Polygraph

VOLUME 37 2008 NUMBER 1

Special Edition – Norman S. Ansley

Selected Articles from his Work

Contents

Published Quarterly

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

A Tribute to Norman S. Ansley

Donnie W. Dutton
President
American Polygraph Association

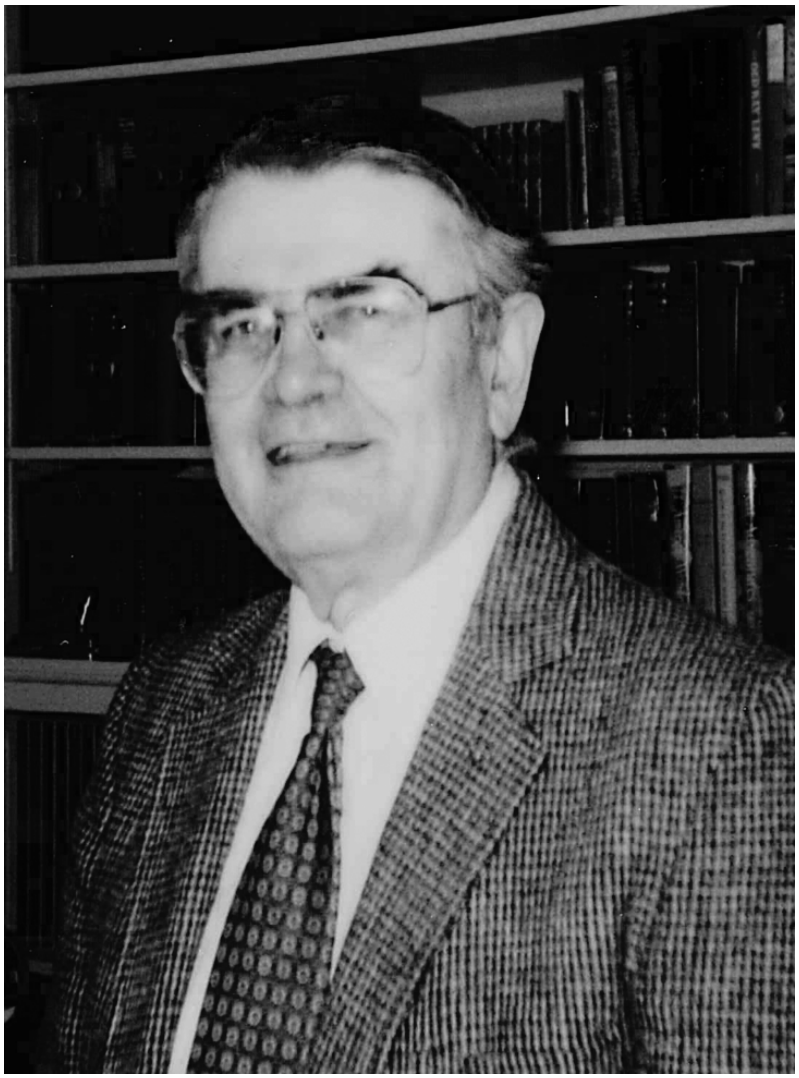
This issue of *Polygraph* is devoted to the published works of Mr. Norm Ansley. By necessity it can only cover a small fraction of the polygraph-related publications Norm produced in over 50 years of writing. His articles have spanned the full spectrum of polygraphy, from law to history, physiology to techniques, science to data analysis to legislation. There is scarcely an aspect of our field that has been left untouched by Norm's pen. Despite age and health challenges, Norm keeps current with the field, and continues to read the APA publications he created.

When I was a student going through the then-United States Army Military Police School (USAMPS) polygraph program I was fortunate to have been taught the Relevant/Irrelevant Technique by Norm. At that time I was like a sponge trying to soak in everything that there was to our profession. As a student I didn't know how much Norm had already contributed to the field of polygraph with the research that had been conducted, nor did I appreciate the importance of what his research has now led us to today. As the years went by I became aware of the multitude of contributions he made and eventually, I too got to teach at the USAMPS which later became known as the Department of Defense Polygraph Institute (now the Defense Academy for Credibility Assessment, or DACA). While instructing there I was assigned the task of putting a history lesson together. Wanting to impress Mr. Ron Decker, the school Director at the time, I did a lot of research and developed an extremely challenging block of instruction. As luck would have it, Norm was at the school during the unveiling of my new and improved history block and I asked him for his feedback on my product. Norm took the examination and upon completion missed one of the answers. I just knew I had stumped a master, and was preparing for my proud moment. He expressed appreciation for the test and then informed me that my research was wrong in one particular area. As the discussions continued I showed him where I had obtained my data. With the care of a great teacher he politely pointed out that he had been the one who had conducted the initial research I was quoting, and went on to correct the error in my test. I had been granted a teaching moment from one of the masters, one that will long be with me. Norm never showed pride or impatience, but simply offered to send me a copy of his research so that we could place it into the school library for future reference. He was good to his word, and the DACA library holds a copy of that research to this day.

Norm not only has conducted research on the various areas of polygraph but was also instrumental in the growth of the APA. Norm was the first Editor of the *APA Newsletter* (predecessor to the *APA Magazine*) and the APA's journal *Polygraph* for over 25 years. His work allowed examiners to conduct polygraph examinations which would produce a higher degree of validity and reliability, he challenged scientist to disprove his polygraph research findings, and provided documents used in Congressional hearings on the polygraph. Norm was awarded the Lifetime

Achievement Award in June of this year for his contributions to the APA and polygraph profession.

Norm's research projects have been the catalyst for other scientists to build upon over the years and without his pioneering efforts, I dare to say we would not be nearly as far in the discovery of truth. I hope you enjoy reading this issue of *Polygraph*, and as you read I hope you too will come to appreciate what Mr. Norm Ansley has done to shape the profession. And so, to this great polygraph writer, researcher, chronicler, editor, teacher, leader, colleague, and inspiration we dedicate this publication.



Research on the Validity of the Relevant-Irrelevant Technique as Used in Screening

Norman Ansley

The use of the Relevant-Irrelevant (RI) polygraph technique in employment screening goes back to 1931 when Leonarde Keeler began a program of testing employees of banks in Chicago on a systematic basis to detect and prevent embezzlement and theft (Keeler, 1931). Federal use of the RI technique in screening began in World War II with a program to protect the atom bomb project (Trovillo, 1951) and another program involving the screening of German prisoners of ware for post-war police assignments in Germany (Linehan, 1978). After the war, the Government developed polygraph screening programs for the protection of intelligence operations and agencies, programs that continue to exist (Hearings, 1974). The screening of Chicago bank personnel also continues, and commercial testing has expanded dramatically into other fields. Although other techniques may be used, much of the commercial employment screening is conducted with the RI technique, and almost all of the Federal security screening is done with RI techniques.

Criticism

Criticism of RI technique has been confounded by inaccurate descriptions. For example, in their textbook on Reid Control Question Technique, Reid and Inbau give the following description of RI technique:

> It contains some questions pertaining to the issue under investigation (relevant questions) and other questions that are irrelevant but chosen because the answers are known to be truthful . . . The responses to the relevant and irrelevant questions are then compared and if the subject responds more to the relevant (issue) questions than to the irrelevant (known truthful) questions, the subject is considered as not telling the truth, but if there are no significant responses to either the relevant or irrelevant questions, the subject is reported as truthful. (Reid, 1977)

This description is significantly incomplete and inaccurate; and ignores the detailed descriptions of RI by the authorities on two current major forms (Weir, 1974; Harrelson, 1973). In fact, RI examiners may use control questions and control methods, question repetition, guilt complex questions, and other types of questions and procedures which were not mentioned above. The book by Leonard Harrelson, *The Keeler Technique* and the lengthy article by Raymond J. Weir, Jr., "In Defense of the Relevant-Irrelevant Polygraph Test" are readily available; the former from the Keeler Institute in Chicago and the latter from the APA Reference Service. Another proponent of control question technique, David C. Raskin, gave the following description of RI technique during his presentation in opposition to employment screening at a symposium on the polygraph at the 1979 meeting of the American Psychological Association:

> In its simplest form, the relevant-irrelevant technique includes questions about the crime (relevant) and questions totally unrelated to the crime (irrelevant) . . . The simple-minded theory of that test is that an individual who is deceptive [about] his involvement in the crime feels very threatened by the relevant questions and shows larger autonomic physiological reactions to those questions. However, the innocent subject does not have those strong concerns, and therefore shows no greater reactions to the relevant as compared to the irrelevant questions (Raskin, 1980).

This article was originally published in *Polygraph*, 1980, Volume 9(3)

While Raskin took care to say this was the simplest form, and put it in the setting of the 1920's and 1930's under John A. Larson and Leonarde Keeler, the casual reader might be misled to think that this is an adequate description of current RI technique. It is, of course, utterly inadequate. In fact, it is not even a fair description of how many of the RI tests were conducted in the 1930's (Cf. Keeler, 1938). Keeler was using recognition factors in RI relevant questions in which the significance of certain details would be known only to the perpetrator (Keeler, 1931) and Lee was teaching and practicing the use of guilt complex questions; threatening questions which appear relevant to the subject, which are answered truthfully. Then, as now, it was common practice to combine RI and peak of tension tests in a single interview (Lee, 1943). It is not at all unusual for that combination to be used now in complex screening cases.

Some of the critics of RI technique are adherents of control question techniques who have been trained in only one method and believe that theirs is the only way to conduct a test. Although Frank Horvath was initially trained in a control technique, he has had the opportunity to study the whole polygraph profession. In an article on the state of the art, Horvath commented particularly on parochial training in which the student acquires knowledge of only one technique:

> Many examiners tend to believe that control-question testing is the ultimate approach – the most sophisticated, advanced, and useful procedure that exists or ever will exist. These people have been trained in a way not unlike the way you train fleas. If you put fleas in a jar and cover it, the fleas will jump so high, but no higher – the will be unable to escape. Examiners who believe that control question testing or one variation of it is the ultimate technique or that it is always the only way to conduct a polygraph examination, are trained fleas (Horvath, 1980).

One of the more serious allegations of critics is that the RI screening technique produces an unacceptably large number of false positive errors, truthful statements erroneously called lies. The statistical probability arguments raised against screening (Lykken, 1974; Raskin, 1980) make the assumption that false positives and false negatives are nearly equally distributed among errors made by examiners. Research suggests that assumption may be false, rendering the analyses specious so far as RI is concerned. None of the research shows a false positive rate as high as the false negative rate. Research at the University of Georgia (Correa, 1979) did not produce any false positive errors in their simulation of employment screening. There were no false positive errors among those employees examined by Professor MacNitt (1941). Similarly, Blum and Osterloh had a false positive rate of less than one-tenth of one per cent, which represented only one false positive in the four errors he made in testing the truthfulness of 2,120 subsets of information (Blum, 1968).

In the laboratory setting, it is customary to make decisions in every case, something which does not happen in the world of daily practice. In the laboratory, when a decision is made where the evidence is not clear, the error rate may be increased. In the real world, the examiner often exercises his option to conduct additional examinations at a later date. Even with a reexamination, there are cases that must be reported as inconclusive. The problem of risk versus accuracy in screening is addressed in an interesting study in Israel in which different cutoff points in scoring the charts were established to match the purposes of the examination (Ben Shakhar, 1969).

Research

Research on the RI polygraph technique as used in employment screening has demonstrated that the technique has high validity (Correa, 1980; MacNitt, 1942). One study which compared RI with the validity of the highly regarded peak of tension technique, showed that the RI method was superior to the peak in detecting both prior activity and specific knowledge (Gustafson, 1964). The RI technique is also highly valid in detecting deception where stories may be partly true and partly false, all true, or all false, a situation common to screening (Blum, 1964). That research, using RI technique and police informants as subjects, indicated correct decisions in 99.9% of the decisions. A report

on the use of RI technique in specific criminal cases, and using only the electrodermal parameter, indicated an accuracy rate of 94.1% (Guertin, 1954). In a criminal case involving 81 suspects, using the most basic RI methodology, supplemented with peak of tension tests, produced correct results in finding all of them innocent (Bitterman, 1946). It turned out that the thief was not among the suspects examined by Bitterman.

Critics have suggested that screening may be ineffective in those cases where the motivation is quite low, particularly when the lying is cognitive rather than emotional. However, in a study where motivation was low and stimulus intensity was strictly controlled, a laboratory screening test to detect false biographical entries on application forms indicated that significant electrodermal responses occurred to the pure act of deception (Hemsley, Heselgrave & Furedy, 1979). Moreover, practical experience suggests that the motivation of most applicants is quite high.

