to work in non-threatening and low motivation conditions, however. The conditioned response theory is not the prevailing explanation for PDD.

conductance

Capacity of a material to permit the flow of electricity. Skin conductance is a common measure used in PDD. A related measure, resistance, is the reciprocal of conductance.

confabulation

The reporting of information from imaginary experiences to fill in gaps of memory without any overt intention to deceive, though the information is most likely false. Confabulation can be a symptom of some organic brain disorders, though developmental factors explain other cases. The effect on deception tests has not been investigated.

confession criterion

A potential confound in field research on polygraphy. If cases are selected for research that use confession as a form of confirmation, then the study sample may be biased for the following reason. Standard practice in polygraphy is to only interrogate after a deceptive outcome on the examination. Therefore, confessions will only be obtained from examinees who failed the examination, but not from those who managed to defeat the examination. The sample will then represent those cases in which the examinee was caught by the original examiner, rather than all cases in which the examinee was deceptive. It has been asserted that the use of the confession-verified cases in blind scoring studies to assess polygraph validity may overestimate the accuracy of the polygraph, because they may have charts in which deception is the easiest to interpret. Most field studies that have examined this source of research error have not a meaningful effect, though the issue is still hotly debated. See: Horvath (1977); Honts (1996); Iacono (1991); Krapohl, Shull & Ryan (2002); Patrick & Iacono (1991); Raskin, Kircher, Honts, & Horowitz (1988).

confirmatory testing

PDD examination used to verify the statements of witnesses, victims and confidential reporting sources.

conflict theory

One of several theories that attempt to explain the underlying mechanisms of PDD. According to conflict theory, the simultaneous activation of two conflicting tendencies, such as the motivations to lie and tell the truth, results in physiologic arousal. The greater the conflict, the larger will be the response. This explanation arises from the work of Luria in the 1920s and 1930s. The conflict theory predicts that psychopaths, by virtue of a defective conscience, do not produce arousal responses as large as non-psychopaths, and this effect has been demonstrated in laboratory studies for psychopaths as a group. However, it does not explain well why phasic responses occur even when an examinee is not required to answer the question, or even when the examinee answers truthfully. The conflict theory is rarely cited as the principal explanation for PDD. See: Gardner (1937).

conspecnificance

Mnemonic device used in the instruction of PDD. It stands for consistency, specificity, and significance, three characteristics of a physiologic response indicative of deception. In order for response patterns to support a PDD outcome of deception, they must appear regularly to the same questions, manifest themselves uniquely to those questions, and be of a magnitude to be distinguishable from baseline variability. See Weir (1976).

constant current method

Measurement of skin resistance where the current applied to the skin is held constant.

constant voltage method

Measurement of skin conductance where the voltage applied to the skin is held constant.

containment approach

Criminal justice system, treatment team, and polygraph examiner working together in a team concept sharing information equally with the other members of the team. As an offspring some containment teams now include one or more of the following: law enforcement, child protective services, rape crisis centers, prosecuting attorneys, judges, and in some cases school counselors, victim advocates, and medical staff. See: Cooley-Towel, Pasini-Hill, & Patrick (2000); English, Pullen, & Jones (1996); Heil, Ahlmeyer, McCullar, & McKee (2000).

control group

The group in research that differs from the experimental group only in that the latter receives the positive manipulation pertaining to the independent variable. A control group is necessary to infer that the changes observed in the experimental group are the result of the independent variable. For example, if a researcher wished to know the effects of the drug diazapam (independent variable) on electrodermal responses (dependent variable), the group that did not receive the medication would be called the control group and the medicated group would be the experimental group.

control question

Superseded term, now call the *comparison question*. Class of questions used in deception examinations that serves to elicit larger physiologic responses from innocent examinees when compared to the relevant questions. There are several types, such as the exclusionary, non-exclusionary, probable-lie, directed-lie, the positive, and minor variations. The term "control" in PDD traces its roots to the 1930s and to what are now called stimulation tests. These tests were used as "controls" for the production of deception response patterns that would later be compared with responses to relevant questions in the Relevant/Irrelevant technique. In 1947 John Reid published a paper in which he referred to two types of questions as controls—one he called a "guilt complex" and the other a "comparative response" question, the latter being a probable-lie question. The "comparative response" question was called a "control question" in a paper published by Fred Inbau in 1948, and the name became the standard terminology in PDD for nearly 50 years. This was not the first use of this class of question, however. Walter Summers used similar questions with his Pathometer technique which he labeled *emotional standards* as early as 1939, and they were used by New York State Troopers from 1939 until at least 1952. Elizabeth Marston, widow of William Marston, and Olive Richard, Marston's former secretary, reported that they had participated in deception examinations with Marston some years before in which "hot" questions were used for comparison. A typical hot question would be, "Did you ever think of stealing money from that safe?" Elizabeth stated during an interview that they did not believe it wise to publish these types of questions, and consequently they have not been generally credited with this contribution to the science. Beginning in the 1970s, critics of PDD noted that the word "control" as used in PDD tests did not meet the criteria of the

term as used in science. The term has since been replaced by *comparison question* in publications of the American Polygraph Association, American Society for Testing and Materials, federal polygraph programs, and scientific papers. See: Waller, 2001.

comparison/control question validation test

Test procedure in which probable-lie comparison questions are tested against relevant questions from a contrived crime to theoretically verify that the subject will respond to the comparison questions. These questions are then carried into the actual testing. Proposed and taught by James Matte and used in his testing technique, though not a widespread practice nor are there any published data in support of the theory. See: Matte (1976).

control test

An alternate term for a known numbers test. See *stimulation test*.

correlation

Measure of how one variable changes with another, for example, chart scoring ability and years of experience are positively correlated. Measures of correlation range from -1.00 (indicating a perfect negative correlation—as variable *x* increases, variable *y* decreases) to +1.00 (indicating a perfect positive relationship—as variable *x* increases, variable *y* increases). A correlation equal to 0.00 means that two variables are not linearly related. One should not infer cause from correlation; two variables can be correlated without one causing the other. For example, in most cities the number of churches and the number of criminals are positively correlated, but one does not cause the other. Both are correlated to a third factor, population.

cortisol

A glucocorticoid or steroid hormone substance made from cholesterol found in the bloodstream which is produced by the adrenal cortex in response to stress. Cortisol replenishes energy stores depleted during an adrenaline rush by converting a variety of food sources into storage forms such as glycogen or fat.

counter-countermeasures

Methods used to detect and neutralize those efforts an examinee has engaged in while trying to defeat the PDD examination. For example, if a PDD examiner concluded from an examinee's behavior that the examinee had tried to dissociate during testing, he could insist that the examinee respond to the test questions with both a key word and the answer, thus ensuring attention to the content of each question. Lynn Marcy is credited with distinguishing counter-countermeasures (reactive) from anti-countermeasures (proactive).

Counterintelligence-Scope Polygraph (CSP)

PDD screening examination administered by the Federal Government on individuals with sensitive security clearances to detect and deter espionage, security breaches, sabotage, or other acts against the Government. Sometimes referred to as a *loyalty examination*.

countermeasures

Generally, methods used to mislead an observer. In polygraph research it has been labeled as actions taken by the examinee to influence the physiological responses to produce a truthful test outcome. There are several typologies for countermeasures, depending on the definition used. Under some circumstances polygraph countermeasures have been found to be effective, such as when an examinee receives special training and feedback. Most spontaneous attempts are crude and ineffective. Various methods have been devised by PDD practitioners to deter and detect countermeasures. See Honts (1987); Krapohl (1996).

CPS

Computer polygraph developed by Drs. David Raskin and John Kircher of the University of Utah. The CPS, which stands for Computer Polygraph System. Like its predecessor, the CAPS, the CPS has a discriminant analysis algorithm that weighs and combines physiological measures to calculate the probability of deception.

craniosacral division of autonomic nervous system

An anatomical division of the autonomic nervous system (ANS) that represents the sites for outflow from the parasympathetic division of the ANS, i.e., some of the cranial and sacral nerves carry parasympathetic nerves.

credibility assessment

An umbrella expression for the multiple-disciplinary field that relies on physiological and behavioral measures to test the agreement between an individual's memories and statements. Credibility assessment approaches have included reaction time tests, word association tests, polygraph, central nervous system measures, and behavioral analysis. See: Krapohl & Trimarco (2005).

critical item

In the Peak of Tension or Concealed Information Test, the critical item is the stimulus the guilty persons recognize from among the other items as being related to the event of interest. Sometimes called a "key".

current exclusive comparison question

An exclusionary probable-lie comparison question in which the scope includes the time period of the relevant issue, but is excluded from the relevant issue by category, place, or some other delimiter. See: Matte (1996).

Daubert v Merrell Dow Pharmaceuticals, Inc.

Although not a PDD case, the Daubert case set aside the Frye Rule's "general acceptability" provisions in favor of the Federal Rules of Evidence. The case paved the way for the admissibility of PDD evidence in most jurisdictions. The *Daubert* standard is a rule of evidence regarding the admissibility of expert witness testimony. Essentially it states that the judge is gatekeeper and determines what gets into evidence. The judge determines if it is relevant to this case and reliable. The judge also decides whether the evidence is based on scientific knowledge or methodology: falsifiable, refutable, and testable; subjected to peer review and publication; known or potential error rate; existence and maintenance of standards and controls concerning its operation, and; degree to which the theory and technique is generally accepted by a relevant scientific community. See: Daubert v Merrell Dow Pharmaceuticals, Inc. (1993).

deception

The act of deliberately providing or omitting information with the intention of misleading. The most critical element of the definition is the intention of the information provider. With no intent, deception does not take place regardless of the accuracy of the information being conveyed. This is an important concept to consider with regard to PDD. Inconsistencies between ground truth and examinee beliefs may erroneously appear to be mistakes of the PDD technique.

deception exemplar

Physiological response pattern which examinees expect to be used as a sample, or exemplar, of that which occurs when they are deceptive, and will be used for comparison against their responses to relevant questions. See: Matte (2000).

Deception Indicated (DI)

Along with *NDI* (*No Deception Indicated*) and *Inconclusive*, a conventional term for a polygraph outcome. A decision of DI in PDD means that (1) the physiological data are stable and interpretable, and (2) the evaluation criteria used by the examiner led him to conclude that the examinee is not wholly truthful to the relevant issue under investigation. The DI and NDI decision options are used primarily in single- issue testing, and they correspond with SR (*Significant Response*) or *SPR* (*Significant Physiological Responses*) and NSR *(No Significant Response*) or *NSPR* (*No Significant Physiological Responses*) in mixed-issue, or screening, examinations.

deception test

These methods ask directly about the matter to be assessed, are capable of addressing multiple behavioral issues of concern and may or may not depend on the existence of a known incident or known allegation. There are two broad categories of deception tests, the historically older Relevant-Irrelevant and the Comparison Question Tests (CQT). Term is used in contrast to *recognition test*.

decision rule

Generically, decision rules determine when data meet the criteria for inclusion in a particular category. Decision rules are the final steps in polygraph numerical scoring, producing categorical classifications. Optimal decision rules require the following: tracing feature selection; development of best scoring rules, consideration given for base rates; assessing and weighting collateral or countervailing information, and; performance of a cost and benefit analysis to determine the achievable level of accuracy and errors that meet the needs of the consumer. In polygraphy, feature selection and scoring rules have been thoroughly investigated. There are also decision rules in some polygraph analysis systems that include extra-polygraphic information as part of the decision process, though there is no validated method yet published. However, few published decision scoring procedures allow for consideration of the base rate issue. Also, few models publish a sufficient level of detail to allow a formal cost-benefit analysis to identify the appropriate cutting scores for a set of conditions. See: Swets, Dawes & Monahan (2000).

defensible dozen

An expression to connote the ensemble of tracing features used in chart interpretation that have replicated scientific evidence. There are four often-overlapping validated features for breathing (suppression, apnea, slowing, baseline rise), three for electrodermal (phasic response amplitude, duration, and complexity), three for cardiograph (baseline rise and duration, and pulse amplitude increase) and one for vasomotor (reduction in tracing width). The actual number of features can vary depending on how the tracings are characterized. For example, respiratory apnea is a form of respiratory suppression, but in some chart interpretation regimens they are counted separately. Moreover, there are wide individual differences in that certain examinees may present idiosyncratic features that do not generalize to other examinees.

Nevertheless, the "Defensible Dozen" is used as a convenient mnemonic device that focuses on those features that tend to be the most universal. See: Kircher, Kristjansson, Gardner, & Webb (2005); Krapohl, Stern & Bronkema (2003).

degrees of freedom (df)

For any set of values, every value within a set can be freely selected except the last, which is determined. Or, in other words, when there is only one value remaining, the final selection is not free to vary. Technically, the concept of degrees of freedom refers to the number of independent observations minus the number of parameters being estimated. The degrees of freedom are essential in the calculation of the threshold or critical value of a test distribution.

delayed answer test (DAT)

Experimental methodology used to determine whether physiologic responding was elicited by the stimulus question or an untruthful answer. It was experimentally demonstrated that the physiological arousal was more closely tied to the stimulus presentation than the act of giving the deceptive answer. The DAT is not used in field PDD examinations. See: Dawson (1980).

dendrite

Process of a neuron specialized to function as a postsynaptic receptor region of the neuron.

Department of Defense Polygraph Institute (DoDPI)

See: National Center for Credibility Assessment.

dependent variable

Variable that changes as a result of the experimenter's manipulation of the independent variable. For example, electrodermal activity could act as a dependent variable to changes in stimulus intensity, an independent variable.

devil's finger

See: *finger of death*.

Diagnostic and Statistical Manual of Mental Disorders, 5th Edition

The official system for classification of psychiatric and psychological disorders prepared and published by the *American Psychiatric Association* in 2013.

diastole

Portion of the heartbeat cycle when the heart muscles have relaxed and the chambers fill with blood. The left ventricular diastole is represented in the PDD sphygmograph tracing as the descending limb of the pulse wave (catacrotic limb).

diastolic blood pressure

Lowest blood pressure value occurring during the relaxation phase of the cardiac cycle.

dichotomization theory

A hypothesis that holds there is a difference in habituation rates for relevant and comparison question responses. According to this theory, guilty subjects habituate more slowly to relevant questions, while innocent examinees continue to respond more strongly to the comparison questions over time. Dichotomization theory has not been generally accepted in PDD. See: Ben-Shakhar, G. (1977).

dichotomous variables

Those for which there are only two mutually exclusive and exhaustive possibilities: yes or no, 0 or 1, heads or tails, etc.

dicrotic notch

A regular feature in diastolic limb of the pulse wave recorded on the sphygmograph of the polygraph. It occurs between the systole and subsequent diastole of the waveform, and its proximity to either of these two cardiac events is a function of the air pressure in the recording system. The greatest contributor to the dicrotic notch is the rebounding of the blood against the closed aortic semilunar valve after systole. It has not been found to be a reliable diagnostic feature in PDD. Also called *incisura*.

differential responsivity (reactivity)

Term frequently reported in PDD literature regarding the differences in responses to two types of questions. In Comparison Question Technique (CQT) formats, differential responsivity (differential reactivity) relates to the relative response magnitudes from relevant questions as compared to those from comparison questions. Within normal limits, this differential responsivity (reactivity) persists even when overall responsivity is attenuated, or as it habituates. For this reason, CQT formats are unlikely to render false negative decisions that are attributable to countermeasures that attack general arousal, such as the use of drugs, relaxation, or dissociation. By contrast, the Peak of Tension and Concealed Information Tests rely on differential responsivity (reactivity) between critical and neutral items. Decisions in these concealed information tests are based, not on the type of question that induces the response, but rather the presence or absence of significant responses to the critical items only. Therefore, the type of differential responsivity used in decision making in PDD, along with vulnerability to certain types of countermeasures, will depend on the polygraph technique.

differential salience

Expression that characterizes the positive correlation between the degree of psychological significance and intensity of the physiological response. The concept of differential salience is based on the premise that responsivity can reveal underlying mental processes which can be exploited to detect deception or recognition under controlled and structured conditions. It does not restrict the operating mechanism to fear, but assumes a common pathway for the physiological expression of those cognitive and emotional processes that gave rise to the psychological salience. It is proposed as a substitute for the older "Psychological Set" hypothesis. See: Handler & Nelson (2007); Handler, Shaw & Gougler (2010); Senter, Weatherman, Krapohl & Horvath (2010).

directed-lie comparison (DLC) question

Type of question used to elicit a response that is compared with the response evoked by the relevant question. The DLC question is different from the probable-lie comparison (PLC) question in that the examinee is instructed by the examiner to answer the DLC question untruthfully, whereas the principle of the PLC requires the examiner to lead the examinee to be untruthful to that question without revealing the purpose. DLCs can be further delineated into the *trivial DLC* and the *personally significant DLC*, which, as their names indicate, depend on the content of the DLC. The true strengths of DLCs are that they can be standardized much easier than the PLCs, they are less intrusive, and their effectiveness is less subject to examiner skill. DLCs are being used in many quarters of the PDD profession. See: Horowitz, Kircher, Honts, & Raskin (1997); Raskin & Honts (2002).

Directed Lie Screening Test (DLST)

A screening test for law enforcement based on the technique and procedures for the Test for Espionage and Sabotage (TES). The DLST uses a repeated series of two relevant and two directed-lie comparison questions, and the conventional 7-position scoring system. See: Blalock (2009); Handler, Nelson & Blalock (2008). Reed (1994); Research Staff (1995,1998).

disclosure examination

See: *sexual history examination*.

discontinuous blood pressure method

Deception test procedure developed by William Marston before 1915. Marston's instrumentation was a standard sphygmomanometer that he used to take intermittent systolic blood pressure measurements during questioning on relevant and irrelevant topics. He plotted these measurements by hand, creating a curve that was interpreted for assessing deception. In 1923 Marston attempted to have the results of his deception test entered into evidence in a murder trial in Washington, DC. The Frye case, which was the first to consider deception tests, established the precedent for exclusion of "lie detector" results. The discontinuous blood pressure method did not enjoy widespread field acceptance, and there are no reports of its use after the 1930s. In the 1920s William Marston included a cardio-pneumo polygraph to augment his discontinuous blood pressure method. In practice, Marston and his wife, Elizabeth, would either ask the examinations questions or take the blood pressure measurements, while Olive Richard, an assistant, operated the equipment. If a stenographer were present, there were four participants in the administration of the examination in addition to the examinee. While William Marston was usually the examiner, Elizabeth Marston and Olive Richard did conduct examinations on occasion without him, making them the first women in this field. Given the great methodological and instrumentation differences, Marston discontinuous blood pressure method is not truly in the lineage of modern polygraphy, though it is frequently included in history lessons at polygraph schools. See: Marston (1917; 1938).

discovery test

See: *disclosure test*.

discriminant analysis

Regression analysis with a categorical criterion—that is, attempting to predict group membership from one or more predictor variables. The CPS algorithm is based on discriminant analysis.

disguised comparison question

Comparison questions that are designed to be difficult to identify as such. Typically used in multiple-issue screening examinations. For example, the examinee might have the following disguised comparison question: Did you ever cheat in school? It is expected that nearly every examinee would have doubts about the integrity of his entire academic career, and also believe that truthfulness on the question is required to pass the polygraph examination. Sometimes call *hidden comparison question*.

dissociation

Psychologists use the term *dissociation* to denote largely unconscious processes by which normal relationships in thought, memory, attitudes, or other psychological activities do not adhere to their established relationships and become separate or independent. It is also used in PDD to characterize the deliberate disengagement of attention by examinees from the testing situation. Such tactical redirection of attention is considered a mental countermeasure. Examinees who use this approach are hopeful that they may eliminate physiological responding by virtue of their mental distance from the test questions. Dissociation may be effective in test formats where examinees respond with the same answer to each

question, such as with the Peak of Tension. In that format the examinee needs only listen for the point when the examiner stops speaking and give the rote answer. Dissociation is considered a more unlikely countermeasure for those formats that require both yes and no answers, and where the order of the questions is not predictable to the examinee, or where examinees are required to use key words from the test question in the answer. See: Elaad & Ben-Shakhar (1991); Kircher, Woltz, Bell & Bernhardt (2006).

domestic violence offender testing

A program intended to incorporate the polygraph in an attempt to verify compliance with the conditions of probation for convicted domestic violence offenders. As with post-conviction sex offender testing (PCSOT) and intoxicated drivers on probation (IDOP), it is designed to contribute to community safety by detecting and discouraging behaviors that pose a threat to the public by individuals with a demonstrated propensity to engage in certain criminal acts.

double verification test

Alternate term for *card test*. See: *stimulation test*.

dyspnea

Labored or difficult breathing, generally resulting from disease.

Easterbrook Hypothesis

The premise that attentional resources are more restricted as the level of arousal increases. See: Easterbrook (1959).

eccrine glands

One of two types of sweat glands, the eccrine glands influence electrodermal activity as measured in PDD. They are found throughout the skin surface of the body, but in highest concentration on the hands and feet. See: Handler et al. (2010)

efferent nerves

Neurons that carry nerve impulse from the central nervous system to the effector organ or muscles. Also called *motor nerves*.

Either-Or Rule

PDD scoring rule forwarded by Cleve Backster and used exclusively in the Backster Zone Comparison Technique. According to the rule, if a relevant question does not evoke a physiological reaction, it is scored against the adjacent comparison question with the larger reaction. If the relevant question does produce a significant reaction, it is compared to the comparison question with the smaller reaction. See: Matte (1996); Meiron, Krapohl & Ashkenazi (2008).

electrocardiogram (EKG or ECG)

Tracing of the electrical activity of the heart. This endosomatic waveform consists of the P, Q, R, S, T, and U waves. The search for diagnostic information in the ECG for PDD purposes has not been fruitful to date. There are preliminary data that suggest that the pre-ejection period (PEP), which is derived from the ECG and the impedance cardiograph (ICG), is a reliable gauge of sympathetic nervous system arousal. Inter-beat interval has also shown promise. See: Kircher, Packard, Bell & Bernhardt (2003).

electrodermal activity (EDA)

All exosomatic and endosomatic changes in the electrical properties of the skin. See: Handler et al. (2010).

electrodermal response (EDR)

Reaction of skin measured by changes in its electrical properties, including skin resistance (SR), skin conductance (SC), and skin potential (SP). See: Handler et al. (2010).

electroencephalogram (EEG)

Recording of the electrical activity of the brain generated by the firing of clusters of neurons. In recent years EEG methodology has been applied to deception and concealed information tests. See: *P300*.

electromyograph (EMG)

Tracing of the endosomatic electrical properties of the voluntary muscles. This activity is recorded through sensors placed on the skin near the muscles of interest. EMG could be used for the detection of physical countermeasures when the sites are correctly chosen by the examiner.

electrooculograph (EOG)

Recording of the electrical activity produced during eye movements. EOGs have had two principal uses in deception testing. One is as a deception indicator. Some research has shown lateral eye movements have diagnostic information useful in assessing whether a person harbors concealed information. The second application is with evoked cortical potentials, which also have been used in concealed information paradigms. Brain wave activity generates very small voltages, and eye movements generate electrical potentials that interfere with these signals. EOGs are often collected for subtraction from the brain wave signals

embarrassing personal question (EPQ)

A question, frequently with a sexual theme, sometimes used with the Keeler Relevant/Irrelevant test in the late 1940s and early 1950s. Leonarde Keeler experimented with the embarrassing personal question, hypothesizing that guilty subjects would not respond to it, remaining instead focused on the relevant questions, while the innocent examinees would produce significant responses in the opposite pattern. The embarrassing personal question did not have widespread use and was not taught at Keeler's school after 1951.

emotional standard

A term coined by Rev. Walter Summers in his research into lie detection. The emotional standard was an emotion-provoking question to which the examinee answers truthfully, but one that the examinee would prefer to hide. It was included in a test series so the reaction evoked by it could be compared with the reaction elicited by relevant questions. Summers' test format included an established ordering of pairs of relevant and emotional standard questions, interspersed with irrelevant questions, as needed. It is the first report in the literature of this type of question, and it predates Reid's "comparative response question". See: Summers (1939).

empirical

An approach based entirely on observation rather than speculation. Much of PDD research is empirical in nature, though several theories exist.

Empirical Scoring System

The Empirical Scoring System (ESS) is an evidence-based numerical scoring model for manual test data analysis (TDA) of PDD test data from examinations conducted using comparison question test (CQT) formats. The ESS includes a description of the physiological data features that are correlated with truth and deception, mathematical transformation methods for assigning and aggregating numerical scores, decision rules for the classification of numerical scores as indicative of truthfulness or deception, and numerical

cutscores that define the *a priori* thresholds of statistical significance. ESS cutscores are based on normative data that allow for calculation of the probability of an erroneous test result. Thus, the ESS allows for the selection of statistically optimal cutscores based on operational needs for the resolution and precision of the test result. See: Blalock, Cushman & Nelson, (2009); Handler, Nelson & Blalock (2008); Nelson, Krapohl & Handler (2008).

Employee Polygraph Protection Act of 1988 (EPPA)

US legislation that restricts the use of "lie detector tests" by private employers except under specified conditions. Employers may not compel or request employees or applicants to submit to such testing, nor may they use any results for adverse action. All levels of government are exempt from the provisions of EPPA. There are also exemptions for companies that provide security services and those involved in the manufacture and storage of controlled substances, who may use the polygraph for preemployment screening. Employers may request an employee to undergo PDD testing if it is part of an investigation of a loss to the employer, the employee had access to the property lost, there is a reasonable belief that the employee was involved in the loss, and the employee is given 48 hours notice prior to the examination that outlines the loss, investigation, and the reasons the employee is under suspicion. The employer is prohibited, however, from taking any action against the employee who refuses to cooperate with a PDD examination. Two of the major effects of EPPA have been a sharp decrease in the number of private examiners in the United States, and a move by the professional polygraph associations to upgrade standards of practice. Full text of the law can be found on the Department of Labor website: www.dol.gov.

endosomatic

Something produced from within the body itself. One type of electrodermal response, skin potential response, is produced by electrical activity generated by the dermis. Its measurement requires the placement of one electrode over an area well supplied with sweat glands (active site) and the other over an area devoid of them (reference site). The active site is negative in relation to the reference site by an amount that varies from a few to 50 to 60 mV. An alerting stimulus generally produces an increase in negative potential, followed by a positive wave, usually commensurate with the production of surface sweat, and sometimes a second negative wave. Similarly, EEG signals are generated by bioelectric processes in the brain, and EKG from the heart. For contrast, See: *exosomatic*.

epinephrine

A hormonal stimulator of the sympathetic nervous system. It acts to constrict peripheral blood flow, raise blood pressure, increase cardiac activity, promote metabolic activity through the release of glucose, and inhibit digestive processes. Epinephrine is considered a psychogenic hormone because it alters psychological processes when released in large quantities, such as under stress. It is produced in the adrenal medulla, located immediately above each kidney. Called *adrenaline* in British reports.

error-related negativity (ERN)

Brain wave time-locked to response selection that corresponds with incorrect choices. One of several electrocortical phenomena being investigated as a deception detection measure.

eupnea

Normal quiet breathing.

event marker

Annotations, markings, or abbreviations placed on the chart to alert a reviewer of a significant event. Examples include; start and stop of examination announcement, reading of question, examinee's answer, movements, talking, deep breathes, etc. Many event markers have generally accepted universal meanings.

event-related potentials (ERPs)

A change in electrical activity of the brain in response to a stimulus, recorded as changes in voltage at the scalp surface. Current signal processing approaches allow averaging of EEG activity, and ERPs are extracted by the averaging of brain waves over several repetitions of stimulus items. ERPs have been useful to scientists as markers for specific processes in the brain. There are several types of ERPs: the N100, N200, P300, and N400, to name a few. The "N" and "P" designators are conventions for the polarity of the signal, negative and positive. The number denotes the latency after stimulus presentation, measured in milliseconds. The P300 has been reported to be a dependable indicator of concealed information, though its application to deception detection is not established. ERPs are generated by the central nervous system, and their use in deception tests is attractive because they are expected to be resistant to countermeasures.

evidence-connecting question

Test question in which the examinee is asked about a particular piece of physical evidence that would incriminate the guilty person. It could be items left at the crime scene by the perpetrator, stolen property that could be discovered in his possession, or even doubts about leaving incriminating fingerprints. For example, if a PDD examination was being administered to resolve a fraudulent use of a credit card, a test question could center on the signature on a receipt, the possibility of a photograph being taken at an ATM where the card was used, or possession of property wrongfully obtained by use of the card. The evidence-connecting question could be more salient to the examinee than the "did you do it" relevant question because the examinee knows it can lead to physical evidence that will implicate him.

evidentiary decision rules

Decision rules proposed by Krapohl (2005) that begin with asymmetric cutting scores for 7-position scoring: if the grand sum of scores is -6 or lower, the call is DI; if the grand sum of all scores is +4 or greater, the call is NDI. In those cases where the grand sums ranged from -5 to +3, the sub-totals (Spot Scores) are evaluated. For those cases, if a single relevant question has a sub-total of -3 or below, the decision is DI. All other cases are called Inconclusive. See: Krapohl (2005); Krapohl & Cushman (2006).

evidentiary examination

A polygraph examination in which the written and stated purpose agreed to by the parties involved is to provide a diagnostic opinion as evidence in a pending judicial proceeding.

evoked cortical potentials

Brain waves that are induced by stimuli controlled by the experimenter.

examination

The entirety of the PDD process, including pretest, test, and posttest elements, from onset to completion.

exclusive (exclusionary) comparison question

Probable-lie comparison questions that do not overlap the event covered by the relevant issue questions. There is a school of thought that examinees may confuse the relevant questions with the comparison questions unless these two types of questions are designed to avoid any degree of overlap. This is accomplished by constructing comparison questions that are different from the relevant issue by time period, location, or type of activity. Although exclusive comparison questions have better face validity over non-exclusive comparison questions, research has supported the non-exclusionary version. See: Amsel (1999); Podlesny & Raskin (1978); Horvath (1988); Horvath & Palmatier (2008).

exculpatory examination

A PDD examination offered to an accused against whom other strong evidence exists. The exculpatory examination is often used in the military services when urinalysis has indicated a service member has used an illegal drug. The service member is not obligated to undergo the PDD examination with the military investigative services, but because adverse action can be taken based solely on the urinalysis, many avail themselves of the opportunity. Exculpatory exams are so named because they are intended to offer an opportunity to present evidence to support one's assertion of innocence.

exosomatic

Something generated from outside the body. Electrodermal recordings that apply a voltage or current to the skin are called exosomatic and in polygraphy a direct current (DC) is used to measure aspects of EDA. Constant voltage DC systems record EDA as skin conductance (SC) for which the units are Siemens (S) or mhos, which is the inverse of ohm in both spelling and in computation. Constant current systems measure and record skin resistance (SR), which is measured in ohms. EDL is the accepted abbreviation for electrodermal level and refers to the tonic or baseline level at any given moment, while EDR is reserved for the phasic response or reactions to stimulation. The designators R and L may be appropriately applied to the type of measurement taken, for example SRR (skin resistance response) or SCL (skin conductance level). Both skin conductance and skin resistance are exosomatic measures because electrical currents are applied from outside sources to detect the electrodermal activity. As opposed to *endosomatic*. See: Handler, Nelson, Krapohl & Honts (2010).

ex parte phenomenon

Literally, from the Latin meaning taken from one side or party. Legal term that has been used to indicate the "friendly polygrapher" hypothesis. See: Orne (1973).

experimental group

In research, subjects fall into two broad categories: *experimental* and *control groups*. The experimental group is subjected to the independent variable—that is, the variable of interest to the experimenter. The control group is treated exactly the same, except that it does not receive the independent variable. When the dependent variables of the experimental and control groups are compared, their differences are attributable to the independent variable. Consider research examining the validity of PDD: one group would be assigned to the innocent condition, and the other to the guilty condition. Since the variable of interest is the detection of deception, the guilty would be the experimental group and commit the mock crime, and the innocent would not commit the mock crime and would be the control group.

exploratory test

Zone Comparison Technique test format for dealing with multiple issues.

extrapolygraphic

That which is not derived exclusively from the polygraph waveforms or tracings. Some polygraph schools teach that there are sources of information to assist the polygraph examiner in rendering a decision that are not registered in the physiological data. These sources of extrapolygraphic information include case facts, behavioral indicators, and base rates. Blind interpretation of polygraph charts is one way of parsing out what information is available in the test recordings and that which comes from other sources.

extrasystolic beat (ESB)

See: *premature ventricle contraction*.

false key

A term coined by Arther (1970) for the deliberate placement of a false distracter item among other items on a Peak of Tension test that the examiner has hinted is the correct item. Arther reported it to be a useful diversion of attention of the innocent examinee who does not know the true key or critical item. The false key is the most plausible item in a list for the naive examinee, though it is by design the incorrect item. The false key is used exclusively in Known Solution Peak of Tension tests. No published research is available on the use of the false key.

F3

See: *Fight, Flight and Freeze*.

false negative

The failure to detect the presence of a particular event or item. A false negative in PDD refers to the incorrect decision that deception was not practiced by the examinee. Also called a Type-2 error.

false positive

The false detection of something that is not actually present. In PDD, it is the incorrect decision that deception was practiced by the examinee. Also called a Type-1 error.

fear of detection model

One of several theoretical explanations for the psychophysiological mechanisms underlying arousal during deception. According to this 'concern-based' model, examinees physiologically respond to test questions to which they are lying out of concern that their deceptions will be detected and adverse consequences will follow. The greater the fear, the greater will be the response. While incomplete and unproven, this remains a prevailing theory taught to PDD practitioners because it appears to have some face-validity. Unfortunately, it ignores the large body of scientific literature dealing with cognition, emotion and behavioral conditioning. One obvious exception relates to those instances where the PDD test continues to be effective even when there are no or trivial consequences for detection, or when directed-lie comparison questions are used.

fear of error

Concept forwarded by James Matte to account for a portion of false positive errors in polygraphy. According to the theory, the innocent examinee is inclined to physiologically react to relevant questions if he is excessively concerned about a polygraph error. To correct this confound, Matte advocates the insertion of an "inside track" pair of questions among the test questions. Empirical support for the inside track is not yet available. See: Matte (1996); Nelson & Cushman (2011).

feature

In polygraphy, the term refers to a specific aspect of a waveform, pattern or measurement in a tracing. Features are the fundamental components of chart interpretation on which scoring and decision rules depend. Currently there are about 12 individually validated manual scoring features. In the breathing channel they are: apnea, baseline increase, suppression, and increase in cycle time (slowing). For the electrodermal channel they are peak amplitude, complexity, and duration. In the cardiograph, the features are amplitude and duration. The finger plethysmograph relies on the duration and magnitude of the constriction of the pulse amplitude. Other features are sometimes taught as part of scoring systems, though their validity is disproven or absent. See: Kircher & Raskin (1988); Bell, Raskin, Honts & Kircher (1999).

field research

Scientific investigation using actual PDD cases conducted by practicing examiners on suspects, witnesses, and victims. In contrast to *laboratory research*.

fight, flight, freeze

Three stereotypic behavioral responses to threat, sometimes simply called *F3*. The physiological responses concomitant to these behaviors are the same, namely mobilizing bodily resources for an expenditure of energy, and narrowing attentional focus to the features of the threat. This preparation activity of the body has been used as a rudimentary explanation for the pattern of arousal responses that are recorded during PDD. Handler and Honts (2007; 2008) offered an alternative based on the Behavioral Inhibition System theory proposed by Gray and Mc Naughton (2003).

finger of death

Somewhat whimsical informal expression for a tracing pattern sometimes found in the electrodermal channel that is putatively associated with deception. It is the sudden plunging of the electrodermal tracing shortly after the presentation of a relevant question followed by a normal return to baseline, creating the visual impression of a "finger." The sudden drop and recovery may or may not have been preceded by a phasic response. Some writers have attributed the phenomenon to a loss of contact between the sensors and the skin, such as when certain types of physical countermeasures are practiced. The cause of the phenomenon is not well understood, and it has not been scientifically established as a reliable indicator of deception. Sometimes called the *devil's finger*.

finger pulse amplitude (FPA)

Cardiographic measure of the pulse wave recorded by plethysmograph (both occlusion and photo type) at the finger. Constrictions in amplitude are associated with sympathetic nervous system arousal. See: Handler & Krapohl, (2007); Kircher & Raskin (1988).

Fleiss' kappa

Statistical measure for the degree of agreement among multiple raters for their classifications of items. In PDD, it provides a metric for the reliability of decisions among different scorers interpreting the same test charts and is the preferred method for gauging inter-rater agreement. See: Fleiss (1971).

foil

An irrelevant item in a Concealed Information Test. Sometimes called *padding*, *buffer*, *control*, or *non-critical item*.

forensic psychophysiological detection of deception examination

A process that encompasses all activities that take place between a forensic psychophysiologist and an examinee during a specific series of interactions. These interactions include the pretest interview; the use of the polygraph to collect physiological data from the examinee while presenting a series of tests; the diagnostic phase, which includes the analysis of physiological data in correlation with the questions asked during each test to support a diagnostic decision; and the posttest phase, which may or may not include interrogation of the examinee. See: Yankee (1995).

forensic psychophysiologist

Proposed alternate title for polygraph examiner. It is a person who has successfully completed an academic program in Forensic Psychophysiology, including the appropriate internship, which has been inspected and accredited by the American Polygraph Association.

forensic psychophysiology

Defined by Dr. William J. Yankee in 1992 as the science that deals with the relationship and applications of PDD tests to the legal system. It is the academic discipline that provides the student, the practitioner, and the researcher with the theoretical and applied psychological, physiological, and psychophysiological fundamentals for a thorough understanding of PDD tests, and the skills and qualifications for conducting PDD examinations. The modifier "forensic" delineates and delimits this discipline from the broader discipline of psychophysiology. See: Yankee (1992).

format

A particular order of question presentations, or rules that govern the order, along with the types of questions. "Format" is sometimes incorrectly used interchangeably with "technique," a broader term that encompasses not only the format, but all practices in the pretest and test phase.

frame of reference

The circumstances or facts (crime report, criminal complaint, victim allegation, etc.) presented to the polygraph examiner which form the basis for the PDD examination. See: Holden (2000).

"friendly polygrapher" hypothesis

A hypothesis proposed by Martin Orne that a deceptive examinee would not be as detectible by an examiner who conducts a polygraph examination on behalf of the examinee's attorney because the examinee has no fear of adverse consequences. All field studies that have investigated it have failed to find the effect. See: Honts (1997); Ishida & Sevilla (1981); Matte & Reuss (1990); Orne (1973); Raskin (1976).

functional Magnetic Resonance Image (fMRI)

An image of the brain processes created from the metabolism of neurons. There are three types of fMRI imaging, which use blood flow, blood volume and blood oxygenation level-dependent (BOLD) signals. The BOLD method is the most common method used as it has the highest functional contrast. As the brain regions are engaged in a task they require more blood. The fMRI is an image of changes in blood flow based on task demands. The fMRI is one tool being researched to find central nervous system indicators of deception. See: Kozel, Johnson, Grenesko, Laken, Kose, Lu, Pollina, Ryan & George (2009); Pollina, Horvath, Denver, Dollins & Brown (2008).

Functional Near-Infrared Spectroscopy (fNIRS)

An approach to neuroimaging in which lasers emit infrared light through the skull into the brain and light-detecting sensors record the light that returns. The absorption and reflection of the infrared light can indicate where in the upper brain regions oxygenated red blood cells are concentrated, thereby allowing inferences for those areas that are active during a given task. fNIRS has been tried experimentally in deception detection. See: Butta et al., 2015; Li et al, (2018); Tian et al. (2009).

Galvanic Skin Response (GSR)

A superseded term for the electrodermal response measured exosomatically by the change in the electrical resistance of skin. GSR is sometimes erroneously called Galvanic Skin Resistance or Galvanic Skin Reflex. The modern term is *electrodermal response (EDR)*.

galvanograph

Polygraph component responsible for producing the graphic recording of skin resistance.

ganglion

A cluster of nerve cell bodies. (pl. *ganglia*).

general state countermeasures

Attempts to defeat the polygraph examination by influencing tonic physiological activity, or altering phasic lability. Typical approaches include the use of drugs, meditation, biofeedback, and fatigue. The goal of state countermeasures is to diminish the body's responses to all polygraph questions. State countermeasures may affect testing techniques that rely on the presence or absence of responses to diagnose deception, such as the Concealed Information Test or Peak of Tension tests. Because comparison question tests use differential responsivity to different types of questions, state countermeasures are more likely to result in inconclusive findings than errors. See: Honts & Amato (2002).

generalizability

Extent to which a set of research results can translate to other research paradigms or to the real world.

general nervous tension (GNT)

Expression used in the practice of PDD to characterize recorded physiological patterns that suggest the examinee's basal level of arousal is high. This arousal is not indicative of deception in itself. GNT is sometimes indicated by very fast heart rates, unusually labile electrodermal activity, and uneven breathing cycles. PDD examiners try to bring examinee's arousal state to a median level to optimize the interpretability of the test charts.

general question technique

Alternate expression for the Keeler Relevant/Irrelevant Technique.

global analysis

Evaluation of the polygraph recordings as a whole, as opposed to making systematic comparisons among questions. Global evaluation can also represent the use of extra-polygraphic information such as examinee behavior and case facts when rendering a polygraph decision, an approach championed by Reid and Arther. When information beyond the physiological tracings is considered to produce the final outcome, it is also called the *clinical approach*.

green zone

Term used by Cleve Backster to describe a 20- to 35-second block of polygraph chart initiated by an exclusionary comparison question which has a unique psychological focusing appeal to innocent (truthful) examinees. See: Backster (1963c).

ground truth

Reality. In the PDD context it is the veridical state of truthfulness or deception against which polygraph outcomes are compared in validity studies. Ground truth is an elusive feature in field studies because it is difficult to independently verify guilt or innocence in many cases. In laboratory studies, it is delineated into programmed guilty and programmed innocent groups.

GSG

Expression by a polygraph manufacturer to represent a measure of skin conductance. An adaptation of GSR, substituting the letter "R" with "G," the engineering shorthand for conductance. However, GSR stands for Galvanic Skin Response, not Resistance. The phenomenon called GSG is more correctly denoted as *skin conductance (SC)*.

guilt complex reactor

Hypothetical personality trait that causes innocent examinees to physiologically respond to any question that they consider accusatory. Guilt complex questions have been used in many of the contemporary formats at one time or another in an attempt to identify those examinees who would produce a false positive outcome because of this tendency. No empirical support exists for the existence of guilt complex examinees nor for the benefit of using a test question aimed at identifying them.

guilt complex test

A PDD test format in which an examinee is tested on a fabricated crime. The guilt complex text has several hypothetical purposes, primarily in avoiding false positive outcomes. The guilt complex test was taught in the earliest years of the Reid and Keeler schools. See: Abrams (1977).