In regard to the reliability of chart interpretation, there is evidence that RI charts from screening examinations can be read blind with a high degree of reliability. In a report on a government research project involving charts from real screening cases, there were agreement rates between the examiner and the blind rater on reaction or no reaction on each of three channels (7,590 decisions) or 96% for cardiovascular, 95% for electrodermal, and 96% for respiratory responses (Edel & Jacoby, 1975).

Abstracts of RI Research

The purpose of this paper is to summarize the research which has been done in connection with the RI technique, particularly as it relates to employment screening. The following is a series of abstracts of each of the published research papers on the topic of the relevant-irrelevant technique that have some bearing on the use of the technique in screening.

Bersh, Philip J. "A Validation Study of Polygraph Examiner Judgments," A report prepared for the Department of Defense Joint Services Group on Lie Detection Research, April 1968, United States Army Behavioral Science Research Laboratory, Washington, DC. Subsequently reprinted in *Journal of Applied Psychology*, 1969, 53, 399-403.

The lie detection judgments of polygraph examiners in criminal investigations conducted by the military services were validated against unanimous guilt-innocence decisions by a panel of four Judge Advocate General (JAG) attorneys. The panel of lawyers had access to the complete investigative file, except that all references to the polygraph examination were removed. Cases involving confession were ruled out because some polygraph examiners could have made their judgment of deception after the subject had confessed. Cases were selected at random from those conducted from 1963 to 1966. There were about an equal number of control question and relevant-irrelevant tests, and within each type, an equal number of judgments of deception indicated and no deception indicated. No cases where the examiner reported inconclusive were employed.

A total of 323 case files were submitted to the panel. Each member of the panel was initially required to judge whether a file contained sufficient evidence to warrant a decision of guilt or innocence. Files with inadequate evidence were eliminated from further consideration, and 157 remained. Then each attorney made an independent judgment of the guilt or innocence of a suspect. The attorneys were given explicit instructions to disregard all legal technicalities and to judge each case solely on the evidence contained in the file.

The polygraph examiners and the JAG panel agreed on 92.4 percent of all cases. In seven cases the examiner reported "deception indicated" and the panel decided "not guilty" (4.5%). In five cases the panel decided "guilty" and the examiner reported "no deception indicated" (3.2%). There are no separate figures for RI and control question techniques. Because these were real cases, there is no way in which to decide who was correct, the panel or the examiner, when they disagreed. It is also possible that when the panel and examiner agreed, they were both wrong. However, the results do show the agreement between a panel of attorneys and the judgment of a polygraph examiner is significantly above chance.

Bitterman, M.E. and Marcuse, F.L. "Cardiovascular Responses of Innocent Persons to Criminal Interrogation," *Minor Studies from the Psychology Laboratory of Cornell University*, 1946, 407-12.

A theft of $100 took place in one of the rooms of a campus dormitory. The psychologists, who had no polygraph training but did have some books and articles on the topic, conducted polygraph examinations on the 81 men who lived in the dormitory. They were unable to devise a method for evaluating the pneumograph pattern, but did devise a systematic method to evaluate the cardiosphygmograph pattern. They prepared a seven question test which opened with two irrelevant questions, then a relevant question, an irrelevant question, and three relevant questions. The last question was a general one, asking the subject if he had answered all of the questions truthfully.

In their first series of examinations, they cleared of guilt all but seven of the subjects (8.6%). Those seven, and twenty other students, were also given peak of tension tests after the relevant-irrelevant tests. The authors thereupon concluded that none of the students was guilty of the theft. Obviously, there were no false positive errors after the peak of tension tests. The question of false negatives, calling the thief or thieves not deceptive, was solved some time later when investigation disclosed that the theft was committed by someone else, who was not among those tested. Accordingly, the results were correct in 100% of the decisions.

Blum, Richard H. and Osterloh, William. "The Polygraph Examination as a Means for Detecting Truth and Falsehood in Stories Presented by Police Informants," *Journal of Criminal Law, Criminology, and Polygraph Science*, 1968, 59, 133-137.

A total of 17 male and 3 female informants who had given information on criminal cases to local or federal agencies were selected and paid to engage in the experiment. Great care was taken to protect their identities and the fact that they were cooperating with the police, as their lives would otherwise be in danger. Some of the stories that the informants were to tell the polygraph examiner were true, some were false, and some were partly true and partly false. The true stories were those which had been provided to their handler on a previous occasion which had stood the scrutiny of investigation. A false story was one jointly invented by the officer and the informant, but compatible with the informant's ordinary role and opportunities and containing credible information. True stories with false information followed the same rules as the true story, and the false items were important to the investigation, such as the name of the offender, the place where goods were hidden or fenced, etc. The story was first written, then rehearsed with the police handler. In all, there were 106 statements to be checked for veracity in each of the 20 cases, for a total of 2,120 statements. The examiner used the RI technique.

Of nine stories that were all true, the examiner was correct in his diagnosis of all of them. None of these true stories were designated as false or partly false.

Of the eleven stories that were either partly or entirely incorrect, the examiner was correct in saying that each of these was either partly or wholly false. None of these were diagnosed as completely truthful.

The examiner made four errors among the seven subjects who told stories that were only partly true. In two cases, the examiner made an error in judging a lie as the truth in one of the five elements of a partly true story. In the other case, the examiner made two errors among the five subsets, calling one truthful element deceptive and one deceptive element truthful. Of the 106 statements made by twenty subjects (2,120 statements), the examiner was correct in all but four (99.8%). The error rate was slightly under .2% in terms of verifying information. If one considers any error in a subset as an erroneous case, then the examiner was incorrect in three of twenty cases, correct in 17 (85%).

Correa, E.I. The validity of the pre-employment polygraph examination and the effects of motivation. Unpublished dissertation, University of Georgia, 1979.

In a research project at the University of Georgia, forty subjects took pre-employment examinations, conducted with

the RI technique. Half were to tell the truth, and half were to be deceptive to three of the nine questions. Polygraph recordings included respiration from a thermister probe at the nostril, EKG for heart rate, and electrodermal activitity. A cardiosphygmograph was not employed.

The experimenter, in separating the truthful and untruthful subjects, was correct in all cases, for 100%. Identification of lies, by subject, ranged from 68% to 100%. There were no cases in which a truthful person was called deceptive (false positives). All of the errors were in failing to identify a lie.

Lying responses were characterized by significantly larger increases in conductance in the electrodermal activity and larger decreases in heart rate, than the responses to telling the truth. Respiration showed no significance in these recordings.

An attempt to separate subjects by motivation, offering half of the group $25.00 if they could deceive the examiners, showed no significant difference.

Edel, Eugene C. and Jacoby, Jacob. "Examiner Reliability in Polygraph Chart Analysis: Identification of Physiological Responses." *Journal of Applied Psychology*, 1975, 60, 632-634.

Actual case charts from 40 polygraph screening cases, involving responses to 2,530 questions were examined in detail by ten experienced examiners, working separately. The technique was relevant-irrelevant.

Because each examiner made judgments of reaction or no reaction to each question, of which there were 2,530, on each of three channels, he made 7,590 decisions. The channels were cardiovascular, electrodermal, and respiratory. The agreement between the original examiner and a blind rater was 96% for cardiovascular, 95% for electrodermal, and 96% for respiratory responses. In terms of total agreement between all examiners who read all of the charts, the agreement was 96% for cardiovascular, 91% for electrodermal, and 96% for respiratory responses. The overall percentage of agreement for rater versus rater was 94%.

Of the total 22,770 judgments there were 21,626 agreements for an overall agreement rate of 95%, including the examiner-rater and rater-rater combinations.

Guertin, Wilson H. and Wilhelm, Paul L. "A Statistical Analysis of the Electrodermal Response Employed in Lie Detection." *Journal of General Psychology*, 1954, 51, 153-160.

The author selected 34 criminal cases in which guilt or innocence was established by independent means. There were 19 cases of guilt and 15 cases of innocence among the 31 men and three women.

The records from an electrodermal unit, with which the examiner used a relevant-irrelevant technique, were subsequently evaluated blindly. The resulting analysis was correct in 32 of the 34 cases (94.1%).

Gustafson, Lawrence A. and Orne, Martin T. "The Effects of Task and Method of Stimulus Presentation on the Detection of Deception." *Journal of Applied Psychology*, 1964, 48, 383-387.

The study was designed to investigate the relative effectiveness of two different polygraph techniques, relevant-irrelevant and peak of tension. Subjects were given two RI and two POT tests. 29 were guilty, and 24 subjects had guilty knowledge. Only an electrodermal measure was scored for this research.

The RI method proved more effective in detection than the POT method when subjects were trying to deceive as to which item of information they possessed, in the guilty information paradigm. There was no significant difference in the effectiveness of the RI method of stimulus presentation between the guilty person paradigm and the guilty information paradigm. The POT method proved significantly less effective than the RI method in the guilty information paradigm, and significantly less effective in that paradigm than it was in the guilty person paradigm. In general, subjects found it easier to deceive in the guilty information paradigm where they could attempt to "appear guilty" on a non-critical item, especially in the peak of

tension because they could anticipate the order of presentation of the items.

Hemsley, Gordon; Heslegrave, Ronald J. and Furedy, John H. "Can Deception be Detected When Stimulus Familiarity is Controlled?" Paper presented at the Annual Seminar of the Society for Psychophysiological Research, Cincinnati, Ohio, October 18, 1979.

Using a pre-employment test paradigm with the relevant-irrelevant technique, ten male and ten female subjects were divided into two groups, one which was to be deceptive to some of the 20 items of biographical information supplied on application forms, and one group which was to be truthful about all of the biographical information on the forms. Because the stimulus familiarity was controlled, with all biographical information known to the subjects in advance, the issue was whether or not larger ANS responses would occur to deception alone. Skin conductance response was used to measure ANS response.

The skin conductance response was significantly, (F=64.1) greater for deceptive (S = 2.28 umhos) than for honest (X = 1.40 umohs) responses.

Neither an habituation effect nor a sex difference emerged. The experimenters concluded that the design allows the skin conductance response to detect, in the laboratory, "pure deception," because that emotional arousal associated with question content, and enhanced stimulus familiarity or signal value associated with the relevant alternative, had been eliminated as confounding sources.

MacNitt, Reginald D. "In Defense of the Electrodermal Response and Cardiac Amplitude as Measures of Deception." *Journal of Criminal Law, Criminology and Police Science*, 1942, 33, 266-275.

Professor MacNitt conducted 59 cases using the relevant-irrelevant technique that involved employees made available for the Columbus, Ohio Merchants Audit Bureau. Prior investigations results were withheld from him. There were employees whose honesty and integrity were above reproach, employees who had confessed to stealing goods and money, but were told to lie, and employees who were suspected of stealing and about whom there was quite a bit of evidence already in the possession of authorities. The latter group was expected to lie.

MacNitt said that his results were "correctly reported . . . and checked by confessions or the verifiable records of the employees, the examiner failing only on a few minor details. In all of these 59 cases, the electrodermal response was the more accurate of the two employed." The other channel recorded cardiac rate and amplitude.

References

Ben Shakhar, G., Lieblich, I., and Kugelmass, S. Guilty knowledge technique: Application of signal detection measures. *Journal of Applied Psychology*, 1970, 54, 409-413.

Bersh, P.J. A Validation study of polygraph examiner judgments. *Journal of Applied Psychology,* 1969, 53, 399-403.

Bitterman, M.E. and Marcuse, F.L. Cardiovascular responses of innocent persons to criminal interrogation. *Minor Studies from the Psychology Laboratory of Cornell University*, 1946, 407-412.

Blum, R.H. and Osterhoh, W. The polygraph examination as a means for detecting truth and falsehood in stories presented by police informants. *Journal of Criminal Law, Criminology and Police Science*, 1968, 59, 133-137.

Correa, E.I. The validity of the pre-employment polygraph examination and the effects of motivation. Unpublished dissertation, University of Georgia, 1979.

Edel, E.C. and Jacoby, J. Examiner reliability in polygraph chart analysis: Identification of physiological responses. *Journal of Applied Psychology*, 1975, 60, 632-634.

Guertin, W.H. & Wilhelm, P.L. A statistical analysis of the electrodermal response employed in lie detection. *Journal of General Psychology*, 1954, 51, 153-160.