Guilty Knowledge Test (GKT)

A test published by Dr. David Lykken and is based on a concealed information paradigm. While similar tests are described in the literature as early as 1904 (Wertheimer & Klein), and Hugo Munsterberg outlines a comparable approach in his 1908 book *On the Witness Stand*, Lykken formalized the procedures and advocated its use in place of the CQT. Recent writers have renamed this method the *Concealed Information Test* (CIT). See: Lykken (1959); Verschuere, Ben-Shakhar & Meijer (2011).

habituation

Adaptation to a stimulus over time. As an organism habituates to a stimulus or environment, its response diminishes both in intensity and frequency. In PDD, habituation has been found within tests, but little or none between tests. See: Dollins, Cestaro, & Pettit (1998); Kircher, Raskin, & Honts (1984).

halo effect

Tendency of an observer to be unduly influenced by a single trait of an individual. This term was coined by Thorndike in 1920 in the context of psychological assessment. For the PDD examiner, it is a potential source of error if examineee-examiner interactions are factored into the final PDD decision. See: O'Sullivan (2003); Thorndike (1920).

heart rate

Rate of ventricular contractions, usually measured in beats per minute. It is one index of physiological arousal. Some recent research indicates that after stimulus onset cardiac arousal takes the form of an immediate decrease in heart rate if the response is an orienting response (OR). Heart rate and the interbeat interval are reciprocals of one another.

hertz

Term for frequency, in cycles per second. For example, a heart rate of 80 beats per minute would equal 1.33 hertz. Frequency measures in psychophysiology are often reported in hertz, particularly when identifying engineering specifications of instrumentation. Named for German physicist Heinrich R. Hertz. Sometimes called *cycles per second (cps)*.

Hg

Chemical symbol for the element mercury. Millimeters of mercury is the reference for measures of pressure, such as barometric and blood pressure. Conventional polygraph notation for air pressure in the sphygmomanometer is gauged in millimeters of mercury (i.e., 72 mm Hg). Hg stands for *hydrargyrum*, from Greek for water and silver.

hidden comparison question

Question designed to evoke a response from a truthful person, but appears to be relevant to the examinee, and therefore its true purpose is concealed. Useful for testing victims or those knowledgeable in CQT formats. Sometimes called *disguised comparison question*.

hidden key

Critical item in the Known Solution Peak of Tension test. It is called *hidden* because it is not known to be the critical question to the innocent examinee, and it is embedded in a list of apparently similar questions. There is one key per test.

Hobson's Choice

An expression referring to an apparently free choice that offers no genuine alternative. It was named after Thomas Hobson, a stable owner in the 16th century, who offered patrons the horse nearest the door, or none at all. For Hobson's customers, there was the illusion of choice, but no actual options. Hobson's Choice is used in polygraphy when the probable-lie questions are developed in the pretest interview. The examinee feels as if he or she must pass this question to pass the examination. During the pretest interview the question is presented and refined until the examinee chooses to deceive rather than to accept the much less desirable option of acknowledging socially proscribed behaviors. Truthfulness is not a true choice in that circumstance, and therefore the examinee's decision to lie is based not on a free choice but on a Hobson's Choice. The lack of alternatives or "escapes," which is associated with a state of "learned helplessness," may be a mechanism in the arousal level. See: Vendemia (2002).

homeostasis

Homeostasis is a term in physiology to describe the maintenance of the internal viability of organisms within a prescribed range. The word homeostasis is derived from the Greek homeo, means "same," while stasis means "stable;" thus, "remaining stable by staying the same." Walter Cannon coined the term "homeostasis" from a related idea developed earlier by Claude Bernard. Claude Bernard declared "All the vital mechanisms have only one object, to preserve constant the condition of the internal environment." Studies in physiology and medicine have interpreted that statement to mean certain aspects of the internal milieu are fixed at a specific set point. The historical concept of homeostasis is the basis of

modern concepts of autonomic regulation and control. Organisms are always in a state of homeostasis, even during states of arousal, except during illness or disease. Also see *allostasis*.

hope of error

Concept introduced by James Matte, and a central component of his Quadri-Track Technique. Because guilty examinees usually stand to lose something of importance if their deceptions are uncovered by the polygraph, Matte argues that they are hopeful that there will be an error in the outcome. A challenge to Matte's hypothesis is that truthful subjects are also deceptive during testing - to probable lie comparison questions - and they too might be hopeful for an error to occur. During testing Matte includes a direct question regarding the examinee's hope of an error and scores the question as a relevant question. See: Matte (1996); Matte & Reuss (1989); Nelson & Cushman, (2011).

Horizontal Scoring System

A method devised by Gordon and Cochetti in the 1980s. All responses within each channel are ranked from largest to smallest; ranks assigned to comparison questions are given positive values, while those to relevant questions receive negative values. For example, if a test had three each of relevant and comparison questions, and the magnitude of the responses in a given channel resulted in an order of R3, R1, C1, R2, C3, and C2, their values would be designated as -6, -5, +4, -3, +2, and +1, respectively. This method is repeated for all channels in all tests and then summed for a grand total. Thresholds suggested by Gordon and Cochetti were two points per relevant question per test, and a minimum of two tests. Because of the ranking approach, this scoring system may be limited to single-issue testing situations. Additionally, some of the diagnostic criteria and transformation procedures have not been shown to be empirically supported. See: Gordon (1999); Gordon & Cochetti (1987); Gordon, Mohamed, Faro, Platek, Ahmad & Williams (2005); Krapohl, Gordon & Lombardi (2008); Nelson & Handler (2011).

hydrosphygmograph

Device used by Cesare Lombroso at the end of the 19th century to detect changes in blood pressure during deception, though the hydrosphygmograph that had been invented years earlier for medical purposes. It consisted of a container of water and a rubber seal through which an examinee's fist was placed into the water. Once the container was sealed, changes in relative blood volume changes were transferred to the closed system and could be recorded with tubing leading to a recording pen that wrote on a smoked drum. This is the first mechanical device reported in the literature used specifically for deception tests. See: Trovillo (1939).

hyperventilation

Increase in rate and depth of breathing.

hypnosis

Altered state of consciousness in which the subject is very receptive to suggestion and direction. Hypnosis has been a concern to PDD practitioners because it is thought to be a possible undetectable countermeasure. In a highly suggestible state, guilty subjects could conceivably have memories of their crimes blocked, altered, or replaced so that physiologic responsiveness would be unreliable for diagnosing deception. It could also be used to enhance desensitization training, or autonomic conditioning. The little research on hypnosis has not conclusively settled the issue. See: Weinstein, Abrams, & Gibbons (1970); Timm (1991).

hypothenar eminence

Prominence on the palm corresponding with the musculature of the little finger. One of the most productive recording sites, along with the thenar eminence, for electrodermal activity. See: Handler, Nelson, Krapohl & Honts (2010).

Inbau, Fred

Inbau (d. 1998) is most known in the polygraph community for his collaboration with polygraph pioneer John Reid. Though a lawyer, Inbau joined the new established Scientific Crime Detection Laboratory in Chicago in 1933 to pursue his interest in forensic science and stayed with the lab as director when it was assumed by the Chicago Police Department. He left to be a trial lawyer in 1941 and joined the faculty of Northwestern University School of Law in 1945. Inbau was a prolific writer, and his book *Criminal Interrogation and Confessions* is considered a classic.

imagery

The use of visualization to experience memories or fantasies. Imagery has been shown to produce profound physiologic responses, and because it can be performed covertly by an examinee, it is a concern to PDD examiners as a possible countermeasure. Imagery is one form of *dissociation*.

impedance cardiogram (ICG)

Specialized cardiogram by which the timing and stroke volume of the heart can be derived. Though not currently used in PDD, it has been shown to provide a gauge of sympathetic arousal when used in tandem with the ECG. See: Harrell & Clark (1985).

incisura

A notch or indentation on any form. See *dicrotic notch*.

inclusive (inclusionary, non-exclusionary) comparison question

Comparison question that potentially encompasses the activity of interest in the relevant questions. While contemporary practice tends to favor exclusionary comparison questions, no research has not found them to increase decision accuracy over inclusive comparison questions, and most studies support the inclusive comparison question. See: Amsel (1999); Podlesny & Raskin (1978); Horvath (1988); Horvath & Palmatier (2008).

Incomplete

PDD outcome used in some sectors that indicates that testing was terminated before sufficient physiological information was collected. This may be due to the sudden onset of health problems, extreme emotional distress, or the examinee's unwillingness or inability to remain for further testing. It may also signify that the examinee provided information after initial testing that necessitated subsequent testing, but it was not completed due to examinee fatigue, time limits, or equipment problems. A PDD decision of incomplete implies that testing may continue at a future date.

Inconclusive

PDD outcome where testing was completed, but neither deception nor truthfulness can be diagnosed because the physiological data are inconsistent, inadequate, artifacted, or contaminated. There is disagreement whether an inconclusive outcome should be considered an error when computing validity of PDD. Some argue that examinees are either truthful or deceptive, but never inconclusive; therefore, such an outcome is necessarily in error. Conversely, in the forensic sciences it has been asserted that the inconclusive outcome is used to assess utility, but not validity, because samples in forensic disciplines are often inadequate, or contaminated. Because of this controversy, PDD validity studies report accuracies

both with and without inconclusive results. In practice, inconclusive outcomes are the default results when the criteria for deception or not-deception decisions are not satisfied and are a matter of the decision thresholds employed. Alternate term is *indefinite, or no opinion*.

indefinite
See: *inconclusive*.

independent variable
The variable manipulated by the experimenter to determine the effects on the dependent variable. As an example, if a researcher were interested in sex differences in PDD validity, the independent variable would be the gender of examinee and the dependent variable would be the accuracy of the PDD technique for each sex.

information gain
Statistical approach to determine the usefulness of a technique over the non-use of the technique. In forensic applications, the polygraph has been shown to provide a significant information gain over unassisted lay judgments across a wide range of base rates. In screening, only decisions of deception led to a significant improvement in information gain. See: Honts & Schweinle (2009).

innervation
Provide nerve supply, or to stimulate an organ through its nerves.

inside-issue comparison question
Test question used only the Quadri-Track Comparison Technique. Advocates of the technique assert it is designed to elicit a response from the truthful examinee concerned about a false positive error. Empirical support is mixed between advocate and independent research. See: Matte (1996); Mangan, Armitage & Adams (2008); Nelson & Cushman (2011); Shurany, Stein & Brand (2009).

inside-issue relevant question
Test question used only the Quadri-Track Comparison Technique. Advocates of the technique state that it is designed to elicit a response from the deceptive examinee hoping for a false negative error. Empirical support is mixed between advocate and independent research. See: Matte (1996); Mangan, Armitage & Adams (2008) Nelson & Cushman (2011); Shurany, Stein & Brand (2009).

inside track
One of three tracks in the Quadri-Track Zone Comparison Technique which include the primary, secondary, and outside tracks. The inside track employs two questions. One of these questions addresses an examinee's fear of a false positive error and is used as a comparison question. The other concerns the examinee's hope of a false negative error, which is treated and interpreted as a relevant question. Empirical support is mixed between advocate and independent research. See: Matte (1996); Mangan, Armitage & Adams (2008); Nelson & Cushman (2011): Shurany, Stein & Brand (2009).

inspiration (inhalation) / expiration (exhalation) ratio (I/E ratio)
The duration of inhalation compared with that of exhalation. Normally the ratio is about 1:2 in a resting human and changes during stress. It was first reported by Benussi in 1914. Changes in the I/E ratio are considered by some to be a diagnostic criterion in manual scoring. The I/E ratio as described by Benussi is quite different from that traditionally taught in polygraph schools. Also, more recent research has not found it to be of diagnostic value in manual scoring. See: Krapohl, (2020); Kircher, Kristjansson, Gardner, & Webb (2005).

instant offense examination

A form of *Post-Conviction Sex Offender Testing,* conducted when a subject is in denial of the offense or of some significant element of the offense for which he or she was convicted, and is often used to break down the denial barrier. This is also an examination that can be given when a new allegation has been made while the subject is on probation or parole. The polygraph is used to determine whether the allegations are true. Also called a specific issue examination. See: Cooley-Towel, Pasini-Hill, & Patrick (2000); Dutton, (2000); English, Pullen, & Jones (1996); Heil, Ahlmeyer, McCullar, & McKee (2000).

Integrated Zone Comparison Technique

The Integrated Zone Comparison Technique (IZCT) was developed in 1987 by Nathan J. Gordon, William Waid, and Philip Cochetti at the Academy for Scientific Investigative Training. Much of the design of the IZCT was based on formatting principles from the Backster Zone Comparison Technique although there are significant differences. Developers of the IZCT allow the examiner the flexibility to use the same test structure for both single-issue and multiple-issue cases. The IZCT is unique in that it uses a rank ordering system of analysis, called the *Horizontal Scoring System*. Also, unlike other techniques, the first chart is conducted as a silent answer test, and in the third chart there is a reversal of the positions of comparison-relevant questions to relevant-comparisons. See: Gordon, Fleisher, Morsie, Habib, & Salah (2000); Nelson & Handler (2011).

integument

Covering of the body (skin). Human skin consists of three primary layers: *epidermis, dermis,* and *subdermis*. It is comprised of a complex set of organs that provide protective and sense functions. Skin protects the body from environmental threats such as temperature, chemical, mechanical and infectious agents by acting as a selective barrier. Skin can aid in the removal of substances like water and solutes from the bloodstream through the sweat glands. From a sensory standpoint, skin houses various receptors to provide afferent information related to touch, pain and temperature See: Handler, Nelson, Krapohl & Honts (2010).

intent question

Question used in polygraph testing to determine whether the examinee had engaged in an act with criminal intent, rather than merely committed the act. It is considered the least reliable of all types of relevant questions in PDD testing and is avoided whenever possible. Some behavioral acts include intent by their definition, e.g. sexual contact.

interbeat interval (IBI)

Period between cardiac pulse waves, usually measured from systole to systole. The IBI has been shown to shorten just after the onset of stress in most people if a defense response has been found to occur. Contrarily, IBI has been found to increase initially during an orienting response. IBI and heart rate are reciprocals of one another.

inter-chart stimulation

Examiner-examinee interaction that takes place in the few minutes between individual tests. The interaction might include general reminders for the examinee to answer all questions truthfully (in the case of PLC techniques), or further emphasizing the comparison questions. Some research suggests that inter-chart stimulation may improve the validity of polygraphy, though it remains a controversial procedure. See: Abrams (1999); Honts (1999; 2000); Matte (2000).

introductory test

Alternate term for a stimulation test. See: *stimulation test*.

investigative examination

A polygraph examination which is intended to supplement and/or assist an investigation and for which the examiner has not been informed and does not reasonably believe that the results of the examination will be tendered for admission as evidence in a court proceeding. Types of investigative examinations can include applicant testing, counterintelligence screening, community safety examinations (e.g., post conviction sex offender testing, domestic violence testing, intoxicated drivers on probation, etc.), as well as routine specific issue and single issue or multiple-facet diagnostic testing.

irrelevant question

A question designed to be emotionally neutral to examinees. Irrelevant questions are most often placed in the first position of a question list because an orienting response usually follows the presentation of the first question and is of no diagnostic value. In CQT formats it is also used after a relevant or comparison question that has elicited a strong response to permit physiologic arousal levels to return to baseline before presenting another question. Irrelevant questions are used in nearly every type of PDD test. Also called *norms* or *neutrals*.

jackknife procedure

Statistical technique sometimes used to test a model. All data sets are used to develop a model except one, and the excluded set is tested against the model. This method is repeated until all sets have been excluded once and tested against the model built with the remaining data. This method produces a distribution of outcomes constructed from the individual outcomes of each data set, and is a method sometimes used to validate a model. Jackknife procedures have been used in PDD algorithm development.

Karpman's classification of lying

Classification of lies and their underlying motives. They are *benign lies* (for social conventions), *hysterical lies* (to attract attention), *defensive lies* (to avoid an adverse situation), *compensatory lies* (to impress another), *malicious lies* (for gain), *gossip* (exaggeration), *implied lies* (deceive with partial truths), "*love intoxication*" *lies* (idealistic exaggeration), and *pathological lies* (self-destructive or maladaptive). See: Karpman (1949).

Keeler, Leonarde

Student of John Larson and influential PDD pioneer. Among Keeler's (d. 1949) accomplishments were: the addition of the electrodermal channel to the polygraph, establishing the first PDD school, devising the Keeler Technique, and popularizing the polygraph field.

Keeler Polygraph

Originally manufactured by the Western Electro-Mechanical Company, this instrument was not produced after 1938. It had three tambours: one for the cardiosphygmograph, another for the pneumograph, and a third for either a second pneumograph or a muscular movement device. The kymograph could be geared to move the graph paper 3, 6, or 12 inches per minute. Associated Research, Inc. later produced the Keeler polygraph, similar to the original design except it permitted a galvanograph channel, and the chart speeds were 6 and 12 inches per minute. The Keeler polygraph is no longer in production.

Keeler Technique

A Relevant/Irrelevant testing method devised by Leonarde Keeler and used in single- and multiple-issue testing. Its popularity has declined since the introduction of the CQT formats and it is rarely used today.

key

The critical item in a series of similar but neutral items used in Peak of Tension (POT) tests. In a known solution POT, the key is the relevant question that contains the incriminating information that only a guilty person should know. A key in a searching POT is the test item that holds information that only the guilty person knows and the PDD examiner is trying to uncover. In stimulation tests, the key is the question to which the examiner directs the examinee to lie. Also See: *false key*.

key word method

Procedure employed during PDD testing in which the examinee is instructed to provide not only a yes or no reply but repeat an important word from the test question. Based on the stimulus–stimulus theory in which cognitive activity is involved as an intermediary step between a stimulus and a response. The key word in the test question is associated with the concept it is supposed to represent. The key word method is used to neutralize dissociation countermeasures.

Kircher features

Ensemble of measurable physiological features found in traditional polygraph recordings that correlate highest with deception. They are: respiration line length, electrodermal response amplitude, relative blood pressure amplitude, and finger pulse amplitude. See: Kircher & Raskin (1988).

known numbers acquaintance test

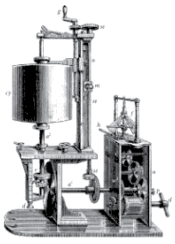
Stimulation test with several variants, it has as a central feature that the critical item is known to both the examinee and the polygraph examiner prior to the test. See: s*timulation test*.

Known Solution Peak of Tension test (KSPOT)

Peak of Tension test in which the critical item, or key, is known only to the investigator, polygraph examiner, a guilty person, or a person with incriminating knowledge. The key is placed in a question series among other items equally plausible to an innocent examinee and presented to the subject to determine if a consistent physiologic arousal occurs to the key. Like all Peak of Tension procedures, trend responses are used in addition to specific responses to interpret the recordings.

kymographion

Mechanical device that used a clockwork mechanism to rotate a drum on which many different types of phenomena could be recorded. The drum was covered with smoked paper. A stylus scratched the surface of the paper, creating a tracing that could be analyzed. The image below is a Ludwig kymographion from the mid-1800s. It was the forerunner of what later would be called a kymograph. .



labile

Unstable, inconsistent, or dynamic. PDD tracings that display a high degree of responsivity or broad amplitude changes are referred to as labile.

laboratory research

Scientific investigation in which experimental procedures are designed to mimic real-world circumstances, but in which there is direct control over the independent variables.

Lafayette Instrument Company

An American manufacturer of polygraphs, both analog and computerized, founded by Max Wastl. Headquarters is located in Lafayette, Indiana.

Larson, John

One of the first modern researchers in PDD, Dr. Larson (d. 1965) first used continuous recordings of respiratory and vasomotor activity with a test format using relevant and irrelevant questions. Dr. Larson's 1932 book, *Lying and Its Detection*, provided the best scientific evaluation of PDD up to that time.

latency

The delay between stimulus presentation and some aspect of the response. Onset latency relates to the delay between the stimulus presentation and the beginning of the response, while the peak latency uses the time of the maximum amplitude of the response as the second point. Latencies of specific physiologic responses vary. The latency of an electrodermal response, for example, from stimulus onset is about one to three seconds for the average person, while hormonal influences on blood pressure require several seconds more. A significant departure from typical latencies can indicate that a given response is unrelated to the stimulus, that there are problems in attention for the subject, or that countermeasures are being engaged. Because of individual differences, within-subject analyses are to be preferred.

law of initial values (LIV)

The magnitude of a given physiologic response will be constrained by the level of arousal present when the response begins. If a response occurs when arousal is already high, the amplitude of the response measured from onset to maximum expression will be less than if the same response occurred during a median level. While there are differing opinions regarding this psychophysiologic principle, it can certainly be said that all biological systems do have upper limits in their potential for response, and ceiling effects can come into play. This is because compensatory systems mediated via the sympathetic and parasympathetic nervous systems work to limit response intensities. Additionally, concrete limitations may exist, such as the finite number of sweat glands establishes the maximum electrodermal response.

Layered Voice Analysis (LVA)

A voice-based technology sold as a means of detecting emotions and deceit. LVA was developed in Israel by Amir Liberman, owner of Nemesysco, Ltd, and is sold in the US through Voice Analysis Technologies in Madison, Wisconsin. The LVA software operates on a laptop computer and applies numerous algorithms to the voice signal to assess a wide range of factors. The company has a very assertive promotional campaign. The company also attempts to distance this technology from the Computer Voice Stress Analyzer (CVSA) in part because of reports of poor validity for the CVSA, and the different approach to analysis of voice data. Research on the LVA has found its validity to be poor to none. See: Damphousse, Pointon, Upchurch & Moore (2007); Harnsberger, Hollien, Martin & Hollien (2009); Hollien & Harnsberger (2006).

law of intensity

Within limits, response magnitudes and stimulus intensities share a log-linear relationship; the stronger the stimulus, the greater the magnitude of the response. Response magnitudes are used in PDD to infer the type of question the examinee considers most salient or threatening.

lens model

Model for studying the decision rules used by human decision-makers, first proposed by E. Brunswik in the early 1950s. Conceptually, the model characterizes the decision process as the selection and evaluation of cues in the assessing of reality. Which cues are used and how they are weighted are central to this model. The term *lens model* springs from the sense that subjects view reality through the lens of these cues. This approach has been applied in the study of PDD decisions at the University of Utah. The lens model is useful to assess the diagnosticity of physiological responses, in identifying how examiners use the physiological information, and to determine the combination and weights of the cues that will maximize decision accuracy. See: Kircher, Kristjansson, Gardner & Webb (2005); Kircher & Raskin (1983); Kircher, J.C., Raskin, D.C., Honts, C.R., & Horowitz, S.W. (1995).