Gustafson, L.A. & Orne, M.T. The effects of task and method of stimulus presentation on the detection of deception. Journal of Applied Psychology, 1964, 48, 383-387.

Harrelson, L.H. *The Keeler Technique (Training Guide)*, Chicago: Keeler Polygraph Institute, 1973.

Hearings, "The Use of Polygraphs and Similar Devices by Federal Agencies," Subcommittee on the Committee on Government Operations, House of Representatives, 93rd Congress, 2nd Session, June 4 and 5, 1974, pp.421-433, 643-663.

Hemsley, G., Heselgrave, R.J. & Furedy, J.J. Can deceptive be detected when stimulus familiarity is controlled? Paper presented at the Annual Seminar of the Society for Psychophysiological Research, Cincinnati, Ohio, October 18, 1979.

Horvath, F. Polygraphy: Some comments on the state of the art. *Polygraph*, 1980, 9, 34-41.

Keeler, L. Lie detector applications. In *Proceedings of the International Association of Chiefs of Police*, Washington, DC: International Association of Chiefs of Police, 1931, p.184.

Keeler, L. The detection of deception. In Keeler, L., Muelberger, C.W., Wilson, C.M., Inbau, F.E., Keeler, K., and O'Neill, M.E. *Outline of Scientific Criminal Investigation.* Chicago: Scientific Crime Detection Laboratory, 1938, pp. 42-50.

Lee, C.D. *Instruction Manual for the Berkeley Psychograph.* Sacramento: Lee & Sons, 1943.

Lykken, D.T. Psychology and the lie detector industry. *American Psychologist*, 1974, 29, 225-239.

MacNitt, R.D. In defense of the electrodermal response and cardiac amplitude as measure of deception. *Journal of Criminal Law, Criminology, and Police Science*, 1942, 33, 266-275.

Reid, J.E. and Inbau, F.E. *Truth and Deception*, 2nd ed. Baltimore: Williams & Wilkins, 1977, p.4.

Raskin, D.C. Comparison of polygraph examinations in criminal justice and personnel selection. *Polygraph*, 1980, 9, 3-16.

Trovillo, P.V. Report on a survey of methods of operation, and accomplishments of the Russell Chatham, Inc. Polygraph Program at Oak Ridge, Tennessee. Unpublished report dated April 14, 1951. 43 pp.

Weir, R.J. Jr. In defense of the relevant-irrelevant polygraph test. *Polygraph*, 1974, 3, 119-166.

Testing the Physically Handicapped

Norman Ansley

Polygraphic testing of persons afflicted by certain handicaps often presents difficulties with respect to modifying procedural matters and equipment. For this reason, examiners tend to avoid testing in such situations, often needlessly so. It is the purpose of this paper to discuss procedures which can be used successfully to overcome difficulties encountered when testing persons who stutter, are deaf and dumb, or are blind.

Stutterers

Stuttering is a relatively common affliction, over a million Americans, probably closer to a million and a half, stutter somewhat. It appears to be essentially psychological in nature and may, in fact, be a cultural phenomenon, since it is known that there are cultures in which stuttering is nonexistent. But, what is stuttering? Stuttering can be defined as a defect in speech which is characterized by a stumbling and spasmodic repetition of syllables resulting from the difficulty in pronouncing initial consonants, presumably caused by spasms in the lingual and palatal muscles. Another definition of stuttering is that it is a disturbance of the smooth flow of speech due to tonic and clonic spasms involving the functions of respiration, phonation, and articulation. The tics and spasms may or may not be near to the speech mechanism. A tonic spasm is a persistent, involuntary, and even a violent muscular contraction. A clonic spasm is one that is marked by the muscular contraction immediately followed by relaxation. Although the two spasms differ, the result is the same, stuttering.

It is said that the stutterer stutters because, first of all, he expects to stutter. And second, he dreads it. He then becomes tense anticipating the stutter and he tries to avoid it, so he stutters. The stutter, then, stutters because he tries not to.

It is important to note that stammering differs from stuttering in that stammering is purely a performance phenomenon, not an emotional one. Stammering is only a defect in articulation; it is hesitant, faltering speech, characterized by mispronunciations and transpositions of words and letters. The subject who stammers can be helped some by avoiding the words that he stammers. This does not always hold true for a subject who stutters.

There is, by the way, no deformity of organs or disease of the nervous system related to stuttering. There is no difference in the general blood pressure, heart rate or metabolism. However, among children there is often a slower development of motor skills. Psychiatrists report that stutterers have an outstanding degree of primary anxiety. In other words, this anxiety is not a result of stuttering. Such primary anxiety is complicated by the secondary anxiety of stuttering.

Many stutters are said to have neurotic traits, being timid and obsessive–compulsive in nature and having repressed hostility. The presence of such traits may be a factor in the case; and may create problems during the testing.

Women may take comfort in the fact that some four to eight times as many men stutter as women; and among the adult population, it may be eight to one men or more. Because people stutter more when they are frightened, a disproportionate number may be encountered during polygraph testing.

Stuttering may involve three types, situational, ritual and traumatic. Situational stuttering occurs when the stutterer is influenced by the presence of certain persons or situations. For example, a child may stutter only in the presence of his parents, or only in school but never at home. Another example is

This article was originally published in *Polygraph*, 1976, 5(1)

the man who may stutter only in front of his boss, or only when he gives a speech.

Ritual stuttering refers not really to stuttering but rather to the manner in which it is controlled; that it, a ritual must be performed to prevent stuttering. This is very common among adults. To control or overcome their stuttering, they find some comfortable physical position or something they do mechanically that will prevent them from stuttering. For instance, they may hold their ear, put their hand on their head, or put a hand in a pocket. If anything is done to prevent them from going through this ceremony, they will stutter. When this kind of stutterer undergoes a polygraph examination, the placing of attachments may make it impossible for him to go through his ritual; hence, he may stutter during the testing, but not during the pre-test.

The third kind of stuttering, one many examiners are probably familiar with, is called traumatic. Such stuttering occurs immediately after a traumatic experience, such as a serious auto accident, and may last for a moment, a few moments, for days, or, rarely, for a lifetime. The V.A. hospitals are still treating a few people who are stuttering as a result of their experiences in the Korean War. Generally, however, traumatic stuttering lends itself to treatment better than other types of stuttering.

Traumatic stuttering may occur because of the immediate tension of the polygraph testing situation. You may come across a subject who has not stuttered before but who does stutter when in the examination room. Joyce Hanson, for instance, tells of one case where the subject stuttered only during the test, and then only when answering the relevant questions. In this case, the stuttering was a reaction, a meaningful deception reaction, later verified by admissions.

The average stutterer does not stutter all the time. He only stutters about 10 percent of his words. Also, some people will not stutter in the morning but only in the afternoon when fatigue sets in. Moreover, some people may botch up one word for a while and then, when that word can be pronounced clearly, have trouble with others, thus, their stuttering is not necessarily consistent.

There are two interesting things about most stutterers of all classes, except the traumatic. They can whisper without stuttering, and they can sing without stuttering. You may not get them to sing a response, but consider the other approach if you have a quiet room.

Many stutters can answer "yes" and "no" without stuttering. Hence, even if a stutterer is difficult to interview it is worthwhile to attempt routine polygraph testing. However, be aware of the occasional situation where a person is malingering; he is hoping that you will not test him because you will think he is unfit for testing. Also be aware that some people stop stuttering under tension. So a man may say, "You know, I usually stutter, but today I don't seem to have any problem." If observations indicate that a person is a stutterer make note of his ability to answer "yes" or "no". If such responses are normal, proceed with polygraph testing. But when "yes" or "no" responses cause problems, consider one of these alternate techniques.

I am indebted to Walter Atwood for the following example. While testing a stutterer, Atwood observed that the first chart was a mess; reactions in the pneumo, cardio, GSR were not systematic. They were simply irregular, and could not be analyzed. Noting this irregularity, Atwood switched to a form of a "no-answer" test, similar to the silent answer test developed by Horvath and Reid (1972). This test is used in conjunction with the Reid Control Question Technique, which requires answers on the first chart and other charts. Although, it is a very fine test, it cannot be used just the way it is for the stutterer. In Horvath's approach the subject is told that he will be asked the same questions as in other tests, but in the silent answer test he is not to give an oral answer, merely to silently and truthfully answer questions. Thus, in this situation, the subject who has been answering truthfully orally doesn't have to change his answers. But if he has been lying orally and he is now told to answer truthfully to himself he must either tell the truth, which means he changes his answer, or he must lie again and disobey the instructions. It creates a dilemma for the liar.

Although Atwood's test differs somewhat from Horvath's, the same dilemma is apparent. Atwood's modification involves instructing the subject to answer silently, to himself, exactly the way he answered the questions during the pretest interview; warning him not to change his answers. Hence, the subject will be either deliberately disobedient if he decides to "beat the test" by changing his silent answers, or will be faced with the act of lying if he answers the questions as instructed. For full effectiveness the pre-test instructions must be made clear to the subject. That is, the subject must know that if he changes answers during the test, when he is answering silently, this is going to create a problem for him, that he will react because of his disobedience. Failure to properly instruct a subject will probably lessen chances of detecting deception. Indeed, this is apparent from the research reported by Gustafsen and Orne (1965) where they demonstrated the value of the direct verbal answers, noting that those who said "no" to questions were more frequently detected than those who were told to say nothing.

Gustafsen used relevant/irrelevant tests in a laboratory situation and was able to detect 19 of 25 people when they answered "no". But only 14 of 24 persons could be picked out when they remained silent. By employing a peak of tension test, using numbers in a known sequence, he was able to pick 20 of 27 correctly when they said "no" but could only pick 13 of 26 from those who remained silent. These statistics suggest that it is important to create a specific dilemma for the subject if you are going to use a silent answer test. And a silent answer test is one of the approaches in handling a stutterer.

Since a stutterer often can whisper without stuttering, if the examination room is quiet, it may be appropriate to tell the subject to whisper his answers; perhaps conducting a trial run to determine if this approach is practicable. I have also been told of cases where the stutterer has been instructed to answer with a very slight nod.

Some subjects engage in false stuttering, which, by the way, is fairly hard to do. In such cases, it may be evident that the stutterer does not sound right. If testing is actually conducted in such instances, the charts may not look right as the stuttering won't be at the end of the inspiration cycle, where it usually occurs. Tom Moore of the Metropolitan Police Department in Washington, D.C. reported a case of a fake stutterer who was told by his attorney to stutter only on the pertinent questions. Moore noted that the pneumo pattern was wrong; the stuttering wasn't at the end of the inspiration cycle. He was right, and the subject confessed.

Deaf and Dumb Subjects

Deaf and dumb subjects present problems different from those of stutterers. The deaf person is not as much of a problem as the person who is both deaf and dumb. When dealing with these subjects the first thing to do is to make extensive preparation. After such preparation, practice the test on someone before you actually conduct it. Then you must be able to answer some questions about the particular subject. First, how deaf is this subject? If he can hear a little or not at all? Can he read lips? Paul F. Rohde had a case where the subject was able to read lips but he refused to cooperate. Incidentally, the reason he did not want to cooperate was that he thought that if he created a problem, the test would not be conducted. With deaf subjects you will spend a lot of time passing notes back and forth. Since you have to do this, bring the subject alongside, like you were playing the piano together. It is easier for the subject to read the pre-test instructions you have written for him, and to write his answers right on them. If you each use a pen of a different color you can later attribute the notes to the right person, but be sure to write on the same pad, taking turns.

You must agree on the questions as you do with anyone else, but if you can stay with your originally planned questions, this is easier as you can then use prepared transparencies or cards, without altering them. We once tried 35 mm slides but gave up on them because there is no way to alter the photograph of the question to agree with the subject's objections or changes. You may use any one of three techniques: a transparency projector, an opaque projector, or a 5 x 8 card technique. The card technique is simple, but there is one point to note. The

appearance of the card causes an initial response. So first you make a dry run with the subject, which also gets your timing down while you mark the chart and handle the cards. When you put the question out in front of the subject, put the blank or the backside out first. Turn the question card over, let him read the question, turn it back to the blank side, and take it away. The first time I did this I had it timed perfectly. I put the card out in front of the subject with the blank side facing him, I turned it over and there was the question-upside down. It really is necessary to practice this technique.

With a projector you may need a second person in the room. If your subject can hear a little, then signal the person to use the projector rather than say, "Ask the question," or you will have a stimulus in front of your question. This technique also needs practice, with someone in your own office acting as the subject, and another as your assistant.

Many deaf people can answer "yes" or "no", and if so have them to do so during testing. But, if it is quite an effort for them to speak, and it takes quite a while to work up to an answer, then I suggest you not have him speak. A slight nod will do. You may also use the Atwood pretest and a silent answer test.

Raymond D. Inglin of the Los Angeles Police Department had a case where the person was not deaf or dumb, but refused to answer questions. He agreed to take a polygraph test but he refused to answer questions during the tests. Inglin conducted a peak of tension test on the caliber of the weapon and the location in which the victim had been shot. The subject reacted perfectly to the right items in this test. He also reacted to the relevant questions during a zone comparison test. He was later found guilty of murdering his girlfriend. But the situation here, where there is no answer, is different from those laboratory studies conducted by Gustafsen and Orne. There was little pressure or stress in the laboratory where subjects were simply picking numbers. In Inglin's case, however, the man had very much to lose and even though he acted like a mute, he was examined with good results. It is, however, an isolated case.

There was a case in Vermont where a deaf mute, who was also illiterate, was successfully tested. He was 23 years old but still living with his mother, incapable of being employed. To explain why he was away all night he alleged that a certain man, whom he identified, had dragged him into a hotel room and forced him to commit indecent acts all night. The mother came to the police and explained the story to them.

The Vermont State Police decided to test the victim first. Now they were presented with something of a problem. This man was deaf, mute, and could not read. They asked the Principal of the nearby school for the deaf to come in and act as an interpreter. In this case, the Principal sat in front of the subject's chair and used sign language for questions that the examiner pointed to. The subject nodded very slightly for answers, "yes" and "no". After the first chart he decided he wanted to change his story and he admitted that he was not forced into the hotel room, but he was lured in on the promise of something to drink and eat. After the third chart he changed his story again. He said he hadn't been lured in, that he had gone willingly, and had enjoyed the whole evening, but he had to explain to his mother why he was out all night. The part of the story as to what they did, however, was true. A difficult test to conduct, but the Vermont State Police took adequate time for preparation, obtained assistance, and completed the examination and interrogation.

If a deaf-dumb subject is able to read lips, this fact may also be a problem. If the subject faces you during the test to read your lips, he will also be reading your reactions to the chart and to his replies. Because the deaf are generally very good at reading facial expressions, I would suggest that you do not conduct the examination with the subject facing you. Rather, use a third person to act as an "interpreter." If the subject reads the facial expressions of the "interpreter", he will not be aware of your reactions to his polygraph charts.

Blind Subjects

When testing blind people you have problems which differ from those presented by other handicapped persons. For instance, you

will have to read the release to them; guide their signature, or at least get their hand in the right place; and, perhaps, adjust the room.

When testing blind persons you will want to know if they have been blind from birth or as a result of a recent occurrence, because it affects the way blind people describe things. In addition, you will want to know whether they are totally blind, or can see some light, or are merely legally blind. Some legally blind people have quite a bit of vision. When you put the attachments on a blind person, let them feel them first. Let them feel the blood pressure cuff, the electrodes and the pneumo tubes, and tell them what each one does. Remember, they cannot see that attachment and they don't know what it looks like. They will be so sensitive to sound and touch that they will notice vibrations you are not aware of. They will note subtle inflections of your voice. They will be aware of your turning your head away while speaking.

In one particular case the subject was not only blind, but was wearing a pacemaker. Despite these handicaps, the test was successful. Although there were no changes in heart rate, blood pressure changes were evident.

One of the toughest cases known involving a handicapped person was conducted by Mr. Paul F. Rhode, in which there was a tremendous response. There was a GSR response beyond belief, there was a spectacular blood pressure rise, an acceleration of pulse rate, and a huge pneumo reaction. Unfortunately none of this was recorded on the chart. What happened was that Paul Rhode was putting the attachments on a blind man when the seeing-eye dog suddenly made vicious growl and jumped at the examiner's throat. The reaction was Paul Rhode's. Be sure that it is all right with the dog when you put those attachments on. A word from the dog's master is a wise precaution.

Miscellaneous Problems

Subjects with palsy create a problem with their constant movement and speech defects. The movement may prevent the use of regular cardio units, but a plethysmograph taped to the subject works very well. The speech is often slow and difficult, even a "yes" or "no" answer may take time. Palsy is common with Parkinson's disease, cerebral tumors and lesions, and sometimes with multiple sclerosis. Subjects who have missing or deformed arms or legs can be tested without trouble. If the arms are deformed, use a blood pressure cuff on the ankle or calf. Elevating the leg on a chair or the use of a reclining chair will improve the pattern. Dwarfs, who are defined as being between two feet and four feet ten inches, are a problem only in that they have a tendency to slide out of an ordinary chair; and require a smaller blood pressure cuff. Use an infant or child cuff. A reclining chair will solve the movement problem. Hemophiliacs should not be tested with a blood pressure cuff at all because of the possibility of vascular or veinous damage. The use of a photoelectric plethysmograph or a cardio activity monitor is recommended as these attachments are passive. If the subject is in a body cast, the pneumograph recording may be difficult to obtain. In one case, the pneumograph tube was placed over one shoulder with the chain under the opposite arm. The polygraph instrument had an amplified pneumograph section which produced a satisfactory pattern. Both the Lafayette model 76164 and the Stoelting Polyscribe have this feature. In another case the body cast did not extend much below the rib cage, and a satisfactory pattern was obtained from a low abdominal position. Although some retarded subjects may be tested; it is not always possible. Research on institutionalized retardates (Abrams and Weinstein, 1974) indicated that they were not fit subjects.

Equipment

Most testing of the physically handicapped requires only some imagination and practice. The use of a polygraph instrument with either the plethysmograph or cardio activity monitor is often useful. So are amplifiers on the pneumograph channels. The use of the highly adjustable polygraph chairs produced by Stoelting or Lafayette, or a commercial model reclining chair which permits tilting the subject back, have advantages with handicapped subjects. The former reduces arm movement and is adjustable to

the subject's size. The recliner reduces movement, is more comfortable, and permits the best examiner observation. The use of an overhead projector or opaque projector is useful in handling deaf subjects. The use of the low pressure cardiosphygmographs is also useful in reducing discomfort.

Conclusion

The successful examination of handicapped subjects depends upon careful preparation of questions and adaptation of techniques. Rehearsal with an assistant is strongly recommended before working with deaf, dumb, and blind subjects. Stutterers may be able to whisper their answers, or a "silent answer" test may be employed. The examiner will need extra time for preparation and conduct of these cases, but there is no reason to avoid them.

References

Abrams, Stanley and Weinstein, Edwin. "The Validity of the Polygraph with Retardates," *Journal of Police Science and Administration,* 2 (1) (March, 1974) pp. 11-14..

Barbara, Dominick A. *Stuttering: A Psychodynamic Approach to its Understanding and Treatment*. New York: Julian Press, 1954.

Battin, R. Ray. *Speech and Language Delay*, 2nd ed. Springfield, Illinois: Charles C. Thomas, 1970.

Berry, Mildred F. *Speech Disorders*. New York: Appleton, 1956.

Bloodstein, 0., Jasper, W., and Tureen, J. "A Study of the Diagnosis of Stuttering by Parents of Stutterers and Non-Stutterers." *Journal of Speech and Hearing Disorders*, 17 (1952).

Canfield, Norton. *Hearing*. New York: Doubleday, 1959.

Carrell, James A. *Disorders of Articulation*. Englewood Cliffs, New Jersey: Prentice Hall, 1968.

Geldard, Frank. *The Human Senses*. New York: Wiley, 1953.

Gustafson, Laurence A. and Orne, Martin T. "The Effects of Verbal Response on the Laboratory Detection of Deception." *Psychophysiology* 2(1)(1965) pp. 10-13.

Hahn, Eugene. *Stuttering: Significant Theories and Therapies*, 2nd ed. Stanford: Stanford University Press 1956.

Horvath, Frank and Reid, John E. "The Polygraph Silent Answer Test." J*ournal of Criminal Law, Criminology, and Police Science*, 63(2) (1972) pp. 285-293.

Kleinfield, Sonny. "Dwarfs," *Atlantic Monthly*, September, 1975, pp. 62-66.

Moore, G. Paul. *Organic Speech Disorders*. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Small, Arnold M., Jr. "Hearing." *International Encyclopedia of the Social Sciences*. New York: Macmillan, 1968, VI, pp. 336-339.

Smith, Karl U. "Deafness." *McGraw-Hill Encyclopedia of Science & Technology*. McGraw-Hill, 1971. IV, pp. 29-37.

Taylor, Eugene J. (Ed.) *Rehabilitation and World Peace: Proceedings of the Eighth World Congress of the International Society for the Welfare of Cripples*. New York: International Society for Rehabilitation of the Disabled, 1960.

Tikofsky, R. S. and Wang, W. "Speech Disorders." McGraw-Hill *Encyclopedia of Science and Technology*. New York: McGraw-Hill, 1971. XII, pp. 733-735.

Wright, David. *Deafness*. New York: Stein and Day, 1969.

Van Riper, Charles. *Speech Correction: Principles and Methods*, 3rd ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1954.

Zweilful, R. L. "Adaptation of the Opaque Projector to Lie Detector Demonstration." *Journal of Criminal Law, Criminology and Police Science*, 43 (1952) pp. 256-258.

Development of Deception Criteria Prior to 1950

Norman Ansley

Abstract

This is a review of the literature published up to 1950 that contributed to the current list of physiological responses considered deception criteria. Even making allowances for differences in terminology, there are deception criteria in the current DoDPI list that had not been observed, or if observed, not described before 1950. An appendix describes Luria's motor movement technique and Wertheimer's word association test. As means of detecting deception, both were discontinued before 1950.

Key words: cardiovascular, deception criteria, electrodermal, motor movement, polygraph history, respiration, terminology, word association.

This is a review of the literature published up to 1950 that contributed to the current list of physiological responses considered deception criteria. That year marked the halfway point for the development of polygraph testing, as we know it in 1999. In 1950 the only formal polygraph training was at the Keeler Polygraph Institute, and most examiners were preceptor trained or self-taught. Most of the instruments were two-channel (cardiograph and pneumograph) mechanical units, although there were some with electrodermal units. The most widely used technique was relevant-irrelevant. A few examiners used one or the other of two published Control Question techniques, one published by Summers (1939), and the other by Inbau (1948). Among the many shortcomings in 1950 was a lack of agreement on what constituted deception criteria. Add inadequate chart markings, and that independent analysis of someone else's charts was difficult, and the results were problematic.

In 1950, Charles M. Wilson, president of the International Society for the Detection of Deception (ISDD), was asked, "Should graphs be released or shown after the test?" Wilson's reply was printed in the *ISDD Bulletin*. He said that in his experience he never released an original record to anyone. He did not think making copies a good policy since possession of the record by an untrained operator represents the first step in the direction of perversion and quackery. Wilson said the charts mean nothing to anyone who was not present when the tests were run, and the only use to which they could be put was to cloud the issue (Wilson 1950).

If one examiner could not reliably read charts from another examiner, what did they know about chart interpretation in 1950? In this paper we list sixteen studies or reports which included something on deception criteria. The sixteen studies or texts did not discuss rank order scoring, only two had a form of numerical analysis, and computers were not yet useful machines. Taking the Department of Defense Polygraph Institute (DoDPI) list as state of the art for hand scoring in 1999, how many of the criteria had been identified by 1950? In the pneumograph tracing, DoDPI lists 12 items. Seven had been identified by 1950: I/E ratio change, amplitude increase, amplitude decrease/ suppression, amplitude progressive decrease and return to homeostasis, respiration baseline change – temporary, baseline change – permanent, and apnea – blocking. By 1950, they had not yet observed respiration rate increase, rate decrease, respiration amplitude progressive increase followed by decrease, amplitude progressive increase and return to

This article was originally published in *Polygraph*, 1999, 28(1).

homeostasis, and apnea – holding. Considering terminology, they might have seen the difference between holding and blocking but did not think the difference mattered. Some of today's examiners might have trouble recognizing DoDPI's more exact definitions of staircases up, staircases down, and staircases up and down. DoDPI lists three criteria under electrodermal, and two, amplitude and duration, were in the pre-1950 literature. Only the complex response, which some examiners call a saddle, was not mentioned. DoDPI includes eight deception criteria under cardiograph including premature ventricular contractions which were not listed prior to 1950, but some would say should not be listed now. Of the seven others on the list, six were known: phasic increase and decrease in baseline, tonic increase in baseline, tonic decrease in baseline, pulse rate increase, pulse rate decrease, and decrease in amplitude of the tracing. The one lacking in 1950 was the increase of the amplitude of the tracing. It would appear that the well-informed examiner of 1950 had enough deception criteria to decide most of his cases, but the more we go back in time, the less he had. The cumulative growth of a body of technical and scientific knowledge is a vital part of a profession. In the text that follows we will see the development of knowledge.