Law Enforcement Pre-Employment Test (LEPET)

A form of the Air Force Modified General Question Test (AFMGQT) which uses specific relevant questions and is used for police candidate screening.

lie detector

A common but inaccurate term for the polygraph.

likelihood ratio

The Likelihood Ratio (LR) provides an index of how much a test result will change the probability or odds of having a condition after a known or assumed prior incidence rate (base rate). In the case of polygraph testing, the condition of interest is involvement in the issue under investigation. The LR+ tells us how much more likely it is that a person is lying than not, after failing a polygraph test, compared with the likelihood before he or she sat in the chair and completed the test. If a person produces a truthful test result, the LR- tells us how much more likely the person is to be telling the truth than before the test. LRs may also be used to compare the efficacy of two or more scoring and decision models, for a given or assumed base rate. The advantage of the LR, compared with traditional Bayesian metrics such as positive predictive value (PPV) and negative predictive value (NPV) is that the LR is inclusive of inconclusive results, and will provide information that more accurately generalizes to field settings.

Limestone Technologies

A Canadian manufacturer of computerized polygraph instruments. Headquartered in Kingston, Ontario, Canada.

Lombroso, Cesare

Italian physician biologist who first employed instrumentation in an effort to detect deception in suspects in live criminal investigations. He reported in 1885 in the second edition of his book, *L'Homme Criminel* the use of the "hydrosphygmograph," a mechanical arrangement invented for medical purposes, to detect blood pressure changes during interrogation.

Luria, Aleksandr

Russian researcher and originator of the conflict theory, one of the theories proffered to explain the psychophysiological mechanisms underlying PDD. Luria did deception detection experiments with a tremograph. See: Luria (1930); Runkel (1936).

Lykken, David T.

Psychologist (d. 2006) and ardent critic of the CQT. Dr. Lykken produced numerous writings for the scientific and general press, including a book, *A Tremor in the Blood*, in which he argued strongly that the CQT is fatally flawed, that it resulted in wrongful criminal convictions, and it was vulnerable to

countermeasures by the guilty. Dr. Lykken did not publish any research of his own on the CQT but used anecdotal histories and interpretations of other research to form his arguments. Lykken endorsed the Guilty Knowledge Test (GKT, now known as the Concealed Information Test, or CIT), an alternate PDD testing format. The CIT has not been widely used outside of Japan. See: Lykken (1998).

Lykken scoring

System of scoring electrodermal responses in the Concealed Information Test (formerly the Guilty Knowledge Test) and establishing the threshold for decisions. The Lykken scoring system compares the responses of the critical test items in a rank order method against those of the neutral items. One variant uses averaged ranks. See: Lykken (1998).

maintenance examination

A form of *Post-Conviction Sex Offender Testing* (PCSOT) that is requested by a treatment provider, and looks at treatment-type issues; i.e., weekly report logs, masturbation habits, boredom tapes etc. See: Cooley-Towel, Pasini-Hill, & Patrick (2000); English, Pullen, & Jones (1996); Heil, Ahlmeyer, McCullar, & McKee (2000).

manual mode

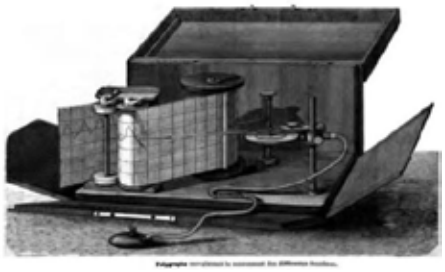
Setting for the electrodermal activity channel in computerized PDD instruments in which filtering of the tonic activity is minimal. All modern polygraphs manufactured in the U.S., have this feature. As opposed to *automatic mode*.

Marston, William

Psychologist, inventor of the discontinuous blood pressure method deception test, and author of the 1938 book *The Lie Detector Test*. Marston was the first to attempt to have instrumental deception test results entered into evidence in court, from which resulted the Frye decision of 1923. Marston's test entailed the use of a conventional blood pressure cuff and sphygmomanometer with which he manually plotted the examinees blood pressure during questioning at several points during the interview. Marston's work represented some of the early approaches to lie detection, though it could be considered polygraphy as it is now understood. Marston had several interests, and he was also the co-creator of the Wonder Woman comic book character. Both William Marston and his wife, Elizabeth, were lawyers and worked together to perform deception testing. See *discontinuous blood pressure method*.

Marey, Étienne-Jules

French researcher and early physiologist who invented the sphygmograph, a device for recording changes in blood pressure and pulse on a smoked drum in a manner not dissimilar from that later added to the polygraph. The diagram below shows a portable physiological recorder devised by Marey. See: Marey (1885).



mean blood pressure

The value of the pressure during the entire cardiac cycle. An approximation of the value of the mean blood pressure can be derived by averaging the systolic and diastolic values, or by summing of the diastolic pressure and one-third of the pulse pressure.

medulla oblongata

A part of the brain stem responsible for automatic control of respiratory and cardiovascular activity. The medulla oblongata is closely associated with physiological events relating to polygraph test data analysis.

mental countermeasure

A class of countermeasures in which the examinee attempts to affect the polygraph recordings through self-manipulation of attention, memory, emotion, cognition, semantics, or arousal. See: Krapohl (1996).

microtremor

Low-frequency oscillation of the human voice in the range of 8 to 12 hertz, and the component used to infer deception with the Psychological Stress Evaluator and the Computer Voice Stress Analyzer. It is claimed by the manufacturers of these devices that there is an inverse relationship between stress and the microtremor. Independent research has not yet found any spectral component of human voice a reliable predictor of deception.

Minnesota Multiphasic Personality Inventory (MMPI)

One of the most common personality inventories employed in psychology and psychiatry. It uses a series of questions to make diagnoses according to standard psychiatric clinical criteria. Among the most cited scales reported in the PDD literature are *psychopathy* and *social introversion-extroversion scales*, both of which have been shown to influence physiologic arousal levels, particularly in the electrodermal response. Evaluation and interpretation of the MMPI is generally based on profiles resulting from individual scales. Test-retest validity for the MMPI averages above 70%, though the validity of the individual personality scales has not been conclusively demonstrated.

Mixed Question Test

A Reid polygraph test in which the straight-through (ST) test questions are changed in order, and some of them are repeated. When the Mixed Question Test is used, it is always conducted after the ST.

mixed issue test

Any polygraph test technique in which each there is little or no overlap in the coverage among the relevant questions. Sometimes called a *multiple-issue test*.

model policy

Non-binding standard that outlines best practices in a given area. The American Polygraph Association uses model policies to help agencies and clients know what good polygraph practices are, and thereby provide a competitive advantage for examiners who adopt these best practices. Among the current APA Model Policies are those for police applicant screening, Post Conviction Sex Offender Testing, and Examinee Suitability.

Modified Relevant Irrelevant (MRI) technique

Specific-issue PDD format based on the Keeler RI format, but it uses situational comparison questions. Users of the MRI discuss all tested issues with the examinee during the pretest interview but

prefer not to review the relevant questions word for word. The prohibition against relevant question review is not an absolute, however. Results from the MRI are based on extrapolygraphic information, and 3- and 7-position scoring. Reported by Paul Minor (1985) but seldom used in the field.

monitoring examination

A form of *Post-Conviction Sex Offender Testing* (PCSOT) that is requested by a probation or parole officer to ensure compliance with the conditions of the offender's release from prison; i.e., alcohol or drug issues, computer violations, contact with children etc. See: Cooley-Towel, Pasini-Hill, & Patrick (2000); Dutton, (2000); English, Pullen, & Jones (1996); Heil, Ahlmeyer, McCullar, & McKee (2000).

Monte Carlo method

A statistical tool which is based on repeated random sampling of data and has been applied to the problem of estimating polygraph decision accuracy. The expression was coined by scientists at the Los Alamos National Laboratory in the 1940s for their approach to estimating radiation shielding.

Mosso, Angelo

Student of Cesare Lombroso, who in 1896 developed the scientific cradle, a device for recording bodily responses to fear.

motor nerves

Neurons that carry nerve impulses from the central nervous system to the effector organ or muscles. Also called *efferent nerves*.

movement sensor

Mechanical sensor that detects covert movements. The movement sensor is used to detect certain types of physical countermeasures.

multiple-facet test

Test format in which the relevant questions are targeted toward different elements of the same crime. For example, in a counterfeiting case, the PDD examiner might use three relevant test questions with a suspect. One could cover printing the bills, the second passing the bills, and the third knowing where the printing equipment is. In such a test the spot scores would determine whether a diagnosis of deception is made, as opposed to the overall score. The Zone "exploratory," the Air Force Modified General Question Test, and criminal RI are three possible formats for this approach.

multiple-issue examination

Typically used in screening, it allows the PDD examiner to determine which of several areas should be followed up with further questioning. It is somewhat uncommon to make decisions of truthfulness or deception in these types of tests because of the known adverse effects on decision accuracy. Such decisions are generally made after subsequent testing on the isolated issue in a single-issue test format. Among the more common multiple-issue test formats are the Law Enforcement Pre-employment Test and the Test for Espionage and Sabotage. See: Barland, Honts, & Barger (1989).

Münsterberg , Hugo

Chairman of the Psychology Department at Harvard who, in his 1908 book *On the Witness Stand*, suggested the possibility of devising deception tests using cardiovascular, breathing, and electrodermal measures. Münsterberg also described the Concealed Information Test. He had as a student William Marston, who later went on to develop the discontinuous blood pressure method deception test.

National Center for Credibility Assessment

The NCCA is the US government polygraph education, oversight and research center for credibility assessment, including the polygraph. Other historical names include: the U.S. Army Polygraph School (1951-1962); the US Army Military Police School (USAMP, 1975-1986); the DoD Polygraph Institute (DoDPI, 1986-2007), and; Defense Center for Credibility Assessment (2007-2009). With its campus located at Ft. Jackson, SC, the NCCA falls under the Defense Counterintelligence and Security Agency. The "Center", as it is often called, provides all polygraph instruction for the US federal government

nervous system

Consists of the brain, spinal cord, and peripheral nerves, each performing specific functions. Processing of nerve impulses in the brain is somewhat localized. Basic functions are mediated in the lower parts of the brain, activities such as hunger, thirst, and thermoregulation. Sensory regions of the brain are located above, along with most voluntary control of muscles. The highest regions of the brain are dedicated to processing and integrating information, and the production of thought. The spinal cord is the primary pathway by which most of the nerve impulses are carried to the brain. Nerves throughout the body send pulses through the spinal cord to the brain where they are processed, and the brain sends back impulses to regulate and control organs and muscles. There are two main divisions to the nervous system: the *central* (brain and spinal cord) and the *peripheral* (nerves and ganglia located outside of the central nervous system). The peripheral nervous system is further divided into the somatic (voluntary muscular movements) and autonomic branches (various unconscious functions such as digestion, sweating, heart rate, pupillary response, vasomotor activity, etc.) Some taxonomies also add a third branch, the sensory nervous system. In polygraphy, the autonomic branch receives special attention due to its association with the physiological data recorded and analyzed with the polygraph.

neuron

Structural unit of the nervous system and is the conducting cell. The typical neuron consists of a soma body, dendrites and axon.

neurotransmitter

Chemical involved in the transport of the neural signal to another neuron or effector organ. Neurotransmission has six stages: synthesis of the neurotransmitter, storage, release, receptor interaction, re-uptake, and inactivation. There are many pharmacological agents that influence neurotransmission, and they are of interest in PDD research due to their effects on tonic and phasic arousal levels.

neutral question

Another term for the irrelevant question in a CQT. Also called a *norm*.

No Deception Indicated (NDI)

In conventional PDD, NDI signifies that (1) the polygraph test recordings are stable and interpretable and (2) the evaluation criteria used by the examiner led him to conclude that the examinee was truthful to the relevant issue. The NDI and DI (Deception Indicated) decision options are used in specific- issue testing and correspond to NSR (No Significant Responses) and SR (Significant Physiological Responses) in multiple-issue, or screening, examinations.

non-current exclusive comparison question

A probable-lie comparison question that is of the same type or category as the relevant issue but excludes the relevant issue by use of a time-bar. It is the type of comparison question developed and advocated by Cleve Backster. See: Matte (1996).

non-exclusive (inclusive or inclusionary) comparison question

Comparison question that overlaps the relevant issue by time, location, or issue. Also called *Reid*, *inclusionary*, or *inclusive comparison question*. As an example, if the relevant issue were the robbery of a particular bank on a specific date, the comparison question might be, "Have you ever stolen anything in your life?" There is a long-running debate in the PDD community regarding the supremacy of the exclusive over the non-exclusive comparison questions. The current body of evidence supports the non- exclusive comparison question. See: Amsel (1999); Horvath (1988); Horvath & Palmatier (2008); Podlesny & Raskin (1978).

non-specific responses

Physiological responses that do not appear to have any relationship to the presentation of an external stimulus.

noradrenaline

British term for *norepinephrine*. See: *norepinephrine*.

norm

Verbal shorthand currently used by some PDD examiners to signify an irrelevant question in Relevant/Irrelevant and comparison question test formats. Much earlier (1922) John Larson referred to "norms" as individuals who were possible-but-unlikely suspects to a crime whom he added to his list of persons to be tested so he could account for variables such as anger, indignation, and fright that he could expect from innocent-but-likely suspects he would be testing.

No Opinion

Alternate term for an Inconclusive. Sometimes used to denote an Incomplete in some sectors.

No Significant Physiological Responses (NSPR or NSR)

Common verbiage for polygraph screening examination outcomes equivalent to No Deception Indicated in single-issue tests. The alternate language comes from an acceptance that screening examinations do not produce the high validity of single-issue tests, and therefore, the results are better reported as the absence of physiologic arousals rather than inferring truthful intent on the part of the examinee.

numerical analysis

Systematic assignment of numbers to physiologic responses, along with decision rules, so that PDD data analysis is more objective and standardized. The first such system was published by Dr. John Winter in 1936. Contemporary numerical analytic methods include the *Rank Order Scoring System, Horizontal Scoring System, 3-position scoring system, 7-position scoring system, Empirical Scoring System, and the Lykken Scoring*. Sometimes referred to as *semi-objective analysis*.

numerical chart analysis

Method of rendering polygraph decisions that are based exclusively on numeric values that have been assigned to physiological responses recorded during a structured polygraph examination. The numerical approach does not consider extra-polygraphic information such as case facts or examinee behaviors. The numerical approach has four primary components. They are: feature identification, numerical value assignment, computation of the numerical values, and decision rules.

Objective Scoring System (OSS)

A form of 7-position scoring where the individually assigned values are derived from ratios that come from measurements of the "Kircher features." Because the scores come from measurements, the OSS eliminates subjectivity in chart interpretation. However, it is very time-intensive when performed manually, and impractical for routine use. The OSS has been automated by some computer polygraph manufacturers. The OSS version 3 (Nelson, Krapohl & Handler, 2008) can accommodate almost all probable-lie CQTs. See: Dutton (2000); Krapohl & McManus (1999).

Oculomotor Deception Test

A credibility assessment technology that uses eye movements, pupillary responses and reaction time of examinees as they read statements and press a key for true or false. See: Webb, Hacker, Osher, Cook, Woltz, Kristjansson & Kircher (2009).

operant conditioning

Type of conditioning in which reward or punishment is given to the subject, depending on the preceding behavior. Rewards increase the likelihood of the recurrence of the behavior and punishment discourages the behavior. Biofeedback uses operant conditioning to help patients reduce blood pressure, electrodermal activity, and other physiological processes. Since operant conditioning can be used to teach individuals to regulate their own autonomic responses, it is a method for teaching PDD countermeasures. The little research that has evaluated the influence of biofeedback training on PDD efficacy has not found an effect. See: Honts (1987).

orienting response (OR)

Heightened sensitivity to specific sensory input that is characterized by increased information processing, narrowed attentional concentration, and physiologic excitation.

Othello error

Expression coined by Paul Ekman to denote the misattribution of the fear or emotional distress of an innocent person as an indication of guilt. See: Ekman (1985).

outlier

A value beyond the normal range of values. For example, the last value in the following series could be considered an outlier: 5,9,2,6,6,8,3,1,6,9,5,32. Outliers may be excluded from data collection because they have inordinate influence on central tendency. What constitutes an outlier, or extreme score, is established a priori.

outside issue question

See: *symptomatic question*.

outside track

One of four tracks in the Quadri-Track Zone Comparison Technique, which include the primary, secondary, and inside tracks. The outside track consists of two symptomatic questions. See: Matte (1996).

overall truth question

PDD test question that addresses the examinee's overall truthfulness or intention to be truthful during testing. Used in some multiple-issue screening tests.

P300

Event-related potential (ERP) of the brain measured at the scalp that has an average latency of 300 milliseconds from stimulus onset and is recorded maximally at site PZ in the International 10-20 System. The P300 is related to unique characteristics of the stimulus and is most often associated with the "oddball" or rarely occurring stimulus. For example, the P300 is known to be evoked by a low incidence auditory tone that is of a different pitch than another tone that is occurring much more frequently. A P300-based Concealed Information Test has been developed for criminal testing, though field testing is incomplete. See: Farwell & Donchin (1991); MacLaren & Taukulis (2000); Rosenfeld (1998).

padding

An alternate expression for irrelevant items in a Known Solution Peak of Tension test. In some reports, padding relates to only the first or last one or two items in these tests.

pain countermeasure

Special type of physical countermeasure in which an examinee will attempt to evoke physiologic responses by covertly self-inducing discomfort. Included in this group are strategies such as biting the tongue, pressing against a sharp object in the shoe, forcing a fingernail into the thumb cuticle, and irritating a wound. While spontaneous use of these tactics has not been found effective, they can be more powerful when the examinee receives training and feedback. See: Honts, Raskin & Kircher (1994); Krapohl (1996).

parameter

Term used in PDD to denote a single physiological data channel, such as the pneumograph, cardiograph, etc.

paradigm

Example or model. Experimental paradigms attempt to explain real world phenomena by assessing the critical elements and their relationships with one another.

parasympathetic nervous system

One of the three divisions of the autonomic nervous system also referred to as the craniosacral system because the preganglionic neurons lie in those areas. Parasympathetic ganglia anatomically lie in or near the organs they innervate thus allowing for more localized control. Functionally, it is involved in conservation and restoration of energy. The parasympathetic and sympathetic divisions of the autonomic nervous system function to maintain homeostasis.

parasympathomimetic

An agent whose effects mimic those resulting from stimulation of parasympathetic nerves, especially those produced by acetylcholine.

Pathometer

Device used by Rev. Walter Summers to perform deception tests in the 1930s. Researchers had to order Pathometers from Fordham University, and they were only assembled upon order. Summers conducted testing on hundreds of subjects using this recording galvanometer and a testing procedure he devised that included what would be later known as comparison question See: Summers (1939).

Peak of Tension (POT)

Family of recognition testing procedures, including *known solution*, *searching* (*probing*), and *stimulation tests*. A Known Solution POT (KSPOT) is used to determine whether the examinee is aware of details of a crime that have been kept from the general public and would presumably only be known to the

perpetrator of the crime or those with incriminating knowledge. A Searching POT (SPOT) is used to determine details of a crime that are not known to officials, such as the location of an unrecovered body, but would be known to a participant in the crime. The evaluation criterion of Peak of Tension strip charts is simply identifying the point in the tracings where physiologic arousal has peaked, hence the name. Peak of Tension tests are not generally used to determine truth or deception, but rather to assist in the investigation or interrogation.

Pearson product-moment correlation

A test of correlation between two sets of interval level data (See: *scale of data*). The coefficients will lie between -1 and +1. A value of 0 would indicate no correlation, while -1 and +1 would mean perfect negative and positive correlations, respectively.

peripheral nervous system

Portion of the nervous system resident primarily outside of the brain and spinal cord. The cell bodies of the preganglionic sympathetic neurons lie in the spinal cord and those of the parasympathetic branch are situated in the brain stem.

pertinent question

Rarely used term for *relevant question*.

phalanx

Any bone of the fingers and toes. In research and as commonly practiced in PDD, electrodermal sensors are attached to the distal phalanx, or end of the fingertips. Plural is phalanges.

pharmacological countermeasures

Class of countermeasures in which the examinee attempts to affect the polygraph recordings through the use of ingested drugs or application of topical preparations. See: Krapohl (1996).

phasic response

A physiological response characterized by a relatively rapid change from and subsequent return toward baseline.

photoplethysmograph (PPG)

The PPG uses the reflection of a red light emitted into the skin to detect changes in the volume of blood in the upper layers of skin, typically recorded at the finger when using a polygraph. Physiological arousal is marked by a constriction in the pulse amplitude as blood is shunted from the extremity during activation of the sympathetic nervous system. See: Geddes (1974); Handler & Krapohl (2007); Kircher & Raskin (1988).

photopolygraph

A polygraph created by C.W. Darrow in the 1930s. It was one of the most elaborate polygraphs of that era, recording relative blood pressure, skin resistance, breathing, reaction time, and bilateral hand tremors. It had two stimulus markers, one activated by hand and the other was a voice key. Costing upwards of $2,000 and requiring a separate technician to operate, Darrow's photopolygraph was primarily a laboratory instrument and was not used extensively by the PDD community. Also called the *Darrow Photopolygraph*. See: Darrow (1932).

physical countermeasures

Class of countermeasures in which the examinee attempts to manipulate the polygraph recordings through the discreet use of movements. Some of these movements are also used to self-induce pain. See: Honts (1987); Krapohl (1996).