One wonders if the pioneers in instrumental detection of deception knew of Daniel Defoe's proposal to take the pulse of a suspected thief. One would think he was discussing a modern polygraph problem when he observed, "It may be true that this discovery by the pulsation of the blood cannot be brought to a certainty, and therefore it is not to be brought into evidence; but I insist, if it be duly and skillfully observed, it may be brought to be allowed for a just addition to other circumstances, especially if concurring with other just grounds of suspicion." (1730) (Moore 1955).

Cesare Lombroso (1911) mentions a case in which he used his recording hydro-sphygmograph. His apparatus measured blood volume and pulse rate. He reports, "The same apathy persisted when he was spoken to of the robbery on the railroad, while there was an enormous depression – a fall of 14mm – when the Torelli theft was mentioned. I concluded, that he had no part in the railway robbery, but he had certainly participated in the Torelli affair; and my conclusions were completely verified." Here we have a measure of a cardiovascular reaction, and a verified decision.

In 1914, Vittorio Benussi published the results of an experiment relating to the symptoms of lying in respiration. At the University of Graz in Leipzig, Benussi had subjects read aloud five statements, some of which were coded and not to be read as stated. Half of the items in the 80 experiments were to be lied about. Panels of witnesses made judgments as to when subjects were lying, and when they were telling the truth. Using a Marey pneumograph which recorded on a polygraph, Benussi measured the distance between the beginning and end (length) of each of three, four, or five cycles of breathing after the subject spoke. For each cycle of breathing Benussi measured the length (time) of the inspiration (I) and the length of the expiration (E), and calculated the ratio (I/E) for each of the cycles before and after the statement. He found that lying produced greater I/E ratios than truthfulness. Of the 80 experiments, I/E analysis resulted in one false positive error and one false negative error, for a total accuracy of 97.5%. The average panel accuracy was 56% for truthful and 58% for deceptive statements. This experiment attracted the attention of Marston, Larson, and others to the diagnostic value of a respiratory recording. Also, the I/E ratio has remained on the deception criteria lists of the DoD Polygraph Institute, The Maryland Institute of Criminal Justice, and other polygraph schools and courses.

John A. Larson (1923) had experience in hundreds of criminal cases as a basis for his description of deception criteria. Larson recorded a continuous cardiograph and pneumograph pattern on a smoked drum apparatus, and observed that the record of the innocent suspect will usually vary but slightly, if at all, from its normal. In describing some guilty test results, Larson describes repression in the pneumograph tracing, and the accompanying chart illustration shows a rise and fall in the cardiograph pattern of the confirmed deception. In another chart we see suppression, loss of baseline, and changes in the I/E ratio and rhythm and regularity in the

pneumograph and a rise and fall in the cardiograph tracing, but his text does not describe this illustration. Larson notes that the cardiac curve is usually more significant than the respiratory curve. In the description of a chart, Larson writes of the extreme blocking effect of deception. In one chart Larson described deception causing a drop in the blood pressure curve with the obliteration of the pulsations. In addition, there was an increase in frequency. Describing another chart segment with a lie, Larson notes in both the cardiac and respiratory curves there was repression. Larson states the following changes have been observed as the effect of deception. These changes may occur in both the cardiac and the respiratory curves or in one alone, more frequently in the cardiac action:

1. Increase in blood pressure – a rise.
2. Decrease in blood pressure.
3. Increase in height.
4. Increase in frequency.
5. Summative effects.
6. Incomplete inhibition.
7. Complete inhibitory effect.
8. Irregular fluctuations, especially noticeable at the base of each cardiac pulsation.
9. Combination of any of the above effects in the same individual.
10. These changes may occur with but little latent period, or then may be accumulative in effect and more generally distributed.

Leonarde Keeler (1930) wanted to compare the peak of tension polygraph technique with the word association method. Seventy-five subjects took a one-chart peak of tension on which of ten cards they had chosen. If the chosen card, placed by chance, was first or last in the sequence, the test was repeated in a different sequence. The deception criteria were a rise in blood pressure followed by a release in tension after the chosen card, and the greatest suppression in the respiratory tracing. There were 71 correct decisions of 75 (95%) on the first trial. Posttest interviews attributed the failures to a lack of interest or concern which resulted in a lack of responses. Here we have pneumograph suppression and a rise and relief in the cardiograph pattern established as valid deception criteria. By comparison, the word association test of 30 students was correct in 19 (63%). For results of another comparison see the work of John E. Winter (1936) in a dormitory theft case.

Professor John E. Winter (1936) investigated thefts in the women's dormitories at West Virginia University with two methods: Jung's word association test with a chronoscope for reaction time, and a Larson type polygraph test employing respiratory and cardiovascular measures, from separate devices. The breathing curve was rated as regular or irregular; light or deep. The blood pressure curve was rated as regular or irregular, and medium or strong. Winter gave three levels of significance to the results of each of the methods: 0 for no significance, "nothing to indicate guilt;" 1 for "some significance and points in direction of guilt;" and 2 for "distinct signs of guilt." There were 25 women suspects and each received two Larson type tests, with consistent responses except for the culprit. The first test of each subject was labeled practice. From the respiration recording there were 24 zeros, including the thief, who confessed. On her practice she scored a 2 on her cardiograph curve, the only one to do so. She was given a post-confession test where she again scored a 2 on the cardiograph curve. This may be the first case of numerical scoring. Word association cleared 19 innocent suspects, and had the thief among the five who scored a 1.

Winter's polygraph apparatus was reported as "an ordinary pneumograph, a Baumanometer, an improved form of the Erlanger capsule for high and low air pressure, and a MacKenzie polygraph for a continuous record of breathing and heart action." For a picture and description of the MacKenzie polygraph, see *Polygraph* (1992) 21(4) 349-350.

C.D. Lee wrote an article, "The Lie Detector," published in the September, 1937 issue of the *Fingerprint and Identification Magazine*. Lee illustrates the article with a picture of a chart from the examination of Jerone Selz who confessed to murder after the test. There was a double rise and fall in the cardiograph pattern and suppressed respiration following the question, "Did you kill Mrs. Rice?" The remainder of the article is about the instrument and testing.

Leon G. Turrou (1938) in his book *Nazi Spies in America* describes several polygraph examinations given to suspects and witnesses involved in a German espionage ring. Turrou describes how the instrument functions (cardiograph and pneumograph), then quotes Keeler on the procedure. Eight suspects or witnesses were tested. Because many questions were asked each examinee, a system of asterisks was devised to give some indication of results. In the report, one asterisk after a question indicated a mild emotional reaction, two a strong emotional reaction, and three asterisks, quite an emotional reaction, "such as would be found when the subject is telling a whopper." One examinee was asked nine relevant questions. There were no asterisks behind four of the questions, two asterisks behind one question, and three asterisks behind four questions; a split call from a multiple issue relevant-irrelevant test format. During the testing of a suspect, 18 relevants were asked, and in the report there were no asterisks behind five of the relevant questions, one asterisk behind four of the relevants, two asterisks behind five relevants, and three asterisks behind four relevants. This evaluation of the charts was unusual, at least unusual to appear in the report. In reality, the asterisks were a numerical system, zero to three, for each question.

William M. Marston published a book in 1938. Under the heading "Judging a Polygraph Record," Marston states that changes in the blood pressure are the chief and only dependable criterion of deception. This is shown by the shifting of the entire mass of pulse tracings toward the upper edge of the recording strip. Variations in the pulse are not significant. Regarding breathing, Marston said marked changes in respiration tracings that accompany changes in the blood pressure justify a judgment of deception. He noted that Benussi's breathing ratios are probably extremely significant of lying, but it has never proved practical. Marston said a sudden hump in the breathing record may be meaningful, as may a "shoulder" in either the inspiration or expiration tracing. Also indicative is a sudden irregularity indicating a "catching of the breath," or an unaccountable flattening out of the whole respiration tracing indicating an extended series of shallow breaths.

The Reverend Walter G. Summers, S.J., prepared a paper on his work before his death on September 24, 1938. Published in 1939, it describes a sophisticated test format and means of chart analysis. In a theft case there would be three relevant questions. In sequence the questions asked about knowledge, guilt, and possession. Called "significant" questions, examples were, "Do you know who took the money?", "Did you take the money?", and "Have you the money on your person?" He said that within one record there were usually included three different but related significant questions, each of which was asked three times. Interspersed among the non-significant questions (irrelevants) are emotional standard questions (controls). An emotional standard question precedes each significant question. The format is three pairs of control-relevant questions, with irrelevants put in as needed. Examples of irrelevants were, "Are you wearing a black coat?" and "Did you eat breakfast this morning?" Examples of emotional standards, developed after extensive interviewing of the examinee were, "Were you ever arrested?" and "Do you own a revolver?"

The analytical system is modern. Summers said "...we contrast and compare the reactions to the significant questions with the reactions to the emotional standards. If the reactions to the significant questions are consistently greater than the deflections to the emotional standards, the individual is consciously trying to deceive the examiner. If, on the other hand, the deflections to the critical questions are not consistently greater than those to the emotional standards, the individual is truthfully expressing his state of mind. This is the essential criterion of interpretation." Professor Summers used a recording galvanometer, the Fordham Pathometer, which he manufactured. A letter to the author from William E. Kirwan in 1952, indicated the New York State Troopers Scientific Laboratory was still using the Summer's technique, with excellent results (Kirwan 1952). Summers, who conducted laboratory and criminal cases, established the control question test concept, including the analytic procedure (Summers 1939).

Paul Trovillo (1942) wrote what is probably the first treatise on the topic of deception test criteria. The illustrations were

taken from real cases. Although the electrodermal unit was not widely used, there is a good section of illustrations of GSR tracings. For the cardiograph he lists and illustrates:

1. Common form of blood pressure rise (and return to baseline).
2. Blood pressure increase . . . complicated by cyclical increase throughout the graph.
3. Rapid rise and decline in blood pressure, accompanied by obliteration of pulse amplitude.
4. Gradual increase in blood pressure.
5. Constriction of pulse amplitude and gradual rise in blood pressure.
6. Slight rise accompanied by rapid decline in blood pressure.
7. Peak of tension.
8. Rapid changes in heart rhythm.
9. Another form of change in heart rhythm (includes general pulse irregularity).
10. Complication of deception pattern – increase in blood pressure and return to baseline, variations in pulse frequency, and reduction of pulse amplitude.
11. Reduction in pulse amplitude.

For the respiration tracing, he lists and illustrates:

1. Suppression at point of deception.
2. Respiratory block.
3. Rise in baseline.
4. Respiratory suppression preceding deception stimulus, followed by deeper respiration at point of deception.
5. Regularity of respiration up to and through the deception stimulus, followed by irregular respiration.
6. Respiratory irregularities up to point of deception, followed by regular respiration.

For the electrodermal he lists and illustrates:

1. Comparatively large area of reaction at point of deception.
2. Comparatively large magnitude of reaction at point of deception.
3. Peak of tension test (experimental age test), reactions to each age up to and including the point of deception, then none.
4. Peak of tension card test. The only large reaction.
5. Peak of tension. Pattern at deception different from patterns at truthful answers.
6. Gradual rise in the electrodermal pattern.

Trovillo then lists and illustrates what he calls ambiguities in the records. In the cardiograph tracing he shows the effect of body movements, a deep breath, general excitement, increase in blood pressure even at irrelevant questions, an absence of blood pressure and pulse rate changes during lying, inconsistency of reactions on questions involving guilt, startle response of innocent subjects, and a cardiac irregularity.

For ambiguous respiratory patterns he lists and illustrates: deception-like suppression found among some innocent examinees, effects of superfluous talking and physical movement, erratic breathing of an innocent person from great fear, a deep breath taken deliberately to obliterate suppression, normal shallow breathing following a deep breath, effect of sinus congestion, lack of response in known guilty subject, and respiratory tremor found in both relevant and irrelevant questions by an excited person.

For the ambiguous electrodermal patterns he lists and illustrates: over-activity of the reaction, effects of bodily movement, effect of deep breath at the very moment of response, unresponsiveness in guilty subject, inconsistent reactions in guilty subject, and guilt reactions in innocent persons.