Pinocchio response

Nonexistent lie-specific physiological response. The expression sometimes used to deride the notion that the act of deception produces stereotypical response patterns.

place bar

One method of restricting the coverage of the comparison question so that it will not overlap the relevant question. For example, if a relevant question concerned whether the examinee had physically assaulted a person in the city in which he now resides, a comparison question with a place bar may inquire whether the examinee had ever deliberately hurt another person while living in another city. There is a school of thought that examinees may confuse the relevant questions with the comparison questions unless these two types of questions are designed to avoid any degree of overlap. Research has not supported this hypothesis, however. See: Amsel (1999); Podlesny & Raskin (1978); Horvath (1988); Horvath & Palmatier (2008). Also See: *exclusive (exclusionary) comparison question*

placebo

Procedure or substance with no intrinsic effect but is useful to convince the patient or subject that an effective treatment has been applied. Placebos often have effects that are attributable to suggestion. They are used extensively in medical research for control purposes during drug testing and for certain psychosomatic illnesses. In PDD it addresses one type of mental countermeasure whereby examinees use ritual objects, incantations, or other ineffectual actions with the expectation that the power of the polygraph to uncover deception will be impeded.

plethysmograph

A device used to measure relative changes in blood volume and pulse volume. The three most commonly used are (a) changes recorded using a strain gauge, (b) impedance changes and (c) photoelectric changes. It is the third technique that is used in modern polygraphy to detect relative changes in pulse volume associated with the vasomotor response, usually at the distal phalanx of one of the examinee's fingers.

pneumograph

A device that records breathing, and one of the three traditional channels of the modern polygraph used in PDD. Most contemporary polygraphs use two pneumograph recordings: abdominal and thoracic. The types of sensors include the traditional corrugated rubber tube, the mercury strain gauge, or the newer piezoelectric.

polygraph

By definition, an instrument that simultaneously records two or more channels of data. The term now most commonly signifies the instrument and techniques used in the psychophysiological detection of deception, though polygraphs are also used in research in other sciences. In PDD the polygraph traditionally records physiologic activity with five sensors: blood pressure cuff, electrodermal sensors, two breathing sensors and a sensor designed to detect covert muscular movements. Some instruments also record *finger pulse amplitude* using a photoplethysmograph.

polygraph surveillance
See: *maintenance polygraph examination.*

polygram
Complete graphical recording of physiological data from a polygraph test with the required annotations. Often called a *polygraph chart*.

positive control pair
The combination of the subjective truth question and the subjective lie question, to form a set in the Positive Control Technique. See: Driscoll, Honts & Jones (1987); Gordon & Cochetti (1982); Howland (1981); Reali (1978).

positive comparison question
In the Positive Control Technique, each question is presented to the examinee twice in a row, and the examinee is instructed to answer differently the first time from the second time. Therefore, each question serves as its own comparison question. See: Driscoll, Honts & Jones (1987); Gordon & Cochetti (1982); Howland (1981); Reali (1978).

Positive Control Technique
Technique that employs most of the standard test questions except a probable-lie comparison questions, and each question is presented twice in succession during the testing. The examinee is instructed to answer truthfully to the first presentation, and untruthfully the second time, or vice versa. While the technique is amenable to the 7-position scoring, it has its own unique set of decision rules that are different from the more familiar comparison question formats. The Positive Control Technique is one form of the Yes-No Technique. See: Driscoll, Honts & Jones (1987); Gordon & Cochetti (1982); Howland (1981); Reali (1978).

Post-Conviction Sex Offender Testing (PCSOT)
Specialized application of polygraphy which aids in the management of the convicted sex offender who has been released into the community, though sometimes is employed as part of treatment of offenders who are incarcerated. See: Dutton (2000).

post hoc
(L: after this) Establishment of criteria or analyses after the conduct of the experiment is completed.

posttest
Final portion of a polygraph examination. The posttest could include a debriefing of an examinee who passed the examination, or an interview or interrogation of an examinee who failed the examination. The posttest may or may not be a part of any given polygraph technique and plays no part in the formulation of the results in any polygraph technique.

pre-ejection period (PEP)
Time between the Q wave of the electrocardiogram and the B wave of the impedance cardiogram for the same pulse. It is the period between when the ventricular contraction occurs and the semilunar valves open ejecting blood into the aorta. Shorter periods are thought to correlate with sympathetic nervous system arousal. The sensors for the production of the PEP phenomenon are relatively noninvasive, and if future research validates it as a diagnostic measure, the PEP could be added as an alternate PDD channel.

Preliminary Credibility Assessment Screening System (PCASS)

Device developed by the Johns Hopkins Applied Physics Laboratory, in conjunction with the Lafayette Instrument company, at the request of the US Government in 2005. Its concept of operation is to be used by minimally trained US troops as an initial screening tool in war zones to pare down the number of individuals who would undergo subsequent vetting by the polygraph and other tools. It has two sensors: electrodermal and photoplethysmograph. Test questions are typed into the template of the PDA platform, and the user taps the screen to indicate the place of question onset. The PCASS is a one-chart test that runs about 12 minutes. At the completion of testing an algorithm analyzes the data to produce the screening decision. Five laboratory studies have been conducted with the PCASS using realistic wartime scenarios or mock theft scenarios, with a combined accuracy of 80% when inconclusives were excluded, and about 23% inconclusives. The algorithm was devised to minimize false negatives. The PCASS was approved for use in the Department of Defense in 2007. See: Battelle Memorial Institute (2007); Senter, Waller & Krapohl (2009).

premature ventricle contraction (PVC)

Term loosely applied to distortion in the cardiograph waveform resulting from an ectopic heartbeat. More precisely it is a ventricular contraction between two sinus cycles without a compensatory pause. Sometimes referred to as *extrasystolic beat* (*esb*) in the older literature. See *arrythmia*.

pretest interview

The earliest portion of the PDD examination process during which the examinee and examiner discuss the test, test procedure, examinee's medical history, and the details of the test issues. The pretest interview also serves to prepare the examinee for the testing. The length of the pretest interview ranges from 30 minutes to 2 hours or longer, depending on the complexity of the case, examiner-examinee interactions, and testing technique. All PDD techniques use pretest interviews.

primary track

One of four tracks in the Quadri-Track Zone Comparison Technique, which include the secondary, outside, and inside tracks. The primary track contains a relevant question, and a non-current exclusive probable-lie comparison question against which the relevant question is compared. See: Matte (1996).

probability

Likelihood of an occurrence, expressed as a number. By convention, probabilities are reported in scientific reports as numbers between 0.00 and 1.00. Probabilities are most often reported in PDD studies to characterize the likelihood of the experimental results occurring by chance.

probable-lie comparison question (PLC)

One of two major types of comparison questions. PLCs are questions to which it is likely that the examinee is untruthful or unsure of his or her answer. Their intended purpose is to create a competition of salience such that the anxious innocent examinees will expend more of their physiologic responses on them than the relevant questions, but the guilty examinee will still find the relevant questions more arousing than the PLCs. Most systems of analysis compare the physiological responses elicited by the PLC with those from the relevant questions. A PLC is fundamentally different from a DLC (directed lie) in that the examinee believes he must pass the PLC question to pass the examination, whereas the true purpose of the DLC is more apparent to the examinee. Two main types of PLCs are the *exclusionary* (Backster type) and the *non-*ex*clusionary* (Reid type).

Probing Peak of Tension
See: *Searching Peak of Tension.*

pseudorelevant question
A test question so worded as to appear to be relevant to the examinee. Example: "Did you lie to any question on this test?" or "Do you intend to answer truthfully each question on this test?".

Psychogalvanic Reflex (PGR)
Expression coined by Veraguth for what is now called the *electrodermal response*. See: Veraguth (1906).

psychograph
A term from the 1930s for the polygraph that consisted of the pneumograph, sphygmograph, and a stimulus marker. Sometimes referred to as the *Berkeley Psychograph,* the *Lee Polygraph,* and the *cardio-pneumo-psychograph.*

psychological set
The expression *psychological set* was introduced in PDD by Cleve Backster who initially attributed it to a psychological writer Floyd L. Ruch (Matte & Grove, 2001) but later claimed to have made up the expression himself (Senter, Weatherman, Krapohl, & Horvath (2010). Backster has made the concept central to his Zone Comparison Technique and has tethered the concept to the emotion of fear. According to Backster's PDD hypothesis, examinees are expected to attend more to the category of question that presents the greater threat to their interests, either the relevant or comparison questions. Subjects who are lying to the relevant issues consider these questions more threatening than the others, which, in turn, draw more attention to the relevant questions, and more physiological arousal. Similarly, innocent subjects find the probable-lie comparison questions more disconcerting, and the greater attention paid to them generates the larger arousals. The expression *psychological set,* together with its underlying assumptions, have long been questioned by scientists on both sides of the polygraph debate. Competing concepts include "Differential Salience*"* (Senter, Weatherman, Krapohl & Horvath, 2010) and "Relevant Issue Gravity" (Ginton, 2009.) See: Krapohl (2001); Matte & Grove (2001).

Psychological Stress Evaluator (PSE)
A voice stress device. Dektor Counterintelligence and Security and Allan Bell Enterprises produced the PSE, first introduced in 1971. This device, which is no longer manufactured, is the original voice stress analyzer. See: Horvath (1978; 1979); Lynch & Henry (1979).

psychopath
An individual with a personality marked with superficial charm, habitual lying, no regard for others, showing no remorse after hurting others, having no shame for outrageous and objectionable behavior, impulsivity, inability to form relationships and take responsibility, failure to learn from punishment, lack of empathy and conscience, and need for excitement. Also referred to as antisocial personality. While popular lore holds that the psychopath, with his diminished conscience, is able to defeat PDD testing, all research has found that the guilty psychopath is no different from guilty non-psychopaths in being detected by the polygraph. See: Barland & Raskin (1975); Raskin & Hare (1978); Patrick & Iaconno (1989).

psychophysiological detection of deception (PDD)
Common scientific term to denote the use of the polygraph to diagnose deception.

pulse pressure
The arithmetic difference between the systolic blood pressure and the diastolic blood pressure.

pulse transit time (PTT)
Period of time for the passage of a mechanical pulse wave between two points on the body. Typically, it is derived by an EKG recording of an electrical discharge from the heart signalling the beginning of a ventricle contraction, and then timing when the blood pressure wave arrives at a distant point in the body, often the finger. Faster pulse transit time corresponds with greater sympathetic nervous system activation. PTT may have some usefulness as a PDD parameter.

punishment theory
One of several theories that attempt to explain PDD. It holds that physiologic arousal during deception is activated by the fear of the consequences if detected. This theory fails to explain why polygraph testing still functions well in the absence of fear.

pupillary response
Change in the diameter of the pupil of the eye in response to stimuli. Pupil size is regulated by the sphincter pupillae muscles in the iris, which respond to parasympathetic stimulation, and the dilator pupillae muscles, innervated by the sympathetic nervous system. Dilation can result from sympathetic nervous system stimulation or the suppression of the parasympathetic nervous system. Pupil dilation has been investigated by several researchers as an index of stress and continues to be a phenomenon of interest in PDD. See: Bradley & Janisse (1981); Webb, Honts, Kircher, Bernhardt & Cook (2009).

Purposeful Non-Cooperation (PNC)
An expression first reported by John Reid to denote a PDD outcome in which examinees had used physical countermeasures in an attempt to defeat the polygraph examination. Reid did not consider PNC to be synonymous with the practicing of deception, though he wrote that it was a fairly reliable indicator of the examinee's motives to deceive.

Quadri-Track Zone Comparison Technique
A single-issue polygraph technique developed and advocated by James Matte which extends the method devised by Cleve Backster. This technique has four "tracks:" *primary track, secondary track, inside track,* and *outside track*. The primary and secondary tracks consist of pairings of two relevant questions with two non-current exclusive probable-lie comparison questions. The inside track contains two questions, one relating to the examinee's concern about a false positive error, and the other regarding the examinee's hope for a false negative error. Independent research has to date failed to support the construct of the inside track (See: Nelson & Cushman, 2011). The questions used in the outside track share some similarities with questions historically referred to as *symptomatic questions*. A 7-position scoring system is used for analysis. For a full explanation, see: Matte (1996). Formerly known as the *Quadri-Zone Technique.*

R-wave peak to carotid incisura (RWPCI)
A cardiac response measured by the interval between the peak of the R-wave on an electrocardiograph (indicating ventricular contracture) and the arrival of the pulse waveform at the carotid incisura in the neck. It has been investigated as a possible additional measure of sympathetic nervous system arousal in the PDD setting.

radial artery

Major artery in the forearm and wrist along with the ulna artery, and one of the alternate PDD recording sites for cardiovascular activity using the blood pressure cuff.

random assignment

Research strategy whereby each selected subject is placed in a given group by chance. This can be accomplished through random numbers tables, coin flips, or other chance methods. Random assignment is one way of assuring that experimental effects are not the result of a systematic error in groupings of subjects. Random assignment is not the same as random selection, which addresses which subjects will participate in the study. Most laboratory studies of PDD randomly assign subjects into programmed guilt or programmed innocence.

random selection

Method in research for extracting samples from a population where each individual has an equal opportunity for selection, and the selection of a subject has no influence on the selection of other subjects. Random selection is used to avoid the systematic error that can occur from other strategies. Random selection is not the same as *random assignment*, which addresses to which group the selected subjects will be placed. In practice, true random selection from the larger population is difficult to accomplish in PDD research. Either subjects are drawn from subgroups, such as military recruits or college students, or they self-select in proportions that may not well represent the larger population, such as through newspaper recruitments.

rank order analysis

Any of the methods of PDD scoring that entail the assignment of ranks to response intensities within channel. Rank order analysis methods for the CQT include the Horizontal Scoring System and the Rank Order Scoring System. First published report for ranking of responses was for the Guilty Knowledge Test in the 1950s. See: Gordon & Cochetti (1987); Honts & Driscoll; (1987); Krapohl, Dutton & Ryan (2001); Lykken (1959); Ohnishi, Matsuno, Arasuna & Suzuki (1976).

rationalization

Self deception. In the psychoanalytic model, rationalization is a psychic defense mechanism in which one's true motives or behaviors that are threatening to the psyche are reinterpreted to be other motives that are more acceptable. PDD examiners routinely watch for indications of rationalization during the pretest interview to prevent a false negative through faulty test questions that empower the self-deceit. No research has evaluated the effects of rationalization on PDD efficiency.

reaction tracing

Section of a physiologic tracing in which an arousal is apparent.

Receiver Operating Characteristics (ROC)

Also known as the Relative Operating Characteristics, it is a graphical plot of the sensitivity, or true positive rate (true deceivers correctly identified), vs. false positive rate (percent of non-deceivers falsely implicated), for a given separation of the distribution of all possible scores for all possible choices of threshold. It is a psychophysical conceptual model for detection efficiency based on *signal detection theory* (SDT). The ROC characterizes the sensitivity of the decision criteria versus the specificity, and is useful to predict false positive and false negative rates across all levels of a criterion (cutting score, in PDD). It is an extension of work from the 1940s regarding the ability of radar operators to discriminate radar signals of friendly aircraft from those of enemy aircraft or noise. See: Swets (1995); Swets, Dawes & Monahan (2000).

recognition test

Polygraph techniques can be divided into two major categories, knowledge-based tests, also called recognition tests, and deception-based tests. The recognition test family of PDD techniques includes the Peak of Tension test, acquaintance test and concealed information test. Recognition tests attempt to determine if the examinee has knowledge only available to persons directly involved in an incident of concern. Because this approach depends upon the existence of a known crime or incident facts that remain unknown to the innocent suspect, the recognition test paradigm is not suited for use as a screening test. See: Krapohl, McCloughan & Senter (2006); Lykken (1959); Osugi (2011).

recovery half-time

Interval between response onset and the return of the response to one-half of the maximum amplitude of the phasic response. Recovery half-time has been investigated as a diagnostic feature with electrodermal data using automated analysis. See: Kircher & Raskin (1988).

recovery time

Period between maximum amplitude of a phasic response and the return to a predetermined level.

red zone

In the Backster framework, it is a 20- to 35-second block of polygraph chart time initiated by a relevant question having a unique psychological focusing appeal to the guilty (deceptive) examinees. One of the three primary zones in the Zone Comparison Technique (red, green, black).

regression analysis

Method for mathematically modeling a relationship and is used in prediction and description. Unknown values for the dependent values can be estimated by what is known about the corresponding independent variables. Independent and dependent variables must be at least interval scale. Regression analysis comes in various forms and is very useful to parse out the contributions individual variables make to an outcome. Multivariate and univariate regression are distinguished by the number of independent variables. At least one PDD algorithm uses a form of regression analysis.

Reid, John

One of the first modern PDD examiners, Reid (d. 1982) developed many techniques still in use today. Reid is credited with bringing the probable-lie comparison question into common practice in the field. He also developed the Reid Technique, which includes the Yes Test and the Guilt Complex Test. Reid helped bring about the first state licensure for PDD practitioners in Illinois in 1963. Reid instructed hundreds of students at his school and offered the first accredited Masters program in PDD. Though the Reid School closed in the 1980s, the Reid Technique is used by some PDD examiners today.

Relevant/Irrelevant (RI) Technique

Family of polygraph test formats in which traditional lie comparison questions are not employed. While originally used in criminal testing, RI tests currently are more often found in multiple-issue screening applications. The RI test can trace its roots to word association tests employed in the early 1900s, and these word tests were later used occasionally during the monitoring of electrodermal activity. The RI was used extensively by pioneers John Larson and Leonarde Keeler in the 1920s through the 1940s, and it is still in use today. Among the screening techniques that have undergone validity research, the RI has been shown to have poorer validity and reliability than any of the more modern approaches.

Relevant Issue Gravity (RIG)

A unifying theory proposed by Avital Ginton (2009) to explain the underlying mechanisms in PDD. According to the RIG theory, the force induced by aggregation of qualities possessed by the relevant issue attracts and binds the guilty examinee's attention to the relevant questions in a manner not experienced by the innocent examinee. It is manifested in the preoccupation in the guilty examinee's mind with the relevant issue and in difficulties to divert attention to other topics or issues. It is postulated that the RIG strength for the guilty examinees on average is stronger than the RIG strength for the truth-tellers. Therefore, it is harder to divert the attention of the guilty examinee to the comparison question and relatively easier to do that with the truth-tellers. The theory is compatible with existing experimental evidence, and does not rely on older hypotheses regarding the central role of fear to explain polygraph reactions. See: Ginton (2009).

relevant question

A question that deals with the target issue of concern to the investigation. In addition to "did you do it" types of questions, relevant questions also include evidence-connecting and "do you know who" questions. Strong relevant questions address the "did you do it" type of questions, while moderate-strength relevant questions address evidence connecting and prior knowledge, such as participation in planning, providing help the perpetrators, or knowing the identity of the perpetrators. Moderate-strength relevant questions also address the examinee's alibi or place him at the scene of the crime.

reliability

Stability or consistency of measurement. Reliability studies in PDD often examine the rate of decision agreement among examiners on polygraph test charts. *Interrater (between rater) reliability* denotes agreement among examiners, whereas *intrarater (within rater) agreement* (test-retest reliability) pertains to an examiner agreement with his or her own decisions when evaluating the charts on different occasions. Reliability is not the same as validity, which means accuracy. A technique cannot be more valid than it is reliable since reliability constrains validity.

relief tracing

Section of a physiologic tracing in which a recovery from an arousal is apparent.

resistance

Degree to which a material hinders the flow of electricity. Skin resistance is one of the measures used in conventional polygraphs. Resistance and conductance are reciprocals of one another. See: Handler (2010).

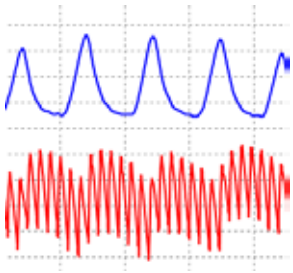
respiratory amplitude

One of several features found in the breathing pattern. Suppression is a reliable indicator of sympathetic nervous system arousal and is one of the features evaluated in the diagnosis of deception.

respiratory blood pressure fluctuations (RBPF)

An undulating waveform observed in the cardiograph channel during PDD testing. During breathing vasoconstrictor neurons are activated in the inspiratory phase leading to rhythmic vasoconstriction of blood vessels controlling blood pressure. Increased vasomotor constriction results in increased blood pressure. Additionally, deep breathing results in negative pressure in the venae cavae resulting in increased blood flow. Increased blood flow results in a larger "pre-load" (the amount of blood returning to the right side of the heart) or end diastolic volume which leads to increased blood pressure. Respiratory Sinus Arrhythmia (RSA) results in increased heart rate during the inspiration cycle. Increased

heart rate results in increased cardiac output which, in turn, results in increased blood pressure. When an examinee engages in a deliberate attempt to control his or her breathing one can expect to see an exacerbation of any cyclic waveform in the cardiovascular channel. The synchronous rise and fall of the relative blood pressure is quite possibly a result of any combination of these physiological factors. The image below shows the undulations in the cardiograph tracing and its relationship with breathing. Note that the timing of cardiograph undulations follows the breathing by about 2 – 3 seconds. See: Handler & Reicherter (2008).



respiratory cycle time (RCT)

One of several diagnostic features found in the respiratory pattern. The RCT is calculated by measuring the time or linear chart distance between two inspiratory maximums, divided by the same measure of two inspiratory peaks after the subject's vocal response. Lower ratios indicate longer breathing cycle time after stimulus presentation and, therefore, sympathetic nervous system arousal. This principle presupposes stable respiratory patterns.

respiration line length (RLL)

Feature in a breathing tracing that changes during arousal. It is a linear measure of the breathing waveform over a specified period time. RLLs are inversely related to autonomic nervous system activation. This is because the typical breathing response to stimulation or orientation is a suppression in the rate and amplitude of breathing. It was first reported in the PDD literature by Dr. Howard Timm, and several studies have since supported the use of this measure in deception tests. The use of RLL as a deception criterion presupposes stable respiratory patterns. See: Timm (1982).

respiratory sinus arrhythmia (RSA)

Variations in heartbeat directly related to breathing; slower during inspiration and faster during expiration.

rise time

Period between the start of a response to its greatest amplitude.

response onset window

Beginning of the period after question onset in which physiological responses are considered for analysis and scoring.

sacrifice relevant question

Question introduced by Cleve Backster and used in most forms of the ZCT as well as other types of tests. The sacrifice relevant is a question that asks the examinee if he intends to answer truthfully to every question related to the relevant issue. Its putative role is to dissipate the responses of innocent persons who frequently react to the first relevant question. The sacrifice relevant question is not numerically scored. Its value has been disputed in independent research. See: Capps, (1991); Horvath (1994).

salience

The state or quality of standing out relative to other stimuli. It is a vital unconscious process that helps an organism efficiently use limited attentional resources, and facilitates survival. In learning theory, salience refers to the strength of the relationship between a response and a reinforcer or outcome. In general, as the intensity of the outcome increases, the intensity of the response increases. In the framework of PDD testing, the intensity of arousal will increase commensurate with perceived salience of stimuli as they relate to the subject's goals, standards and attitudes.

scientific cradle

Device built by Angelo Mosso to record respiratory and cardiovascular responses to fear. It consisted of a fulcrum base onto which was placed a platform for his subject to lie upon. Using counterweights to bring the platform into balance, the device recorded on a smoked drum using a kymographion the changing balance of the platform that accompanied the undulations of breathing and the shifting concentration of blood in the body. The image below is taken from Mosso's book *Fear* (1896).

screening examination

A polygraph examination conducted in the absence of a reported issue or allegation to investigate whether an examinee has withheld information regarding engagement in behaviors encompassed by the relevant questions that cover specified periods of time. Screening examinations may be designed to investigate both multiple and single types of behavior. The strength of screening examinations is in their utility to develop significant information that is most often not obtainable from any other source. Its weakness is that it is not as powerful an examination as is the specific issue test in terms of validity and reliability. See: Krapohl & Stern (2003).

Searching (or Probing) Peak of Tension (SPOT)

Peak of Tension test in which the testing examiner does not know the critical item, and it is used to determine information concealed by a guilty examinee. Applications of this format include determining the location of stolen goods or the bodies of murder victims, the amount of cash stolen, or the name of an accomplice. There is no published research to support the SPOT.

secondary track

One of four tracks in the Quadri-Track Zone Comparison Technique, which includes the primary, outside, and inside tracks. The secondary track contains a relevant question, and a non-current exclusive probable-lie comparison question against which the relevant question is compared. See: Matte (1996).

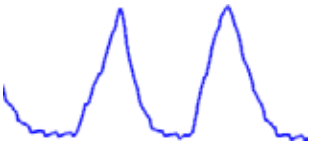
sensitivity

Ability of a test to detect specific features at all levels of magnitude or prevalence. Mathematically, sensitivity can be calculated by dividing the number of true positives by the sum of true positives and false negatives. In PDD testing this term is used to describe how well a test identifies a person engaging in deception to the issue under investigation. It is a measure of "true positive" results generally expressed as a decimal (for example a sensitivity of 0.90 would indicate a particular test identified 90% of the liars.)

sensory nerves
See: *afferent nerves*.

Senter Rules
Also called two-stage rules, Senter Rules are used for single-issue testing. The Senter Rules begin by basing a decision on the total score for the case. If a decision would be inconclusive from the total score rule, the second stage is used in which the spot scores are considered. The net effect of the Senter Rules is to decrease the proportion of inconclusive results while not affecting the proportion of correct decisions. See: Senter & Dollins (2002).

serrated breathing pattern
The breathing tracing that also includes the examinee's pulse due to the proximity of the breathing sensors to the heart. The image below is an example of a serrated breathing pattern.



sexual history examination
A form of *Post-Conviction Sex Offender Testing* (PCSOT) which entails an in-depth look at the entire life cycle of an offender and his or her sexual behaviors up to the date of criminal conviction. Sometimes referred to as a *disclosure examination*. See: Cooley-Towel, Pasini-Hill, & Patrick (2000); Dutton, (2000); English, Pullen, & Jones (1996); Heil, Ahlmeyer, McCullar, & McKee (2000).

sexual offender monitoring
Use of the polygraph to verify that sexual offenders on parole or probation are in compliance with the conditions of their release from incarceration.

signal detection theory (SDT)
One approach used to quantify the capacity of a test or method to discriminate between signal and noise. Its greatest value has been in the field of diagnostics, including polygraphy. Using SDT, optimal cutting scores can be calculated that correspond with the costs and benefits of errors. See: Swets (1995) and Green & Swets (1988).

signal value
The perceived significance of a stimulus to an organism, and is related to the concept of salience. Significant stimuli (those with signal value) can elicit physiological responses, and greater signal value corresponds with greater response magnitude. External significance is assigned to a stimulus when it appears to differ from others based on appearance. In polygraphy, this could be when a test question is much longer or is read in a louder tone of voice. Internal significance is assigned to a stimulus due to its meaning. An objective of a CQT examination is to make the external significance of relevant and comparison questions appear equal, and for their internal significance to vary. An innocent examinee would be expected to find higher internal significance in the comparison questions, whereas the relevant questions would hold higher internal significance for the deceptive. See: Handler & Honts (2007).

Significant Physiological Responses (SPR or SR)

Accepted verbiage in the US polygraph screening programs and is equivalent to *Deception Indicated* in specific-issue tests. This alternate language comes from an acceptance that screening examinations do not produce the high validity of specific-issue tests and, therefore, the results are better reported as the presence of physiologic arousals rather than inferring an examinee's deceptive intent.

silent answer test (SAT)

Specialized procedure in which the examinee is directed to answer to himself instead of making a verbal response. The use of the SAT is prescribed by some PDD experts to help avoid distortions to the pneumograph tracings attributable to speech disorders, or to uncover certain countermeasures. To ensure examinees are attending to the content of the test questions, some PDD examiners instruct the examinees to indicate their responses by slightly nodding or shaking their heads. When the SAT is used with head movements, it is called an SAT Nod. See: Horvath (1972).

Silent Talker

A new-generation "lie detector" which extracts measures of body movements and gestures during recorded videos to base assessments of credibility. Decisions are automated using an Artificial Neural Network. Evidence for its accuracy is limited. See: Rothwell, Bandar, O'Shea, & McLean (2006).

single-issue examination

An event specific or a screening polygraph examination conducted in response to a single known or alleged incident for which the examinee is suspected of involvement or to investigate the examinee's possible involvement in a single behavioral concern for which there is no known or alleged incident. When used in screening, a single-issue examination typically follows a multiple or mixed issue screening examination in the successive hurdles model.

situational comparison question

Question used for comparison in the Modified Relevant/Irrelevant (MRI) technique that elicits physiologic responses for comparison against those of the relevant question. Situational comparison questions are a departure from their conventional counterparts in that the examinee is not faced with a question to which he is deceiving or uncertain. Rather, this type of comparison question addresses how the examinee is linked to the crime, such as having legitimate access to money that later disappeared or being the last person to see a murder victim alive. The truthful examinee is allowed to respond to a question related to the crime that is not a relevant question, but one for which the examinee may feel uncomfortable because it put him on the list of suspects.

skin conductance (SC)

Broad term for two exosomatic electrodermal phenomena, *skin conductance level* and *skin conductance response*. See: Handler et al. (2010).

skin conductance level (SCL)

Basal conductance of the skin. SCL is the tonic measure of SC.

skin conductance response (SCR)

A change in the electrical conductance of the skin elicited by a stimulus. SCR is a *phasic response*.

skin potential (SP)

General term for the endosomatic electrodermal properties of the skin. Though not currently used in PDD, preliminary laboratory research has shown SP to be as diagnostic as the exosomatic measures. See: Handler (2010); Kircher & Raskin (1988).

skin potential level (SPL)

Basal electropotential of the skin. SPL is the *tonic level*.

skin potential response (SPR):

An endosomatically produced electrodermal response, and of interest as a parameter in PDD. It is frequently measured between the forearm and the hypothenar eminence. (Special note: In the Japanese PDD literature the term *skin potential* was sometimes reported in the English translation when *skin resistance* was actually used in the study.)

skin resistance (SR)

General term for the phenomena of skin resistance level and skin resistance response. SR is recorded exosomatically and was the primary means of detecting electrodermal activity throughout much of PDD history until the introduction of computerized instrumentation with SC. See: Handler (2010).

skin resistance level (SRL)

Tonic level of electrical resistance of the skin.

skin resistance response (SRR)

Phasic response measured by electrical resistance of the skin.

S-K-Y

Abbreviation for *Suspect, Know, You*. A standardized Zone Comparison Technique test format that is included under the umbrella of the Backster Zone Comparison Technique. In a structured format, the S-K-Y allows for broadening of the scope of a single-issue test to include questions relating to secondary involvement and knowledge. Along with the questions addressing direct involvement, such as "Did you shoot Henry Jones?", other moderate strength relevant questions could be used, such as "Do you know for sure who shot Henry Jones?" or tertiary issues such as those that place the examinee at the crime scene, address his alibi, or tests prior knowledge. The S-K-Y is rarely used today.

smoked drum recording

An antecedent to the continuous strip chart. A cylinder was wrapped with paper that had been covered with soot produced by a yellow flame. The paper-covered cylinder was rotated against a stylus at a selected rate to produce a graphic recording of physiologic and other events. The recording was preserved by a coat of shellac varnish. The accompanying image is taken from page 153 in Marey's 1885 book *Méthode Graphique Dans Les Sciences Expérimentales Et Principalement En Physiologie Et En Médecine*.

smooth muscles

All of the involuntary nonstriated muscles involved in autonomic functions, except the heart. Smooth muscles are located in the bladder, intestines, and blood vessels. Also called *unstriated muscles*.

Spearman Rank Correlation

Statistical technique for testing the correlation between ordinal-level data (See: *scale of data*). This method can be used to assess interrater reliability for rankings to questions assigned by different evaluators.

specificity

A term most used in the scientific literature to describe the selectivity of a test. It is the proportion of true negatives a test can produce. Mathematically, specificity is the number of true negatives divided by the sum of true negatives and false positives. The specificity of a test will determine its efficiency. If a PDD test can detect deception 100% of the time, but has a high false positive rate, it does not have good specificity and would have lower validity. Specificity and sensitivity are dimensions that characterize the validity of a test.

specific issue polygraph examination

A single-issue PDD examination, almost always administered in conjunction with a criminal investigation, and usually addresses a single issue. Sometimes called a *specific* by PDD practitioners to differentiate from preemployment or periodic testing.

specific point countermeasures

Attempts to defeat a polygraph examination by the self-induction of physiological responses to particular questions. Typical strategies include targeting comparison questions with manipulated respiration, self-inflicted pain, covert tensing of muscles, and sometimes mental imagery. See: Honts & Amato (2002).

sphygmograph

An instrument for graphically recording arterial pulse and blood pressure. A more precise term for the *cardiograph channel* in PDD.

sphygmomanometer

Aneroid gauge used to register air pressure in the polygraph cardiovascular sensor system. Changes in a closed sphygmomanometer system signal changes in relative blood volume at the recording site on the examinee.

spontaneous countermeasures

Efforts used to interfere with PDD testing for which the examinee has not prepared in advance. These types of countermeasures have been shown in laboratory studies to be ineffective. See: Honts & Amato (2002).

spontaneous response

Any reaction not associated with an identifiable stimulus. A high incidence of spontaneous responses can be used as an index of general arousal. Also known as a non-specific response.

spot

A permanently assigned location of a relevant question in a CQT question series.

spot analysis

The numerical evaluation of a relevant question by comparing it to a comparison question to the left or right of that spot location. A "spot" represents the location of a relevant question in a question series; the physiological data at the relevant question (spot) are compared with those of an adjacent comparison question.

spot responder

An as-yet unproven concept wherein an examinee has a propensity to physiologically respond to a question by virtue of its position in a sequence of questions rather than the content of the question. Some techniques rotate the serial position of relevant questions because of the concern of spot responses.

stabile

Resistant to change. Opposed to *labile.*

staircase respiration

Expression for a pattern of respiration in which the tops of the inhalation cycles move higher or lower with each subsequent cycle, forming the characteristic "staircase" appearance. The increase in respiratory amplitude is called an ascending staircase, while the opposite pattern is a descending staircase. The descending version of this pattern has been found diagnostic in at least one study. See: Kircher, Kristjansson, Gardner, & Webb, A. (2005).

standard deviation

Statistical term for a standardized unit of dispersion of scores. When scores are clustered closely together, the standard deviation is small, whereas a wide spread would have a larger standard deviation. Mathematically, the standard deviation is the square root of the variance. Conceptually, the standard deviation is the square root of the average squared deviation from the mean.

statistical significance

Phrase to describe an experimental result that is unlikely to have occurred by chance. In PDD, conventional thresholds of statistical significance are 0.05 and 0.01.

stim marker

See: *event marker*.

stimulation test

Procedure used by many PDD examiners before or between the regular tests. One of its purposes is to demonstrate to examinees that the polygraph works with them, and in doing so, reassure the innocent while heightening the guilty person's concern about the relevant questions. Other purposes include allowing the testing examiner to set the gains properly, to verify that the sensors are properly placed and functioning correctly, and to acquaint the examinee with the examination procedures. Virtually all stimulation tests use a question set of very similar items in which is embedded a single item that the subject is directed to lie about. There are several types of stimulation tests. The more common are the *card test*, *known numbers acquaintance test*, *blind numbers test*, *control test*, and *true blue control test*. Tests of this nature were used by early polygraph examiners for the purpose of comparing the reactions on the stimulation test with those from the relevant questions on the R/I test. Stimulation tests were sometimes referred to in general as *stim tests*.

Stoelting Instruments
C.H. Stoelting of Chicago, Illinois. An American manufacturer of analog and computer polygraphs.

straight-through test (ST)
A Reid polygraph test with a standard sequence of questions, usually conducted before other sequences are used. In contrast to *Reid's Mixed Question Test* (*MQT*). See: Reid & Inbau (1977).

strain gauge
Generally, any sensor for registering changes occurring in the dimensions of a solid or body. In PDD, respiration is sometimes recorded with pneumatic strain gauges placed about the thorax or abdomen, or both.

striated muscles
Includes all skeletal muscles that perform voluntary contractions, as well as cardiac muscle responsible for the involuntary muscular activity of the heart. Also called *striped muscles*.

stroke volume (SV)
Measure of the amount of blood ejected from the heart in each beat.

strong relevant question
PDD question that goes directly to the heart of the matter under investigation, as compared to knowledge or complicity questions.

subjective lie/truth question
In the Positive Control Technique, examinees are presented with each question two times in succession as a set. Examinees are instructed to admit committing the offense under investigation after the first presentation of each question and deny it after the second reading of the same question. The first time the question is read, it is called the subjective lie question, while the second presentation is called the subjective truth question. See: Driscoll, Honts & Jones (1987); Gordon & Cochetti (1982); Howland (1981); Reali (1978).

successive hurdles approach
In screening, it is a method to maximize accuracy when base rates are unbalanced or very low. For example, an agency may have a vital interest in uncovering past activities of applicants before hiring them, behaviors that are very difficult to investigate effectively with any method except the polygraph. The agency may choose to conduct a single-issue polygraph for each of the subject areas. If there only two or three areas, this approach may be effective, however, if the number is greater, it would tax the agency's resources in terms of polygraph examiners and processing time. Alternatively, in the successive hurdles approach, each applicant would begin with a multiple-issue examination. Though multiple-issue examinations have a higher false positive rate (incorrect decision of deception) than do single-issue examinations, they have a very small rate of false negatives. If no significant responses were noted during the multiple-issue examination, the polygraph session would be over. However, if there were significant responses during the multiple-issue examination, the polygraph session would continue. Next, a single- issue examination would be administered in the area the examiner observed significant responses during the multiple-issue examination. The single-issue examination has much better discrimination power and would help to elucidate physiological arousal more specifically. In this way, the more resource-intensive single-issue examinations would be reserved for the smaller subset of applicants who did not pass the multiple-issue examination. The net effect of the two-stage screening process is a better accuracy without increased resources. See: Krapohl & Stern (2003); Meehl & Rosen (1955).

Summers, Walter

Early researcher of deception testing who used an electrodermal device and a structured test series to verify truthfulness and deception. Summers reported the use of what he called *emotional standards*, which are similar in many respects to present day comparison questions. See: Summers (1939).

super-dampening concept

Concept forwarded by Cleve Backster which holds that there will be a suppression of general reactivity to relevant and comparison questions when the examinee is more concerned that the examiner will ask an unreviewed question about another issue outside the scope of the current examination questions. Though once widely accepted, the available evidence suggests that the effect is negligible or nonexistent. See: Backster, (1964); Capps, Knill, & Evans (1993); Honts, Amato, & Gordon (2000); Krapohl (2001); Matte (2001).

suppression

One breathing response pattern indicative of orienting and arousal characterized by breathing that is shallower and slower than tracing average. The PDD breathing tracing of suppression will manifest in a decrease in amplitude, a slower rate, or a temporary increase in the baseline of the waveform. Suppression has long been found to be a reliable indicator of salience and a primary reaction criterion for scoring.

surreptitious breathing tracing

Recording of the breathing cycles of an examinee during a portion of the polygraph session when the examinee is not aware of it. Frequently, though not always, this is done just before or immediately after a test is conducted. Surreptitious breathing recordings are useful to help assess whether an examinee has altered his breathing during testing. Differences can indicate countermeasures. The motives of the examinee may be discerned with a Yes Test or other specialized techniques that help assess the examinee's level of cooperation.

sympathetic nervous system

Thoracolumbar portion of the autonomic nervous system centrally involved in responding to arousing stimuli. Most sympathetic nerves are *adrenergic* and prepare the body to respond to increased demands. Sympathetic nervous activation increases blood flow from the heart, triggers the release of glucose and epinephrine, dilates the pupils, and initiates other responses in preparation for action. Unlike most of the sympathetic nervous system, sympathetic nerves to the eccrine sweat glands are *cholinergic*.

sympathetic chain

A system of 21 to 22 pairs of ganglia located in the thoracic and abdominal areas and is the site of synapse between pre- and postganglionic sympathetic neurons. There is one notable exception—there is no postganglionic sympathetic nerve to adrenal medulla.

sympathomimetic

Drugs that mimic the action of sympathetic postganglionic nerves or their neurotransmitters.

symptomatic question

A question type developed by Cleve Backster that was once thought to identify whether an examinee is fearful the examiner will ask an unreviewed question about an outside issue. In this construct, the examinee's mistrust would dampen his responses to other test questions, and the symptomatic question could determine whether the lack of responsivity was attributable to the outside issue. Symptomatic questions are widely used, though the trend in the research finds they have no meaningful effect. See: Backster (2001a); Honts, Amato, & Gordon (2000); Krapohl & Ryan (2001); Matte (2001).

synapse

Junction between neurons. Site where the nervous impulse is transferred from one neuron to another. Neurotransmitters reside in vesicles of one neuron and are released by the axon into the synapse to chemically induce the next neuron or organ to respond.

systole

Contraction of the cardiac muscles. Left ventricular systole results in movement of blood out of the heart into the aorta. Systoles can be subdivided into three primary components: *preejection*, *ejection*, and *relaxation periods*. The left ventricular systolic peak is represented on a conventional polygraph as the highest vertical point in the pulse wave of the cardiovascular tracing.

systolic blood pressure

Force exerted by blood against the wall of the arteries at the height of ventricular contraction. Also called *maximum pressure* and expressed in millimeters of mercury (mmHg).

t test

A statistical test frequently used with small samples (when the number of observations is less than 30) to determine whether the mean of one sample is significantly different from that of other.

tachycardia

Abnormally rapid heartbeat, greater than 100 beats per minute. Tachycardia can result from poisoning, medications, certain illnesses, and during states of anxiety or excitement. A rapid heartbeat is normal for the very young or during the physical exercise of normal adults. While not a diagnostic feature in itself, PDD examiners pay attention to heart rate during evaluation of the polygraphs as it may signal pharmaceutical countermeasures or atypical levels of stress.

tachypnea

Rapid breathing, usually shallow.

technique

All practices taking place in a polygraph examination, including pretest procedures, question formulation, format, number of tests, test sequencing, and scoring and decision rules.

test

In PDD, the *test* is used to differentiate a single running of a question series (sometimes also called a *chart*) during physiological recording from the examination, which is considered to be the totality of the PDD process. It can also refer to specialized procedures within techniques, such as the Yes Test and stimulation test. *Test* has been erroneously used to refer to polygraph techniques, such as the Zone Comparison Test, or Modified General Question Test.

test data analysis (TDA)

Newer expression for polygraph chart interpretation, a change prompted by digital polygraphs where physiological data are displayed on computer screens rather than paper strip charts.

Test for Espionage and Sabotage (TES).