In 1942, Fred E. Inbau published the first of his three books on *Lie Detection and Criminal Interrogation*. The techniques were relevant-irrelevant and peak of tension. In the section on deception criteria he notes that the criteria differ somewhat for the two techniques. For the cardiograph he mentions an increase in blood pressure and the illustration shows it returning to baseline after first going below the baseline. Other criteria include a sharp drop in blood pressure, and slowing of the pulse rate. For the respiration pattern he lists suppression, and heavy breathing about twenty or twenty-five seconds after the reply to a question. Inbau writes about the EDA and the lack of knowledge about it, and concludes that electrodermal tracings alone cannot be considered as adequate for deception diagnosis, but it may

be occasionally helpful as an adjunct to the other recordings.

In 1943, C.D. Lee prepared an *Instruction Manual for the Berkeley Polygraph*. It is a complete text on conducting examinations and reading the charts. The methods are relevant-irrelevant and peak of tension. He notes the pattern of the innocent is one of regularity and uniformity with no marked difference between the effect produced by neutral questions and those related to the crime. The tension may remain constant, decrease, fluctuate slightly, but seldom increase. In the guilty, the tension is lacking in regularity and uniformity. Illustrations show a phasic rise and fall of the cardiograph pattern associated with deception. In the pneumograph, he shows repressed breathing, followed later by a sigh of relief. Most of the illustrations were of the cardiograph pattern, and the rise and fall of the cardiograph pattern is clearly the primary indication of deception.

Joseph W. Haney (1944) was a forensic psychologist and experienced polygraph examiner in the Chicago Crime Laboratory. He was interested in the catalogue of deception criteria by Paul V. Trovillo. Haney wondered if the respiration responses described by Trovillo might not be produced by a nondeceptive mental task as well as deception. Haney did that, producing charts with blocking (apnea), suppression, and baseline rises. Haney suggested that before using these as deception criteria, one should see if they occur also at irrelevant questions.

In 1948, Fred E. Inbau published the second edition of his book *Lie Detection and Criminal Interrogation*. In addition to the relevant-irrelevant and peak of tension tests there was the Reid control question test. The section on deception criteria has not changed in a significant way. For the cardiograph, Inbau mentions an increase in blood pressure, and the illustration shows it returning to baseline after first dropping below the baseline. Other criteria include a sharp drop in blood pressure, and slowing of the pulse rate. For the respiration system he lists suppression, and heavy breathing about twenty or twenty-five seconds after the reply to a question. In regard to the electrodermal channel, the author said electrodermal responses have been found to be of little practical value in diagnosing deception.

Baesen, Chung & Yang (1948-1949) tell us the chart criteria they used in a laboratory research project employing a two-channel Keeler polygraph. They used pulse rate changes, sudden and delayed drops in blood pressure, duration of rise and fall in blood pressure, and location of the dicrotic notch. Notice was taken of changes in respiration baseline, blocking and suppression of respiration either prior to, during, or immediately following the question.

In 1950, Colonel Ralph W. Pierce, president of Leonarde Keeler, Inc. was writing about the use of the peak of tension test. He wrote, "One man reacted to this test, his blood pressure rising until the question concerning the German Luger was asked, then falling off. He also showed marked irregularity in his breathing up to the question about Luger, followed by regularity to the end of the test. The galvanometer pen also rose sharply at the question concerning the Luger. This man also reacted similarly to the other tests referring to the disposition of the gun, its condition, etc." The examinee confessed to the crime. Colonel Pierce's description has tonic changes in the cardiograph and pneumograph tracings, but a phasic response in the electrodermal, channel showing a combination of deception criteria.

Abandoned Methods for Detecting Deception

In the period before 1950 there were two techniques that were subject to considerable research as means for detecting deception. Their criteria for deception were not related to the methods in the polygraph technique.

Luria (1930, 1932) developed a lie detection method that involved tremors and motor movement. It received some research attention in the United States but was not used in criminal cases (Berrien, 1939; Morgan & Ojemann 1942).

From the turn of the century into the 1930s the word association test was considered a method for detecting guilt in criminal cases (Wertheimer & Klein 1904, Jung 1919). However, Larson (1922) found

"the association words with time reaction do not give as satisfactory results as the cardiorespiratory changes." Larson added, "We can say this definitely in cases where the suspect has subsequently confessed where, although there were marked and striking changes in the tracings, the findings by association method were not significant." Keeler (1930) found the association method performed poorly when compared to polygraph test results. Winter (1936) in a real case of theft involving 25 students, found the association method had the thief among five in a narrowed pool of suspects, but his cardio-pneumo method identified the culprit, followed by a confession. Although word association with reaction time remains as a psychological tool, its use in solving crime has disappeared.

References

Baesen, H. V., Chung, C. & Yang, C. (1948-49). A lie detector experiment. *Journal of Criminal Law, Criminology, and Police Science*, 39, 532-537.

Benussi, V. (1914). Die atmung asymptome der luge (The respiratory symptoms of lying). *Archiv fur die Gestamte Psychologie*, 11, 244-273. Translated and printed in *Polygraph*, 4(1), 52-76.

Berrien, F. K. (1939). Finger oscillations as indices of emotion. I Preliminary Validation, *Journal of Experimental Psychology*, 24, 485-498; and II Further validation and use in detecting deception, *Journal of Experimental Psychology*, 24, 609-620.

Haney, J. W. (1943-44). The analyses of respiratory criteria in deception tests, a possible source of misinterpretation. *Journal of Criminal Law and Criminology*, 34, 268-273.

Inbau, F. E. (1942). *Lie Detection and Criminal Interrogation*. Baltimore: Williams & Wilkins Company.

Inbau, F. E. (1948). *Lie Detection and Criminal Interrogation*. Second Ed., Rev. Baltimore: Williams & Wilkins Company.

Instruction Manual for the Berkeley Psychograph (1943). San Rafael, California: Lee & Sons.

Jung, C. G. (1919). *Studies in Word Association*. New York: Moffet, Yard & Company.

Keeler, L. (1930). A method for detecting deception. *American Journal of Police Science*, 1(1), 38-51.

Kirwan, W. E., (16 Oct 1952). Letter from the Director or the New York State Troopers Scientific Laboratory to Norman Ansley.

Larson, J. A. (1923). The cardio-pneumo-psychogram and its use in the study of the emotions, with practical applications. *Journal of Experimental Psychology*, 5(5), 323-328.

Larson, J. A. (1923). The cardio-pneumo-psychogram in deception. *Journal of Experimental Psychology*, 6(6), 420-454.

Lee, C. D. (Sep 1937). The lie detector. *Fingerprint and Identification Magazine*.

Lombroso, C., M.D. (1911). *Crime: Its Causes and Remedies*. Boston: Little, Brown & Company. Excerpt reprinted in *Polygraph* (1991) 20 (1) 52-53.

Luria, A. R. (1930). Die Methode der Abbildenden Motorik in der Titbestandiagnostik. *Zeitschrift fur Angewandte Psycholgie*, 35, 139-183.

Luria, A. R. (1932). *The Nature of Human Conflicts: On Emotion, Conflict and Will*. New York: Liveright.

Marston, W. M. (1938). *The Lie Detector Test.* New York: Richard R. Smith. Reprinted by the American Polygraph Association, 1989.

Moore, J. R. (1955). Defoe's project for lie-detection. *American Journal of Psychology,* 68, 672.

Morgan, M. I. & Ojemann, R. H. (1942). A study of the Luria method. *Journal of Applied Psychology*, 26, 168-179.

Pierce, R. W. (1950). The peak of tension test. *ISDD Bulletin*, 3 (3) 1, 4, 5.

Summers, W. G. (1939). Science can get the confession. *Fordham Law Review*, 8, 335-354.

Trovillo, P. V. (1942). Deception test criteria: How one can determine truth and falsehood from polygraphic records. *Journal of Criminal Law and Criminology*, 33 (4) 338-358.

Turrou, L. G. (1938). *Nazi Spies in America*. New York: Random House.

Wertheimer, M. & Klein J. (1904). Psychologische Tatbestandsdiagnostik. *Achiv fur KriminalAnthropologie und Kriminalistic*, 15, 78-113.

Wilson, C. M. (1950). Should graphs be released or shown after test? *ISDD Bulletin*, 3 (3), 5, 8.

Winter, J. E. (1936). A comparison of the cardio-pneumo-psychograph and association methods in the detection of lying in cases of theft among college students. *Journal of Applied Psychology*, 20, 243-248.

Attachment 1
Deception Criteria for Lie Detection Pioneers

	Defoe	Lombroso	Benussi	Larson	Keeler	Winter	Lee	Turrou	Marston	Summers	Trovillo	Inbau	Lee	Haney	Inbau	Baesen et al	Pierce
	1730	1911	1914	1923	1930	1936	1937	1938	1938	1939	1942	1942	1943	1944	1948	1949	1950
Respiratory																	
rate decrease																	
rate increase																	
I/E ratio change			X						X								
amplitude increase						X			X		X	X	X		X		
amplitude decrease-suppression				X	X	X	X		X		X	X	X	X	X		
progressive increase-decrease																	
progressive increase & return																	
progressive decrease & return									X								
baseline change - temporary															X		
baseline change - permanent											X			X	X		
apnea – holding (inhalation)																	
apnea – holding (exhalation)				X							X			X	X		
Electrodermal																	
amplitude change										X	X						X
complex response																	
response duration & return											X						
Cardiovascular																	
baseline increase & decrease				X	X		X				X	X	X		X		X
baseline increase				X		X			X		X						
baseline decrease		X		X		X						X				X	
amplitude increase																	
amplitude decrease											X						
pulse rate increase	X			X							X						
pulse rate decrease											X				X	X	
PVCs																X	
Numerical analysis						X			X								

The Frequency of Appearance of Evaluative Criteria in Field Polygraph Charts

Norman Ansley and Donald J. Krapohl

Abstract

Every appearance of each of 22 response patterns considered to be diagnostic for the detection of deception by the US Department of Defense Polygraph Institute (DoDPI) was tabulated for 177 cases (616 polygraph charts) selected from the DoDPI database of confirmed field cases. The sets of charts were in 16 different formats, but all were a form of zone comparison. We found the total number of appearances of these criteria ranged from 5 to 4,793. A rank ordering by frequency of the 22 criteria stayed remarkably constant across questions, gender, and truthful or deceptive status. There was a reduction in the number of reactions in the second and third charts of nondeceptive examinees in all three physiological channels and a similar reduction in the electrodermal and cardiograph channels of deceptive examinees. However, the respiratory pattern showed an increase in reactions in successive charts of deceptive examinees. We also found more reactions and a higher tonic heart rate for the deceptive than the nondeceptive examinees. The 1,780 relevant question presentations produced 6,453 reactions, for an average of 3.6 reactions per question. The 1,932 comparison questions produced 6,777 reactions, for an average of 3.5 per question. The technical questions (irrelevant, symptomatic, sacrifice relevant) were asked 2,154 times and produced 7,484 reactions for an average of 3.5 per question. The pneumograph produced 19% of the reactions, the cardiograph 26%, and the electrodermal 55%.

Key words: cardiovascular, deception criteria, distribution of reactions, electrodermal responses, habituation, normative data, polygraph tracing features, pneumograph, tonic heart rate, zone comparison formats.

Over the last 75 years, lists of diagnostic polygraph tracing features have evolved from the observations of examiners in the conduct of countless field examinations. Polygraph schools over the years incorporated those observations into their curricula, developed scoring rules for them, and the instruction influenced the chart interpretations of generations of new polygraph students. While most of the instruction regarding the diagnostic features in polygraph tracings are shared among different schools, surprisingly little work has been done regarding the frequency and predictive value of those reaction criteria. We know from Capps & Ansley (1992) the types of polygraph tracing features examiners use in their analyses of the charts, but that study did not tell us what reactions were present in the tracings but not used. It would be of interest to explore the incidence of polygraph features in field cases, separate from their diagnostic use.

The US Department of Defense Polygraph Institute teaches that there are 23 specific features in polygraph tracings that are used in numerical analysis. Twelve of the features are found in the two pneumograph channels, three in the electrodermal channel, and eight in the cardiograph channel. These criteria were previously reported by Swinford (1999), and are reprinted here.