Multiple-issue testing format employed by some U.S. Government agencies for screening purposes. The TES uses a repeated series of relevant and directed-lie comparison questions, and the conventional 7-position scoring system. See: Reed (1994); Research Staff (1995;1998). Also see Directed Lie Screening Test

thenar eminence

Prominence on the palm at the base of the thumb. One of the optimal recording sites for electrodermal activity, and a preferred location in psychophysiological research. See: Handler (2010).

thermal imaging

A technique that uses a camera to record the emission of radiant energy from the body. The basis for the technology is that any object with a temperature greater than 0 degrees Kelvin radiates energy in the infrared (IR) range. Although IR energy cannot be seen by the human eye, thermal cameras are equipped to record this form of energy. Changes in thermal patterns on the face have been shown to be associated with physiological arousal, and preliminary evidence suggests that thermal imaging can discriminate between truthtellers and deceivers at better than chance using a probable-lie comparison question technique. See: Pavlidis, Eberhardt & Levine (2002); Pollina (2006); Pollina & Ryan (2003).

thoraco-lumbar division of autonomic nervous system

An anatomical division of the autonomic nervous system (ANS) that represents the sites of outflow from the sympathetic division of the ANS (i.e., nerves from the thoracic and abdominal parts of the body). It is the location of the 21-22 pairs of ganglia that constitute the sympathetic chain.

thumb transducer

Cardiovascular sensor consisting of an external ring and internal bladder, into which is placed the thumb. The bladder is inflated to a small pressure, and the sensor detects changes in the relative blood volume of the thumb. Because of the weakness of the signal, the thumb transducer requires an electronically enhanced cardiograph channel.

time bar

One method of restricting the coverage of the comparison question so that it will not include the time in which the incident under investigation took place. Time bars generally predate the crime. A typical time bar for the comparison question might be phrased "Before the age of X, did you ever..." or "Prior to 1998, did you ever …" There is a school of thought that examinees may confuse the relevant questions with the comparison questions unless these two types of questions are designed to avoid any degree of overlap. Research has not supported this hypothesis, however. See: Amsel (1999); Podlesny & Raskin (1978); Horvath (1988); Horvath & Palmatier (2008). Also see *exclusionary comparison question*

tonic change

Shifting of tonic level to a new baseline, typically at a relatively slow rate compared to phasic responses.

tonic level

Baseline level. This terminology in PDD is frequently used to delineate basal waveform levels from short-term responses induced by stimuli. Tonic levels change slowly compared to phasic activity.

tonic response

Shifting of tonic level, typically in response to changing conditions. For example, the adjustment of electrodermal tonic levels due to temperature changes, reduction of pulse rate between standing and reclining, and faster breathing that accompanies an increase in walking speed. Tonic responses take several seconds or minutes to occur, unlike phasic responses which tend to be very rapid. Some examiners consider changes in tonic activity as diagnostic information in Peak of Tension tests, where a change in the trend of tonic activity can signal that the critical item in the series has passed. Research evidence is lacking despite frequent anecdotal reports.

total chart minutes concept (TCMC)

A concept offered to consider possible variation in the rate of habituation for the individual physiological parameters recorded in PDD. Backster developed a habituation curve for each of the parameters for the amount of time these activities were recorded during testing and published his report in 1963. It has not received much attention among researchers and is not currently taught in the field. See: Backster (1963b; 2001b); Krapohl (2001).

tracing average

Section of the physiological tracing used as a comparative baseline, and in which there are no indications of physiologic arousal. Also referred to as the *baseline level*.

tracing distortion

Contamination of a physiologic tracing, typically by movements of an examinee. Also referred to as *artifact*.

track

Term relating to pairs of polygraph test questions that are used for evaluative purposes, and a concept central to the Quadri-Track Zone Comparison. See: Matte (1996)

transducer

Device for transforming energy of one type into energy of another type. In PDD a transducer is most often used in the context of converting pressure changes into electrical signals.

Traube-Hering-Mayer (THM) wave

Rhythmic low-frequency fluctuation in heart rate (~0.1 hz) linked to a blood pressure feedback loops and involving the carotid baroreceptors. The THM wave is a contributing component to the variability expressed in the respiratory sinus arrhythmia. See: *vagal tone*.

tremograph

Instrument for recording tremors and an early deception detection approach. In the 1930s Luria proposed that trembling could be used as index of emotional arousal, leading to possibilities in detecting lying. His hypothesis was based on his conflict theory (i.e., incompatible emotions affect bodily responses). See: *conflict theory*.

true negative

Correct decision that the variable of interest is not present (i.e., an accurate PDD outcome of innocence).

true positive

Correct decision that the variable of interest is present (i.e., an accurate PDD outcome of guilt).

two-stage rules

Two-Stage Rules function with the sequential use of the Grand Total Rule followed by Spot-Score Rules. Subtotal scores are not permitted to supersede the grand total score. Two-Stage Rules have been shown to reduce the occurrence of inconclusive results and with no decrement to the proportion of correct decisions. See: Senter Rules.

type I error

The probability of rejecting a true null hypothesis. Alternatively, it is the probability of incorrectly accepting the alternate hypothesis. Denoted as α and in PDD consider a false-positive test result.

type II error

The probability of not rejecting a false null hypothesis and denoted as β. In PDD it is consider a false-negative test result.

United States Army Military Police School (USAMPS)

Former polygraph instruction center for U.S. Federal PDD examiners from 1951 through 1986. Now called the National Center for Credibility Assessment (NCCA).

United States v Frye

James Alphonzo Frye was administered a deception test by Dr. William M. Marston in 1923 using Marston's discontinuous blood pressure method. Marston's opinion was that Frye was truthful in his recanting of a murder confession for the killing of a prominent Washington, DC, doctor. At Frye's trial his defense attorneys attempted to have the results entered into evidence but were unsuccessful (United States v Frye 54 App D.C.46, 293 F 1013). The Frye Rule, as it came to be known, stated that "expert testimony based on a scientific technique is inadmissible unless the technique is generally accepted as reliable in the relevant scientific community." The Frye Rule was invoked thereafter in many jurisdictions to bar PDD evidence from admissibility. The Frye Rule has been superseded by Federal Rule 702, cited in Daubert v Merrell Dow Pharmaceuticals, Inc (1993), except in those states that do not follow the Federal Rules of Evidence. See: Daubert v Merrell Dow Pharmaceuticals, Inc (1993): Stern & Krapohl (2003).

Utah Probable Lie Technique (UPLT)

A technique developed by researchers at the University of Utah beginning in the 1970s, and was initially influenced by the Backster Zone Comparison Test. It differs from other ZCT formats in that the relevant questions are not bracketed by comparison questions. Other unique characteristics of the UPLT are: the inclusion of the photoplethysmograph; rotation of the probable-lie questions; recording of five charts when the test would be inconclusive at three charts; and symmetrical cutoffs of +/-6. The UPLT has been validated in mock crime analog studies and has more peer-reviewed publications supporting it than any other probable-lie technique. More recently the developers of the Utah technique have also endorsed the use of directed-lie questions in place of probable-lie questions. See: Handler, (2006); Raskin & Honts (2002).

vagal tone

A measure of parasympathetic influences on the heart as measured by the variability of inter-beat intervals in heart rate (R-to-R wave intervals on an electrocardiogram). This variability, a function of respiratory sinus arrhythmia (RSA), fluctuates at the same frequency as respiratory activity. Characteristically, the inter-beat interval will begin with inhalation, and shorten with exhalation. It has been investigated as a possible alternative measure of stress in a PDD setting. See Miller (1994).

vagus nerve

Tenth cranial nerve that provides parasympathetic innervation to the heart and other visceral organs.

vagus roll (or pattern)

See: respiratory blood pressure fluctuations (RBPF)

validity

Accuracy. There are several types of validity. The degree to which a test measures what it professes to measure is *construct validity*. *External validity* relates to the generalizability of the research results out of the laboratory. While there are other types of validity as well, these two types go to the heart of research in PDD.

Valsalva's maneuver

Performed by forced exhalation against a closed glottis, thus increasing internal abdominal and thoracic pressure. This results in a short-term increase in blood pressure and frequently a phasic electrodermal response. Because Valsalva's maneuver can directly affect physiologic channels recorded with the polygraph, it is considered a type of physical countermeasures when used during testing. Named for eighteenth century Italian anatomist Antonio Maria Valsalva.

variance

A standardized measure of dispersion and is the average squared deviation from the mean. Mathematically, the variance of a sample or population is equal to the square of the standard deviation. Variance is a useful measure for tests of effects.

variable

Variables in the behavioral science context are those elements that influence or are influenced by something. For example, habituation is a variable that affects responsivity.

variograph

Term sometimes used in the Polish literature to denote the polygraph.

vasoconstriction

Narrowing of blood vessels, especially arterioles, thereby reducing blood flow to a region of the body. During sympathetic nervous system activation vasoconstriction takes place in some parts of the body, shunting blood supplies to major muscles in preparation for defense or flight. Some medications influence vasoconstriction.

vasodilation

Expansion of blood vessels, especially arterioles, which deliver increased blood flow to a region of the body. Vasodilation and vasoconstriction are regulated by the *autonomic nervous system* and can be affected by some medications.

vasomotor

Relating to the influences of the smooth muscles on the internal diameter of a blood vessel.

ventilation

Regular movement of the chest cavity during the act of breathing.

vertical scoring system

Numerical evaluation method in which spots are individually scored by relevant-comparison question assessments. Vertical scoring is used in the ZCT exploratory and S-K-Y tests. Term generally used in contrast to *rank order analysis*.

voice stress analysis

Any analytical technique implemented to determine whether changes in the vocal signal are indicative of changing levels of stress. Most techniques assess the frequency or amplitude modulation of the vocal signal in one or more frequency bandwidths. Emphasis is often placed on modulation in the 8-10 Hz frequency bandwidth, otherwise known as *microtremors*. Numerous voice stress analysis devices have been introduced since the first was made available in 1971. All purport the ability to detect deception. Today, this genre of credibility assessment devices enjoys a wide distribution, possibly due to their low cost and brief training requirements relative to the polygraph. However, no independent scientific assessment has validated the use of voice stress analysis for credibility assessment. Because of this, these devices are prohibited from use by the US Department of Defense. Brand names include C*omputer Voice Stress Analyzer (CVSA), Lantern, Psychological Stress Evaluator (PSE), TiPi, VSA Mark, Vericator, and Layered Voice Analysis (LVA).*

voir dire

French term meaning "to say the truth." It is the process of questioning jurors for the purpose of excluding those with biases. Pronounced "vwa dear."

Wheatstone bridge

Early electrical circuitry used to detect electrodermal activity. It employed a null-type resistance-measuring circuit in which the examinee's resistance was continually compared to that of known resisters.

Winter, John

Earliest known developer of a scoring system used to interpret polygraph recordings in the detection of deception. His method did not come into wide use and is of historical interest only. See: Winter (1936).

Yerkes-Dodson Inverted U

Yerkes and Dodson proposed in 1908 that an inverted U characterized the relationship between arousal and performance. The level of performance increases with arousal until it reaches some optimum level, beyond which performance experiences a decrement. This model implies that a subject who is under-aroused will not respond sufficiently on a PDD examination, while a subject who is over-aroused will also not respond adequately.

Yes-No Technique

First reported by this name by Richard Golden in 1969, this technique is similar to the Positive Control Technique. Also *Positive Control Technique*.

Yes Test

PDD test series attributed to John Reid in which all comparison questions are omitted from the question list, and an examinee is told to answer "yes" to all remaining test questions. It is preceded by special instruction so to invite physical countermeasures from guilty examinees. The Yes Test is especially useful in identifying examinees who wish to defeat the polygraph test and is generally used only when the examiner has reason to believe that the examinee is attempting countermeasures.

You Phase

One of the most commonly used of the formats in the Backster Zone Comparison Technique. The standardized test addresses a single issue and single degree of involvement in the issue. The format provides for two or three relevant questions, worded slightly differently from one another, addressing the single issue and degree of involvement. It also requires a repeat of the relevant question wording in the sacrifice relevant question. The You Phase ZCT is a very powerful test because it is so highly focused on

essentially one question. An example of You Phase question wording is: sacrifice relevant—"Regarding whether or not you shot Henry Jones, do you intend to answer truthfully each question about that?"; relevant 1—"Did you shoot Henry Jones?"; relevant 2—"Did you fire the shot that caused the death of Henry Jones?"; relevant 3—"Last Friday night, did you shoot Henry Jones?" At this writing the technique has been inadequately researched.

zone

Concept coined by Cleve Backster. A zone is a 20- to 35-second block of polygraph chart time initiated by a question having a unique psychological focusing appeal to a predictable group of examinees. In his ZCT, Backster used color-coding to identify the three zones in the ZCT: *red*, *green*, and *black*. Respectively, the red zone for relevant questions, the green zone for comparison questions, and the black zone for symptomatic questions.

Zone Comparison Technique (ZCT)

PDD technique developed by Cleve Backster that contains three Zones (red, green and black), with comparison of responses between two of the Zones (red and green) for a determination of truth or deception. There are several variations, including the You Phase, Exploratory, Federal Zone Comparison, Integrated Zone Comparison and Utah Probable Lie Test. The ZCT was the first modern PDD technique in general use to incorporate numerical analysis. The ZCT is probably used more often in forensic applications than any other format.

References

Abrams, S. (1977). A survey of attitudes on the guilt complex technique. *Polygraph*, *6*(1), 123-124.

Abrams, S. (1984). The penile plethysmograph: A new transducer used for detection and therapy with sexual deviation cases. *Polygraph*, *13*(2), 198-201.

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

Abrams, S. (1999). A response to Honts on the issue of the discussion of questions between charts. *Polygraph*, *28*(3), 223-228

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

Backster, C. (1963a). Anticlimax dampening concept. *Military Police Journal*, Oct, 22-23.

Backster, C. (1963b). Total chart minutes concept. *Law and Order*, *11*(10), 77-79.

Backster, C. (1963c). *Standardized polygraph notepack and technique guide: Backster zone comparison technique*. Backster: New York.

Backster, C. (1964). Outside "super-dampening" factor. *Military Police Journal*, Jan, 20-21.

Backster, C. (2001a). Comments on Krapohl & Ryan "Belated look at symptomatic questions." *Polygraph*, *30*(3), 213-215.

Backster, C. (2001b). A response to Donald Krapohl's assessment of the Total Chart Minutes Concept. *Journal of the American Association of Police Polygraphists*, *1*, 32-34.

Barland, G.H., Honts, C.R., and Barger, S.D. (1989). The validity of detection of deception for multiple issues. *Psychophysiology*, *26*(4a Supplement), 13 (Abstract).

Barland, G.H., and Raskin, D.C. (1975). Psychopathy and detection of deception in criminal suspects. *Psychophysiology*, *12*, 224 (Abstract).

Battelle Memorial Institute. (2007). *Efficacy of prototype credibility assessment technologies: PCASS final report*. Prepared for the Defense Academy for Credibility Assessment. GS-23F-0011L. Ref No. MC-RRT-06-0036.

Bell, B.G., Raskin, D.C., Honts, C.R., and Kircher, J.C. (1999). The Utah Numerical Scoring System. *Polygraph*, *28*(1), 1-9.

Ben-Shakhar, G. (1977). A further study of the dichotomization theory in detection of information. *Psychophysiology*, *14*(4). 408-413.

Benussi, V. (1914). Die atmungssymptome der lüge (The respiratory symptoms of lying). *Archiv fuer die Gesamte Psychologie*, *31*, 244-273.

Blackwell, N.J. (1998). *PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examination from actual criminal investigations*. DTIC AD Number A355504/PAA. Department of Defense Polygraph Institute, Ft. McClellan, AL. Printed in *Polygraph*, *28*(2) 149-175.

Blalock, B. (2009). Capitalizing on technology to increase standardization and reliability in a polygraph examination. *Polygraph, 38*(2), 154-166.

Blalock, B., Cushman, B., and Nelson, R. (2009). A replication and validation study on an empirically based manual scoring system. *Polygraph, 38*(4), 281-288.

Bradley, M.T., and Janisse, M.P. (1981). Accuracy demonstrations, threat, and the detection of deception: Cardiovascular, electrodermal, and pupillary measures. *Psychophysiology*, *18*(3), 307-315.

Burtt, H.E. (1918). A pneumograph for inspiration-expiration ratios. *Psychological Bulletin*, *15*(10), 325-328.

Burtt, H.E. (1921). The inspiration/expiration ratio during truth and falsehood. *Journal of Experimental Psychology*, *4*(1), 1-23.

Butta, M.R., Hong, M.J., Kim, Y., and Hong, K. (2015). Single-trial lie detection using combined fNIRS-polygraph system. *Frontiers in Psychology*, 6:707. doi: 10.3389/fpsyg.2015.00709

Capps, M.H. (1991). Predictive value of the sacrifice relevant. *Polygraph*, *20*(1), 1-6.

Capps, M.H, Knill, B.L., and Evans, R.K. (1993). Effectiveness of the symptomatic questions. *Polygraph*, *22*(4), 285-298.

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

Cooley-Towell, S., Pasini-Hill, D., Patrick, D. (2000). The value of the post-conviction polygraph: The importance of sanctions. *Polygraph*, *29*(1), 6-19.

Damphousse, K.R., Pointon, L., Upchurch, D., and Moore, R.K. (2007). *Assessing the validity of voice stress analysis tools in a jail setting*. Final report to the US Department of Justice. University of Oklahoma.

Darrow, C.W. (1932). The behavior research photopolygraph. *Journal of General Psychology, 7*, 215- 219.

Daubert v Merrell Dow Pharmaceuticals, Inc. (1993). 509 U.S. 579, 125 1.Ed 2d 469.

Dawson, M.E. (1980). Physiological detection of deception: Measurement of responses to questions and answers during countermeasure maneuvers. *Psychophysiology, 17*(1), 8-17.

Dollins, A.B., Cestaro, V.L., and Pettit, D.J. (1998). Efficacy of repeated psychophysiological detection of deception testing. *Journal of Forensic Science*, *43*(5), 1016-1023.

Dollins, A.B., Krapohl, D.J., and Dutton, D.W. (1999). *A comparison of computer programs designed to evaluate psychophysiological detection of deception examinations: Bakeoff 1*. Department of Defense Polygraph Institute, Ft. Jackson, SC. DoDPI99-R-0001, DTIC # ADA 379990.

Driscoll, L.N., Honts, C.R., and Jones, D. (1987). The validity of the positive control physiological detection of deception technique. *Journal of Police Science and Administration, 15*(1), 46-50.

Dutton, D.W., (2000). Guide for performing the objective scoring system. *Polygraph*, *29*(2), 177-184.

Dutton, D.W. (2000). Introduction. *Polygraph, 29*(1), 1-5.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. W. W. Norton & Co.: New York.

Elaad, E., and Ben-Shakhar, G. (1991). Effects of mental countermeasures on psychophysiological detection in the guilty knowledge test. *International Journal of Psychophysiology, 11*(2), 99-108.

Elkins, A., Bolob, Elyse, Nunamaker, J., Burgoon, J., and Derrick, D. (2014, Oct). *Appriasing the AVATAR for Automatic Border Control*. Available at: https://www.europarl.europa.eu/RegData/questions/reponses_qe/2019/002653/P9_RE(2019)002653(ANN3)_XL.pdf

English, K., Pullen, S., and Jones, L. (Eds)(1996). *Managing adult sex offenders: A containment approach*. American Probation and Parole Association: Lexington, KY.

Farwell, L.A., and Donchin, E. (1988). Event-related potentials in interrogative polygraphy: Analysis using bootstrapping. *Psychophysiology*, *25*(4), 445 (Abstract).

Farwell, L.A., and Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, *28*(5), 531-547.

Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378-382.

Gardner, J.W. (1937). An experimental study of the Luria technique for detecting mental conflict. J*ournal of Experimental Psychology*, 20(6), 495-506

Gastwirth, J.L. (1987). The statistical precision of medical screening procedures: Applications to polygraph and AIDS antibody test data. *Statistical Science*, *2*(3), 213-238.

Geddes, L.A. (1974). What does the photoplethysmograph indicate? *Polygraph, 3*(2), 167-176.

Ginton, A. (2009). Relevant Issue Gravity (RIG) strength - A new concept in PDD that reframes the notion of psychological set and the role of attention in CQT polygraph. *Polygraph, 38*(3). 204-217.

Gordon, N.J. (1999). The Academy for Scientific Investigative Training's horizontal scoring system and examiner's algorithm system for chart interpretation. *Polygraph, 28*(1), 56-64.

Gordon, N.J., and Cochetti, P.M. (1982). The positive control concept and technique. *Polygraph, 11*(4), 330-342.

Gordon, N.J., and Cochetti, P.M. (1987). The horizontal scoring system. *Polygraph*, *16*(2), 116-125.

Gordon, N.J., Fleisher, W.L., Morsie, H., Habib, W., Salah, K. (2000). A field validity study of the Integrated Zone Comparison Technique. *Polygraph*, *29*(3), 220-225.

Gordon, J.J, Mohamed, F.B., Faro, S.H., Platek, S.M., Ahmad, H., and Williams, J.M. (2005). Integrated zone comparison polygraph technique accuracy with scoring algorithms. *Physiology & Behavior, 87*(2), 251-254.

Green, D.M, and Swets, J.A. (1988). *Signal detection theory and psychophysics*. Peninsula Publishing: Los Altos, CA.

Handler, M. (2006). Utah Probable Lie Comparison Test. *Polygraph, 35*(3), 139-149.

Handler, M. (2010). An EDA primer for polygraph examiners. *Polygraph, 39*(2). 68-108.

Handler, M. and Honts, C.R. (2007) Psychophysiological mechanisms in deception detection: A theoretical overview. *Polygraph, 36* (4), 221-232.

Handler, M. and Krapohl, D.J. (2007). The use and benefits of the photoelectric plethysmograph in polygraph testing. *Polygraph, 36*(1), 18-25.

Handler, M., and Nelson, R. (2007). Polygraph terms for the 21st Century. *Polygraph, 36*(3), 157-164.

Handler, M. and Nelson, R. (2008) The Utah approach to comparison question polygraph testing. *European Polygraph*, 2(2).

Handler, M., Nelson, R., and Blalock, B. (2008). A focused polygraph technique for PCSOT and law enforcement screening programs. *Polygraph, 37*(2). 100-111.

Handler, M., Nelson, R., Krapohl, D.J., and Honts, C.R. (2010). An EDA primer for polygraph examiners. *Polygraph, 39*(2), 68-108.

Handler, M., Reicherter, J., Nelson, R. and Fausett, C. (2009). A respiration primer for polygraph examiners. *Polygraph, 38*(2) 130-144.

Handler, M., and Reicherter, J. (2008). Respiratory blood pressure fluctuations observed during polygraph examinations. *Polygraph, 37*(4), 256-262.

Handler, M., Rovner, L., and Nelson, R. (2008). The concept of allostasis in polygraph testing. *Polygraph, 37*(3), 228-233.

Handler, M.D., Shaw, P. and Gougler, M., (2010). Some thoughts about feelings: A study of the role of cognition and emotion in polygraph testing. *Polygraph*, *39*(3), 139-154.

Harnsberger, J.D., Hollien, H., Martin, M.D., and Hollien, K.A. (2009). Stress and deception in speech: Evaluating Layered Voice Analysis. *Journal of Forensic Sciences, 54*(3), 642-650.

Harrell, J.P, and Clark, V.R. (1985). Cardiac responses to psychological tasks: Impedance cardiographic studies. *Biological Psychology, 20*(4), 261-283.

Harwell, E. (2000). A Comparison of 3- and 7-position scoring scales with field examinations. *Polygraph, 29*(2), 195-197.

Heil, P., Ahlmeyer, S., McCullar, B., and McKee, B. (2000). Integration of polygraph testing with sexual offenders in the Colorado Department of Corrections. *Polygraph, 29*(1), 26-35.

Holden, E.J. (2000). Pre- and post-conviction polygraph: Building blocks for the future – Procedures, principles and practices. Polygraph, 29(1), 69-115.

Hollien, H., and Harnsberger, J. (2006). The use of voice security evaluations. *The Journal of Credibility Assessment and Witness Psychology, 7*(2), 74-78.

Honts, C.R. (1987). Interpreting research on polygraph countermeasures. *Journal of Police Science and Administration, 15*(3), 204-209.

Honts, C.R. (1999). The discussion of questions between list repetitions (charts) is associated with increased test accuracy. *Polygraph*, *28*(2), 117-123.

Honts, C.R. (2000). A brief note on the misleading and the inaccurate: A rejoinder to Matte (2000) with critical comments on Matte and Reuss (1999). *Polygraph*, *29*(4), 321-325.

Honts, C.R. (1996). Criterion development and validity of the CQT in field application. *Journal of General Psychology, 123*(4), 309-324.

Honts, C.R., and Amato, S.L. (2002). Countermeasures. In M. Kleiner's (ed.) *Handbook of polygraph testing*, pp. 151-264. Academic Press: London.

Honts, C.R., Amato, S.L., and Gordon, A. (2000). *Validity of outside-issue questions in the control question test*. Final report to the DoD Polygraph Institute, Grant no. N00014-98-1-0725. DTIC # ADA 376666.

Honts, C.R., and Devitt, M.K., (1992). *Bootstrap decision making for polygraph examinations*. Department of Defense Polygraph Institute, Ft. McClellan, AL. DoDPI92-R-0002.

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

Honts, C.R., Raskin, D.C., and Kircher, J.C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology, 79*(2), 252-259.

Honts, C.R., and Schweinle, W. (2009). Information gain in psychophysiological detection of detection in forensic and screening settings. *Applied Psychophysiology and Biofeedbac*k, *34*, 161-172.

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

Horvath, F.S. (1972). The polygraph silent answer test. *Journal of Criminal Law and Criminology and Police Science*, *63*(2), 285-293. Reprinted in *Polygraph*, *11*(1), 100-113.

Horvath, F.S. (1977). The effect of selected variables on interpretation of polygraph records. *Journal of Applied Psychology*, *62*(2), 127-136.

Horvath, F.S. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception. *Journal of Applied Psychology*, *63*(3), 338-344.

Horvath, F.S. (1979). Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response. *Journal of Applied Psychology*, *64*(3), 323-330.

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

Horvath, F.S. (1994). The value and effectiveness of the sacrifice relevant question: An empirical assessment. *Polygraph*, *23*(4), 261-279.

Horvath, F.S., and Palmatier, J.J. (2008). Effect of two types of control questions and two question formats on the outcomes of polygraph examinations. *Journal of Forensic Sciences, 53*(4):889-99.

Howland, D.P. (1981). Positive control question technique pre-test interview and chart interpretation. *Polygraph*, *10*(1), 37-41.

Iacono, W.G. (1991). Can we determine the accuracy of the polygraph tests? In J.R. Jennings, P.K. Ackles, & M.G.H. Coles (Eds) *Advances in Psychophysiology*, *4*, 202-208. Jessica Kingsley Publishers: London.

Ishida, J., and Sevilla, C.M. (1981). The friendly polygrapher concept and admissibility. *Polygraph*, *10*(3), 175-178.

Jones, E.E., and Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*(5), 349-364.

Karpman, B. (1949). Lying - A minor inquiry into the ethics of neurotic and psychopathic behavior. *Journal of Criminal Law and Criminology*, *40*(2), 135-157.

Kircher, J.C., Kristjansson, S.D., Gardner, M.K., and Webb, A. (2005). *Human and computer decision-making in the psychophysiological detection of deception*. Final report to the U.S. Department of Defense. Salt Lake City: University of Utah, Department of Educational Psychology.

Kircher, J.C., Packard T., Bell, B.G., and Bernhardt, P.C. (2003). Effects of deception on tonic autonomic arousal. *Polygraph, 32*(3), 166-187.

Kircher, J.C., and Raskin, D.C. (1983). Clinical versus statistical lie detection revisited - through a lens sharply. *Psychophysiology*, *20*(4), 452.

Kircher, J.C., and Raskin, D.C. (1987). Comment: Base rates and the statistical precision of polygraph tests in various applications. *Statistical Science*, *2*(3), 226-238.

Kircher, J.C., and 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., and Raskin, D.C., and Honts, C.R. (1984). Electrodermal habituation in the detection of deception. *Psychophysiology*, *21*(5), 585 (Abstract).

Kircher, J.C., Raskin, D.C., Honts, C.R., and Horowitz, S.W. (1995). Lens model analysis of decision making by field polygraph examiners. *Psychophysiology*, *32* S1, S45 (Abstract).

Kircher, J.C., Woltz, D.J., Bell, B.G., and Bernhardt, P.C. (2006). Effects of audiovisual presentations of test questions during Relevant/Irrelevant polygraph examinations and new measures. *Polygraph, 35*(1), 25-54.

Kozel, F.A., Johnson, K.A., Grenesko, E.L., Laken, S.J., Koze, S., Lu, X, Pollina, D., Ryan, A., and George, M.S. (2009). Functional MRI detection of deception after committing a mock sabotage crime. *Journal of Forensic Sciences, 54*(1), 220-231.

Krapohl, D.J. (1996). A taxonomy of polygraph countermeasures. *Polygraph*, 25(1), 35-56.

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

Krapohl, D.J. (2020). A brief comment on the inhalation/exhalation ratios in polygraph scoring. *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice, 49*(2), 79 – 81.

Krapohl, D.J. (2000, Oct). An assessment of the Total Chart Minutes Concept with field data. *Journal of the American Association of Police Polygraphists*, *4*, 31-37.

Krapohl, D.J. (2001). A brief rejoinder to Matte and Grove regarding "psychological set." *Polygraph*, *30*(3), 203-205.

Krapohl, D.J. (2005). Polygraph decision rules for evidentiary and paired-testing (Marin Protocol) applications. *Polygraph*, *34*(3) 184-192.

Krapohl, D.J., and Cushman, B. (2006). Comparison of evidentiary and investigative decision rules: A replication. *Polygraph*, *35*(1), 55-63.

Krapohl, D.J., Dutton, D.W., and Ryan, A.H. (2001). The Rank Order Scoring System: Replication and extension with field data. *Polygraph*, *30*(3), 172-181.

Krapohl, D.J., Gordon, N.J., and Lombardi, C. (2008). Accuracy demonstration of the Horizontal Scoring System using field cases conducted with the Federal Zone Comparison Technique. *Polygraph, 37*(4). 263-268.

Krapohl, D.J., Grubin, D., Benson, T., and Morris, B. (2020). Modification of the AFMGQT to accommodate single-issue screening: The British One-issue Screening Test. *Polygraph & Forensic Credibility Assessment: Journal of Science and Field Practice, 49*(2), 176 – 183.

Krapohl, D.J., and McManus, B. (1999). An objective method for manually scoring polygraph data. *Polygraph, 28*(3), 209-222.

Krapohl, D.J., McCloughan, J.B., and Senter, S.M. (2006). How to use the Concealed Information Test. *Polygraph*, *35*(3), 123-138.

Krapohl, D.J. and Norris, W.F. (2000). An exploratory study of traditional and objective scoring systems with MGQT field cases. *Polygraph, 29*(2), 185-194.

Krapohl, D.J., and Ryan, A.H. (2001). A belated look at symptomatic questions. *Polygraph*, *30*(3), 206-212.

Krapohl, D.J., Shull, K.W., and Ryan, A.H. (2002, July). Does the confession criterion in case selection inflate polygraph accuracy estimates? *Forensic Science Communications*.

Krapohl, D.J., and Stern, B.A. (2003). Principles of multiple-issue polygraph screening a model for applicant, post-conviction offender, and counterintelligence testing. *Polygraph, 32*(4), 201-210.

Krapohl, D.J., and Trimarco, J.R. (2005). Credibility assessment methods for the new century. *National Academy Associate, 7*(1), 8-9, 24, 32.

Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & Trumbull (eds.), *Psychological Stress*. Appleton-Century-Crofts: New York

Landis, C., and Wiley, L.E. (1926). Changes of blood pressure and respiration during deception. *Journal of Comparative Psychology, 61*(1), 1 – 19.

Larson, J.A. (1923). The cardio-pneumo-psychogram in deception. *Journal of Experimental Psychology, 6*(6), 420 – 454.

Li, F., Zhu, H., Xu, J., Gao, Q., Guo, H., Wu, S., Li X. and He, S. (2018) Lie detection using fNIRS monitoring of inhibition-related brain regions discriminates infrequent but not frequent liars. *Frontiers in Human Neuroscience*. 12:71. doi: 10.3389/fnhum.2018.00071

Luria, A.R (1930). The method of recording movements in crime detection. *Zeitschrift Fuer Angewandte Psychologie, 35*, 139-183. (Text in German).

Lykken, D.T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43, 385-388. Reprinted in 1979 in *Polygraph*, *7*(2), 123-128.

Lykken, D.T. (1998). A tremor in the blood: Uses and abuses of the lie detector. Plenum Trade: New York.

Lynch, B.E., and Henry, D.R. (1979). A validity study of the psychological stress evaluator. *Canadian Journal of Behavioural Science*, *11*(1), 89-94.

MacLaren, V., and Taukulis, H. (2000). Forensic identification with event related potentials. *Polygraph*, *29*(4), 330-343.

MacLaren, V. (2001). A qualitative review of the Guilty Knowledge Test. *Journal of Applied Psychology*, *86*(4), 674-683.

Mangan, D.J., Armitage, T.E., and Adams, G.C. (2008). A field study on the validity of the Quadri-Track Zone Comparison Technique. *Physiology & Behavior, 95*, 17-23.

Marey, P.E.J. (1885) *Méthode Graphique Dans Les Sciences Expérimentales Et Principalement En Physiologie Et En Médecine*. G. Masson: Paris.

Marin, J. (2000). He said/She said: Polygraph evidence in court. *Polygraph 29*(4) 299-304.

Marin, J. (2001). The exclusionary standard and the "Litigation Certificate" program. *Polygraph*, *30*(4), 288-293.

Marston, W.M. (1917). Systolic blood pressure symptoms of deception. *Journal of Experimental Psychology*, *2*(2), 117-163. Reprinted in *Polygraph*, *14*(4), 289-320.

Marston, W.M. (1938). *The lie detector test*. Richard R. Smith: New York.

Matte, J.A. (1976). A polygraph control question validation procedure. *Polygraph, 5*(2), 170-177.

Matte, J.A. (1996). *Forensic psychophysiology using the polygraph: Scientific truth verification – Lie detection*. J.A.M Publications: Williamsville, NY.

Matte, J. A. (2000). *Examination and cross-examination of experts in forensic psychophysiology using the polygraph*. J.A.M. Publications: Williamsville, NY.

Matte, J.A. (2001). Comments on Krapohl & Ryan criticism of Capps, Knill & Evans research. *Polygraph*, *30*(3), 216-217.

Matte, J. A. (2002). *2002 supplement - Forensic psychophysiology using the polygraph*. J.A.M. Publications: Williamsville, NY.

Matte, J.A., and Grove, R.N. (2001). Psychological set: Its origin, theory and application. *Polygraph*, *30*(3), 196-202.

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

Matte, J.A., and Reuss, R.M. (1998). An analysis of the psychodynamics of the directed-lie control question in the control question technique. *Polygraph*, *27*(1), 56-67.

Matte, J.A., and Reuss, R.M. (1990). A field study of the "friendly polygraphist" concept. *Polygraph*, *19*(1), 1-8.

Meehl, P.E., and Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, and cutting scores. *Psychological Bulletin*, *52*(3), 194-216.

Meiron, E., Krapohl, D.J., and Ashkenazi, T. (2008). An assessment of the Backster "Either-Or" rule in polygraph scoring. *Polygraph, 37*(4), 240-249.

Miller, J.C. (1994). *Cardiovascular indices of guilty knowledge.* Department of Defense Polygraph Institute, Ft. McClellan, AL. DoDPI94-R-0016. DTIC AD Number A305954.

Minor, P. (1985). *The modified relevant/irrelevant (MRI) technique.* Paper presented at the 20th annual seminar of the American Polygraph Association, Reno, NV.

Mosso, A. (1896). *Fear*. Translated from the fifth edition of the Italian by E. Lough and F. Kiesow. Longsmans, Green and Co.: London

Murphy, K.R. (1987). Detecting infrequent deception. *Journal of Applied Psychology*, *72*(4), 611-614.

Nelson, R., Krapohl, D and Handler, M. (2008). Brute-force comparison: A Monte Carlo study of the Objective Scoring System version 3 (OSS-3) and human polygraph scorers. *Polygraph, 37*(3), 185-215.

Ohnishi, K., Matsuno, K., Arasuna, M, and Suzuki, A. (1976). The objective analysis of physiological indices in the field detection of deception. *Reports of the National Institute of Police Science, 29*, 181-188.

Orne, M.T., (1973). Implications of laboratory research for the detection of deception. *Polygraph, 2*(3), 169-199.

Osugi, A. (2011). Daily application of the Concealed Information Test: Japan. In (Verschuere, Ben- Shakhar & Meijer, Eds.) *Memory Detection: Theory and Application of the Concealed Information Test*. Cambridge University Press: New York.

O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy who cried wolf effect. *Personality & Social Psychology Bulletin, 29*(10), 1316-1327.

Patrick, C.J, and Iacono, W.G. (1989). Psychopathy, threat and polygraph test accuracy. *Journal of Applied Psychology*, *74*(2), 347-355

Patrick, C.J, and Iacono, W.G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, *76*(2), 229-238.

Pavlidis, I., Eberhardt, N.L., and Levine, J.A. (2002). Seeing through the face of deception. *Nature, 415*, 3.

Petty, R., and Cacioppo, J. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. William C. Brown: Dubuque, IA.

Podlesny, J.A., (1993). Is the guilty knowledge polygraph technique applicable in criminal investigations?: A review of FBI case records. *Crime Laboratory Digest*, *20*(3), 57-61.

Podlesny, J.A., and Raskin, D.C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology*, 15, 344-358.

Podlesny, J.A., and Truslow, C.M. (1993). Validity of an expanded-issue (modified general question) polygraph technique in a simulated distributed-crime-roles context. *Journal of Applied Psychology, 78*(5), 788-797. Reprinted in *Polygraph*, *23*(3), 195-218.

Pollina, D.A., Horvath, F., Denver, J.W., Dollins, A.B., and Brown, T.E. (2008). Development of technologies and test formats for credibility assessment. In (Alexandra M. Columbus, Ed.) *Advances in Psychology Research*, *58*, 1-36.

Pollina, D., and Ryan, A.H. (2003). *The relationship between facial skin surface temperature reactivity and traditional polygraph measures used in the psychophysiological detection of deception: A preliminary investigation.* DoD Polygraph Institute, Ft. Jackson, SC. DoDPI02-R-0007. DTIC AD Number: ADA414911.

Pollina, D. (2006). Emerging methods and measures for detecting stress and deception: Thermal imaging. *Journal of Credibility Assessment and Witness Psychology, 7*(2), 108-115.

Raskin, D.C. (1976). *Reliability of chart interpretation and sources of error in polygraph examinations*. Report No. 76-3, National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice (Contract No. 75-NI-99-0001). Department of Psychology, University of Utah

Raskin, D.C., and Hare, R.D., (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology*, 15, 126-136.

Raskin, D.C., and Honts, C.R. (2002). The comparison question test. In: M. Kleiner's (ed.) *Handbook of polygraph testing*, pp. 1-47. Academic Press: London.

Raskin, D.C., Kircher, J.C., Honts, C.R., and 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.

Reali, S.F. (1978). Reali's positive control technique: A new concept of polygraph procedures. *Polygraph*, 7(4), 281-285.

Reid, J.E., and Inbau, F.E. (1977). *Truth and deception: The polygraph ("lie detector") technique* (2nd Ed.). Williams & Wilkins: Baltimore, MD.

Reed, S. (1994). A new psychophysiological detection of deception examination for security screening. *Psychophysiology*, 31(Supplement 1), S80, (Abstract).

Rosenfeld, J.P. (1998). Event-related potentials in detection of deception. *International Journal of Psychophysiology*, 30(1), 27. (Abstract).

Rothwell, J., Bandar, Z., O'Shea, J.D., and McLean, D. (2006). Silent talker: A new computer-based system for the analysis of facial cues to deception. *Applied Cognitive Psychology, 20(6). 757 – 777.*

Ruch, F.L. (1948). *Psychology and Life*. Scott Foresman: Chicago.

Runkel, J.E. (1936). Luria's motor method and word association in the study of deception. *Journal of General Psychology*, *15*, 23-37.

Senter, S. M., and Dollins, A. B. (2002). *New decision rule development: Exploration of a two-stage approach*. (DoDPI01-R-0006). Fort Jackson, SC: Department of Defense Polygraph Institute.

Senter, S.M., Waller, J., and Krapohl, D.J. (2008). Air Force Modified General Question Test validation study. *Polygraph, 37*(3), 174-184.

Senter, S.M., Waller, J., and Krapohl, D.J. (2009). Validation studies for the Preliminary Credibility Assessment Screening System (PCASS). *Polygraph, 38*(2). 115-129.

Senter. S.M., Weatherman, D., Krapohl, D.J., and Horvath, F.S. (2010). Psychological set or differential salience: A proposal for reconciling theory and terminology in polygraph testing. *Polygraph, 39*(2). 109-117.

Shurany, T., Stein, E., and Brand, E. (2009). A field study on the validity of the Quadri-Track Zone Comparison Technique. *European Polygraph, 1*(7), 5-23.

Stern, B.A., and Krapohl, D.J. (2003). The infamous James Alphonso Frye. *Polygraph, 32*(3), 188-199.

Summers, W.G. (1939). Science can get the confession. *Fordham Law Review*, *8*, 334-354.

Swets, J.A. (1995). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Lawrence Erlbaum Associates.

Swets, J.A., Dawes, R.M, and Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*(1), 1-26).

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.

Thorndike, E.L. (1920). A constant error on psychological rating. *Journal of Applied Psychology, 4*, 25-29

Tian, F., Sharma, V., Kozel, F.A., and Liu, H. (2009). Functional near-infrared spectroscopy to investigate hemodynamic responses to deception in the prefrontal cortex. *Brain Research*, 120 – 130. Doi:10.1016/j.brainres. 2009.09.085

Timm, H.W. (1982). Analyzing deception from respiration patterns. *Journal of Police Science and Administration*, *10*(1), 47-51.

Timm, H.W. (1991). Effect of posthypnotic suggestions on the accuracy of preemployment polygraph testing. *Journal of Forensic Sciences*, *36*(5), 1521-1535.

Trovillo, P.V. (1939). A history of lie detection. *Journal of Criminal Law and Criminology*, *29*(6), 848-881 and *30*(1), 104-119. Reprinted in *Polygraph, 1*, 46-74 and 151-160.

Van Herk, M. (1990). Numerical evaluation: Seven point scale +/-6 and possible alternatives: A discussion. *The Newsletter of the Canadian Association of Police Polygraphists*, *7*(3), 28-47.

Vendemia, J.M.C. (2002). Hobson's choice: The relationship between consequences and the comparison question. *Polygraph*, *31*(1), 20-25.

Veraguth, S. (1906). Das psycho-galvanische Reflexphänomen. *Monatsschrift für Psychiatrie und Neurologie*, Bd. XXI, Heft 5.

Verschuere, B., Ben-Shakhar, G., and Meijer, E. (2011). *Memory detection: Theory and application of the Concealed Information Test*. Cambridge University Press: New York.

Waller, J.F. (2001). A concise history of the comparison question. *Polygraph*, *30*(3), 92-195.

Weaver, R.S. (1985). Effects of differing numerical chart evaluation systems on polygraph examination results. *Polygraph*, *14*(1), 34-42.

Webb, A. K, Honts, C. R., Kircher, J. C., Bernhardt, P., and Cook, A. E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. *Legal and Criminological Psychology, 14*, 279-292.

Webb, A.K., Hacker, D.J., Osher, D., Cook, A.E., Woltz, D.J., Kristjansson, S., and Kircher, J.C. (2009). Eye movements and pupil size reveal deception in computer administered questionnaires. *Foundations of Augmented Cognition, Neuroergonomics and Operational Neuroscience: Lecture Notes in Computer Science, 5638*, 553-562.

Weir, R.J. (1976). Some principles of question selection and sequencing for Relevant-Irrelevant testing. *Polygraph, 5*(3), 207 – 222.

Winter, J. (1936). Comparison of the cardio-pneumo-psychograph and association methods in the detection of lying in cases of theft among college students. *Journal of Applied Psychology*, *20*(2), 243-248.

Yankee, W.J. (1992). A case for forensic psychophysiology and other changes in terminology. Paper presented to Advisory Committee, Director of Counterintelligence; and to the DASD(CI&SCM). Reprinted in *Polygraph, 23*(3), 188-194.

Yankee, W.J. (1995). The current status of research in forensic psycho-physiology and its application in the psychophysiological detection of deception. *Journal of Forensic Sciences*, *40*(1). 63-68.


AMERICAN POLYGRAPH ASSOCIATION
DEDICATED
TO
TRUTH