This article originally appeared in *Polygraph*, 2000, 29(2)

Acknowledgements

This article is based on a larger study funded by the US Department of Defense Polygraph Institute. The full study can be requested through the Defense Technical Information Center, 8726 John J. Kingman Road, STE 0944, Fort Belvoir, VA 22060-6218. The study was prepared under ONR Grant Number N00014-98-1-0863.

Respiration
R1. Rate decrease
R2. Rate increase
R3. Inhalation/exhalation ratio change
R4. Amplitude increase
R5. Amplitude decrease (suppression)
R6. Progressive increase or decrease
R7. Progressive increase and return
R8. Progressive decrease and return
R9. Baseline change – temporary
R10. Baseline change – permanent
R11. Apnea – holding (inhalation)
R12. Apnea – blocking (exhalation)

Electrodermal
E1. Amplitude change
E2. Complex response
E3. Response duration and return

Cardiovascular
C1. Baseline increase and decrease
C2. Baseline increase
C3. Baseline decrease
C4. Amplitude increase
C5. Amplitude decrease
C6. Rate increase
C7. Rate decrease
C8. Premature ventricle contractions

The purpose of the present paper is to look at the incidence of the DoDPI reaction criteria in field cases. In addition to generic normative data regarding the frequency of reactions, we wanted to know if the distribution of criteria differed between deceptive and nondeceptive cases. We were also interested in evidence of habituation of responding across successive charts, or within charts across questions. Finally, we wanted to know if the tonic heart rate of deceptive examinees was different from the heart rate of truthful examinees. Though most of the DoDPI reaction criteria have been used in field practice and various schools of instruction for about 50 years, reports of normative field data are sparse. A modest investigation of cardiograph responses was reported by Jensen (1981), and his results were compared to the present findings.

Method

Cases

A total of 177 polygraph cases were selected at random from the DoDPI database of confirmed cases by the junior author. All cases had been conducted in the field by federal or law enforcement polygraph examiners using Axciton computer polygraphs (Axciton Systems, Houston, TX). The only criteria for selection of cases were that they be identified as single-issue field zone comparison examinations. For the 161 cases where gender was identified, there were 115 men and 46 women. There were 111 deceptive and 66 nondeceptive cases. Of the 115 males, 71 were deceptive, and 44 nondeceptive. Among the 46 women, 31 were deceptive and 15 were nondeceptive. For the remaining 16 cases where gender was not recorded, 9 were deceptive and 7 were nondeceptive.

Human evaluator

The frequency counts were performed by the senior author, who has 49 years of polygraph experience. He was blind to ground truth and gender until the tabulations were complete.

Tabulation procedure

The list on Forensic Research, Inc. (FRI) Form 1 (Appendix A), and the definitions of scoring criteria used in this study are from the DoD Polygraph Institute. However, FRI Form 1 deleted the premature ventricle contraction (PVC) criterion because it is not generally deemed an autonomic response. Moreover, of the 5,866 question presentations in these cases, PVCs occurred 30 times, of which 18 were in one set of charts. Given the low incidence, they were not considered further.

An FRI Form 1 was made for each chart. The experienced examiner noted the presence of each of the criteria for each question presented on each chart. The data were then tabulated, and sorted for type of question, ground truth (deceptive or nondeceptive), gender, and polygraph channel.

Results

The 1,780 relevant questions produced 6,453 reactions, for an average of 3.6 per question. The 1,932 comparison questions produced 6,777 reactions, for an average of 3.5 per question. The 2,154 technical questions (irrelevant, sacrifice relevant, symptomatic) produced 7,484 reactions, for

an average of 3.5 per question. In terms of types of questions, 31% were relevant, 33% were comparison, and 36% were technical. Of the 20,714 reactions, 3,848 or 19% were from the pneumograph, 11,414 or 55% were from electrodermal, and 5,452 or 26% were from the cardiograph. It is interesting to note that the percentages shown here are similar to the weights give by some scoring algorithms.

The number of times each of the 22 reaction types appeared is on Table 1. Next to the number is the percentage of the total reactions the number represents. At the top is electrodermal amplitude change (E1) which appeared 4,793 times. E1 accounted for 26% of the total appearance of all reactions from all polygraph channels. At the bottom is cardiograph amplitude increase (C4) which appeared only five times, or less than one-half of one percent.

Table 1. Reaction criteria ranked by frequency.

Criterion	Description	Frequency	Percent
E1	Amplitude change	4793	26
E3	Duration	4496	24
C1	Baseline increase & decrease	2778	15
E2	Complex response	1051	6
C5	Amplitude decrease	940	5
R4	Amplitude increase	704	4
R9	Baseline change - temporary	683	4
C2	Baseline increase	578	3
R5	Amplitude decrease/suppression	476	3
C3	Baseline increase	400	2
RI0	Baseline change - Permanent	389	2
R1	Rate decrease	318	2
R8	Progressive decrease & return	265	1
R12	Apnea - (exhalation)	182	1
R2	Rate increase	154	1
R7	Progressive increase & return	107	1
R6	Progressive increase/decrease	102	1
R3	I/E Ratio change	62	less than .5%
C6	Rate increase	25	less than .5%
C7	Rate decrease	23	less than .5%
R11	Apnea - Holding (inspiration)	9	less than .5%
C4	Amplitude increase	5	less than .5%

Codes
P=pneumograph
E=electrodermal
C=cardiograph

Number = criterion number. See FRI Form 1 in Appendix A

In 1981 Carl W. Jensen published a study entitled "Frequency of occurrence of specific reaction criteria as observed in the cardio tracing." When the terminology of Jensen's study is matched with DoDPI's, and both data sets are ranked by frequency, the lists are strikingly similar (Table 3). This finding is reassuring for two reasons. First, the data from each study support the other, lending credibility to both. Second, Jensen's data were produced by analog instruments, and the present data were recorded digitally. The highly similar outcomes of the two studies suggest that the output signals from the two

recording instruments have much in common, and may alleviate concerns in some quarters that the cardiograph tracings of computer polygraphs are different in a significant way from the older analog instruments.

Table 2. Ranking of response criteria by frequency and deceptiveness status.

	All Cases		Nondeceptive Cases		Deceptive Cases	
Rank	Criterion	Percent	Criterion	Percent	Criterion	Percent
1	E1	26	E1	25	E1	26
2	E3	24	E3	24	E3	24
3	C1	15	C1	15	C1	15
4	E2	6	E2	6	C5	5
5	C5	5	R5	4	E2	4
6	R4	4	R4	3	R9	4
7	R9	4	R9	3	R4	4
8	C2	3	C2	3	C2	3
9	R5	3	C3	3	R53	3
10	C3	2	R5	3	C3	2
11	R10	2	R10	3	R10	2
12	R1	2	R8	2	R1	2
13	R8	1	R1	1	R8	1
14	R12	1	R7	1	R12	1
15	R2	1	R2	1	R2	1
16	R7	1	R6	1	R6	1
17	R6	1	R12	0	R3	0
18	R3	0	R3	0	R7	0
19	C6	0	C7	0	C7	0
20	C7	0	Rll	0	R4	0
21	R11	0	C4	0	C4	0
22	C4	0	C6	0	C6	0

Table 3.
Frequencies of cardiograph criteria for the present data, and the Jensen study (1981).

Present Study		Jensen's Study	
Criteria	Frequency	Criteria	Frequency
Baseline increase & decrease	2778	Baseline increase and decrease	363
Pulse amplitude decrease	940	Pulse amplitude decrease	326
Baseline increase	578	Baseline increase	172
Baseline decrease	400	Pulse amplitude increase	52
Pulse rate increase	25	Baseline decrease	48
Pulse rate decrease	23	Pulse rate increase	43
Pulse amplitude increase	5	Pulse rate decrease	20

We looked at the serial effects of questions by deceptiveness and non-deceptiveness. See Tables 4, 5, and 6 for frequencies and proportions of reactions in the cardiograph, electrodermal, and respiratory channels, and the summary on Table 7. There was a consistent reduction of reactions in the second and third charts of nondeceptive examinees in all three channels compared to the first chart. However, the respiratory pattern showed an increase in reactions in successive charts of deceptive examinees. The unusual effect can be seen in eight of the ten questions on Table 6. The underlying cause of this anomaly warrants further study.

Table 4. Change in the frequency of DoDPI diagnostic features in the cardiograph by question and by chart.

	Deceptive Cases (n=111)					Nondeceptive Cases (n=66)				
	Chart 1	Chart 2	change	Chart 3	total change	Chart 1	Chart 2	change	Chart 3	total change
Question										
1	105	106	1	77	-28	61	71	10	55	-6
2	114	112	-2	99	-15	67	64	-3	56	-11
3	103	101	-2	98	-5	83	72	-11	50	-33
4	98	119	21	122	24	77	63	-14	50	-27
5	136	129	-7	101	-35	68	62	-6	63	-5
6	96	103	7	105	9	72	65	-7	60	-12
7	134	113	-21	123	-11	74	59	-15	40	-34
8	79	87	8	83	4	61	50	-11	44	-17
9	105	93	-12	76	-29	59	133	74	51	-8
10	106	100	-6	100	-6	59	28	-31	26	-33
Total	1076	1063	-13	984	-92	681	667	-14	495	-186

Table 5. Change in the frequency of DoDPI diagnostic features in the electrodermal channel by question and by chart.

	Deceptive Cases (n=111)					Nondeceptive Cases (n=66)				
	Chart 1	Chart 2	change	Chart 3	total change	Chart 1	Chart 2	change	Chart 3	total change
Question										
1	222	210	-12	191	-31	136	136	0	120	-16
2	234	213	-21	216	-18	155	137	-18	120	-35
3	231	221	-10	190	-41	172	131	-41	94	-78
4	185	199	14	193	8	141	129	-12	121	-20
5	223	229	6	224	1	155	125	-30	141	-14
6	247	215	-32	209	-38	180	136	-44	132	-48
7	234	230	-4	214	-20	163	120	-43	111	-52
8	203	190	-13	192	-11	124	110	-14	97	-27
9	225	189	-36	149	-76	136	175	39	114	-22
10	217	185	-32	182	-35	128	60	-68	48	-80
Total	2221	2081	-140	1960	-261	1490	1259	-231	1098	-392

Table 6. Change in the frequency of DoDPI diagnostic features in the pneumograph channel by question and by chart.

	Deceptive Cases (n=111)					Nondeceptive Cases (n=66)				
	Chart 1	Chart 2	change	Chart 3	total change	Chart 1	Chart 2	change	Chart 3	total change
Question										
1	56	54	-2	60	4	38	41	3	38	0
2	63	73	10	77	14	46	42	-4	47	1
3	77	86	9	89	12	48	33	-15	40	-8
4	59	86	17	92	13	56	43	-13	38	-18
5	82	84	2	97	15	47	38	-9	37	-10
6	81	88	7	91	10	50	44	-6	34	-16
7	75	87	12	93	18	47	32	-15	38	-9
8	76	78	2	83	7	41	29	-12	33	-8
9	73	70	-3	59	-14	34	7	-27	37	3
10	83	55	-28	61	-22	41	26	-15	22	-19
Total	725	761	26	802	57	448	335	-113	364	-84

Table 7. Reaction totals by channel and chart.

	Deceptive Cases (n=111)			Nondeceptive Cases (n=66)		
	Chart 1	Chart 2	Chart 3	Chart 1	Chart 2	Chart 3
Respiration	745	761	802	448	335	364
Electrodermal	2221	2081	1960	1490	1259	1098
Cardiovascular	1076	1063	984	681	667	495
Average	1347.3	1301.7	1248.7	873.0	753.7	652.3

Tonic Heart Rates

From Table 8, we see average tonic heart rates of deceptive examinees were faster than the tonic rates of nondeceptive examinees, and the difference was significant ($z=2.87$, $p<.05$). The pattern held true for men and women, and at the beginning and the end of charts.

Table 8. Heart beats per minute for men and women during polygraph testing.

	Deceptive		Nondeceptive	
	Men	Women	Men	Women
Beginning	89	98	84	91
End	88	97	84	91

Discussion

The principal purpose of this study was to develop normative data for the DoDPI evaluative criteria in field polygraph charts. Two findings of the study are worthy of special note and comment. First, when the evaluative criteria are placed in rank order by frequency of appearance, it is apparent that some appear very rarely. Given the low incidence of some criteria, the present writers suggest that the list of evaluative criteria could be shortened to some extent without hampering day-to-day chart interpretation. For example, amplitude increase (C4) could be combined with amplitude decrease (C5) as simply cardiograph amplitude change. The criteria of pulse rate increase (C6) and rate decrease (C7) could be dropped entirely, as they each constituted less than one-half of one percent of all responses. However, if instrument manufacturers would add a cardiotachometer as an optional feature, these criteria might have utility. Some of the automated algorithms do make good use of this measure. The present findings with respect to pulse rate may simply reflect the difficulty in discerning subtle rate changes with the instrumentation used here. With regard to the respiration channel, experience with older instruments suggests that the inhalation/exhalation ratio (R3) might be more prevalent than was evident with these digitized instruments. If computer instruments do not manifest more inhalation-exhalation ratio changes than appeared here, the criterion might be considered for deletion. Apnea -holding (R11) could be combined with apnea - blocking (R12), as just apnea. One more combination of respiration criteria would make sense; merge baseline change - temporary (R9) with baseline change – permanent (R10). The differences between the two are not always clear, and the distinction does not appear to add to the probative value of the test.

The second noteworthy finding was the shrinking number of reactions across successive charts, suggesting the influence of generalized habituation. This was not an unexpected finding. However, we did not anticipate the increasing number of respiration reactions over charts that occurred exclusively with deceptive cases. The reason that respiration responses for deceptive examinees ran counter to the trend of habituation found for all other channels for both deceptive and nondeceptive examinees is beyond the scope of this study. Moreover, such a pattern would not be predicted from the published literature on polygraphy. If the finding is confirmed in other research, it may point to an unresolved area in polygraph theory.

The normative data in this paper are a small part of a study conducted by the first author for DoDPI. Those interested in the complete report should contact the Defense Technical Information Center, 8726 John J. Kingman Road, STE 0944, Fort Belvoir, VA 22060-6218. The study was prepared under ONR Grant Number N00014-98-1-0863.

References

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

Jensen, C.W. (1981). Frequency of occurrence of specific reaction criteria as observed in the cardio tracing. *Academy Journal,* 4(2), 5-7.

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

Appendix A

Data Evaluation Form 1 Forensic Research, Inc. 1998

File Name:_____ Chart____of____ Test Format:_____

Heart Rate – Beginning:_____ End:_____ PVC's useful? Yes No Other:_____

	1	2	3	4	5	6	7	8	9	10	T	TC	TR
RESPIRATION													
1. Rate Decrease													
2. Rate Increase													
3. I/E Ratio Change													
4. Amplitude Increase													
5. Ampltd Decrease/Suppression													
6. Progressive Increase/Decrease													
7. Progressive Increase & Return													
8. Progressive Decrease & Return													
9. Baseline Change – Temporary													
10. Baseline Change – Permanent													
11. Apnea – Holding (Inspiration)													
12. Apnea – Blocking (Exhalation)													
ELECTRODERMAL													
1. Amplitude Change													
2. Complex Response													
3. Response Duration & Return													
CARDIOVASCULAR													
1. Baseline Increase & Decrease													
2. Baseline Increase													
3. Baseline Decrease													
4. Amplitude Increase													
5. Amplitude Decrease													
6. Rate Increase													
7. Rate Decrease													
8. P.V.C.													

Gender: Male Female Race: Caucasian African American Hispanic Asian Age:_____

Comments:_____

The Irrelevant Question: A Descriptive Review

Norman Ansley

Abstract

This paper reviews the literature on irrelevant questions, using excerpts to provide definitions, describe their functions, placement in formats, and research.

Keywords: Control/comparison tests, definitions, irrelevant questions, norms, neutral questions, Relevant-Irrelevant Tests, test formats

Definitions

Krapohl and Sturm (1997) in their terminology reference define the irrelevant question as follows:

> A question designed to be non-emotion provoking. Irrelevant questions are most often placed in the first position of a question list, since an orienting response of no diagnostic value usually follows the presentation of the first question. In CQT formats it is also used after a relevant or control/comparison question that has elicited a strong response so as to permit physiologic arousal levels to return to baseline before presenting another diagnostic question. Irrelevant questions are used in every type of PDD test. Also called norms or neutrals.

The Department of Defense Polygraph Institute describes the irrelevant question in a 1993 handout on test question construction as:

> 1. An irrelevant question is the first question asked during a polygraph examination. It is designed to allow the orienting response to habituate. It is a neutral question unrelated to the testing issue. It is worded so the examinee answers "Yes."
>
> 2. Irrelevant questions can be used to establish a norm when some type of distortion is occurring on the chart.

Stan Abrams, in his *Complete Polygraph Handbook* (1989), in the glossary, describes an irrelevant question as:

> A neutral question developed to bring the subject's level of reactivity down after arousal or placed in a position such as first on a test when a reaction would occur because of its position rather than the question itself.

James Allan Matte has a section defining terms in his textbook (1996):

> Neutral (Irrelevant) Question: It is of a non-stimulating nature. In a Zone Comparison test format, this type of question is usually used as the first question in the test, to absorb an examinee's orienting response and reduce general nervous tension. The examinee's place of birth, last name or first name is usually used for that purpose. In other techniques it is also used to create a "norm' pattern at the beginning and end of each chart. It is also used between relevant questions where necessary to terminate lingering reactions due to extraneous stimuli or extended thought process.

In their 1977 text, *Truth and Deception*, second edition, John Reid and Fred Inbau say:

> The primary purpose of the irrelevant questions is to ascertain the subject's normal reactions – his "norm" – under

This article originally appeared in *Polygraph*, 1998, 27(4)

test conditions. Additional reasons for using irrelevant questions are:

1. To terminate a "lingering" type of reaction...

2. To nullify or terminate shock reactions due to noises occurring outside the examining room and heard by the subject during the test.

3. To provide an outlet for a relief response after relevant questions...

4. To separate the reactions to pertinent questions by using an irrelevant question in between.

5. To accentuate a deceptive response to a relevant question by the subject's failure to answer or by giving a delayed answer to the following irrelevant question...

6. To invite the scheming type subject to cause false or fraudulent reactions on irrelevant questions for the purpose of misleading the examiner into believing these reactions are greater, by comparison, than the reactions to the relevant test questions.

Fred Inbau, writing on the lie detector technique in 1942, before his adoption of control questions, wrote:

> The first two or three questions should be irrelevant to the matter under investigation. Questions such as "Is your name John Smith?" "Do you live in Chicago?" are helpful in ascertaining the nature of the subject's reactions to the test situation alone, irrespective of possible deception...

Fred Inbau gave the same definition and examples in his 1948 book, which included the control question test called the "Revised Questioning Technique." In the thirdedition of Fred Inbau's book in 1953, now listing John E. Reid as co-author, they state:

> The primary purpose of irrelevant questions ... is to ascertain the subject's norm under test conditions.

In a 1976 lesson plan on the relevant-irrelevant General Question Test the Army said about the irrelevants that they are used to determine the examinee's norm plus excitement level with a verbal stimulus.

Weir (1974), writing about the Relevant-Irrelevant Test, had a paragraph on terminology. He described the irrelevant question as:

> A polygraph question, of supposedly neutral impact, which does not relate to the matter under inquiry. Frequently called 'norm' questions or 'neutral' questions.

Harrelson (1964), writing about the relevant-irrelevant Keeler Technique, defines the irrelevant question as "A question formulated from information about which the subject would not normally lie, which does not pertain to the matter under investigation, and to which the examiner knows the answer or can reasonably infer same."

Among the items in the descriptions we note that an irrelevant question is:

1. Not emotion provoking;

2. Not related to the matter under inquiry; and

3. Is not a question the subject would normally lie to.

We note that some of the purposes of an irrelevant question are to:

1. Allow the orienting response to subside;

2. Allow other reactions to return to baseline;

3. Reduce general nervous tension;

4. Establish a norm or baseline; or norm plus excitement level;

5. Provide an outlet for relief from a response;

6. Separate reactions to relevant questions; and

7. Invite countermeasures.

Formats and Usage

Of the widely used techniques, the following employ an irrelevant question as the first question: Backster Zone Comparison (Backster, 1969), DoDPI Zone Comparison (DoDPI, 1991), DoDPI MGQT (DoDPI, 1989), Reid CQT (Reid & Inbau 1977), Relevant-Irrelevant (Harrelson, 1964; Minor, 1989; Weir, 1974), PCQT (Forman & McCauley, 1986), Arther CQT (Arther, 1987), and Marcy CQT (Matte 1996). In 1939 the Fordham Law Review published a paper by the late Reverend Walter G. Summers. His technique, used in research and criminal cases, paired his emotional standards questions (control/ comparison) with significant (relevant) questions, three of each on each chart, with the emotional standard in front of the significant in each pair. Summers wrote "These are interspersed among a larger number of nonsignificant questions, as, Are you wearing a black coat? and Did you eat breakfast this morning?" (Summers, 1939). His technique was used for many years by the New York State Troopers (Kirwan, 1952).

The Summers test format was a CQT, essentially a zone comparison. The number and placement of irrelevants was at the examiner's discretion. The relevant-irrelevant Keeler test also called for irrelevant questions as needed. Consider this question sequence by Leonarde Keeler who was testing Alfred de Marigny, acquitted of murdering Sir Harry Oakes.

1. "Is your name Alfred de Marigny?" "Yes."

2. "Do you live in Nassau?" "Yes."

3. "Do you know who killed Sir Harry Oakes?" "No."

4. "Have you had something to eat today?" "Yes."

5. "Did you kill Sir Harry Oakes yourself?" "No."

6. "Were you born in Nassau?" "No."

7. "Did you put your hand on that screen?" No. "

Keeler said there was no evidence of lying. Four of the seven questions were irrelevant, including the opening two. Note that irrelevant question number 6 was answered "No." This account was published in a book by Marshal Houts (1972), a friend of Keeler and no stranger to polygraph testing. A British author adds three relevant questions, lengthens the last one, but does not mention any irrelevant questions (Lessor, 1983).

Unlike the Summers and Keeler tests where irrelevants are placed as needed, fixed format tests put irrelevant questions where they expect they will be needed. An example is the DoDPI MGQT (1989):

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

The mixed series for the MGQT third chart is: 4-1-5-6-3-10-9-6-8-10. The MGQT test question sequence is based on the Reid technique. While three irrelevants are in the first two charts, only two are in the mixed series, both at the beginning.

The U.S. Air Force OSI MGQT differs from the DoDPI format in several ways. In fact, the OSI format looks more like a zone comparison than a MGQT. When Michael Koll (1987) lectured on the OSI, his handout showed several samples from cases, and each one had only one irrelevant question, and that in the first position. Each of his examples was "Is your name_____?"

Prominent among the standard polygraph test formats is the Backster Zone Comparison Test. In his 1969 notepack, Backster lists the following three irrelevant questions for use in his you phase test. The Backster Exploratory Test lists only 13 (j) as the first question.

1. 14 (j) "Were you born in the United States?"

2. 13 (j) last name

3. 13 (k) first name

In 1993 Backster appeared on a panel on question formulation at an APA seminar (Smith, 1993). The prepared examples in the handout used one of the following three to begin his you Phase test:

1. 13. (a) "Is your first name_____?"

2. 13. (b) "Is your middle initial_____?"

3. 13. (c) "Is your last name_____?"

One of these three irrelevant questions was at the beginning of each Backster Zone Comparison S-K-Y test and each Backster Zone Comparison Exploratory Test.

Richard 0. Arther has usually had an irrelevant question in the number one position. However, in 1961 he experimented with removal of the irrelevant question. He said he soon realized that was a mistake and that there must be an irrelevant as the first question. Although he comments that several irrelevant questions is an unnecessary feature, in 1985 he added the irrelevant question "Do you live in Canada?" as the last question, and it is answered "No" while the first question "Do you live in the United States?" is answered "Yes." (Arther, 1987). Matte (1996), in describing the Marcy Control Question Test format of 13 questions, notes irrelevants in positions 1, 4, 8, and 13, and like Arther, the last irrelevant is answered "No."

William M. Marston (1938) in describing a continuous type of test, had a most unusual placement of the irrelevant questions. After a short test in which apparently no questions were asked, a second record was made to "establish a record of b.p. behavior during irrelevant questions and answers." Marston said the examinee may be asked "Have you been in this room more than one minute?" "Have you drunk any liquor since you came into this room?" and so on. The relevant questions were all in the next chart. Marston added that some experts prefer to mix innocent and crucial questions, interspersing them in unpredictable order: "Have you smoked since coming into this room?" "Did you break into Mrs. Trimwell's apartment?" "What is your first name?" "Did you steal the diamond ring?" and so on. Note that one of Marston's irrelevant questions asked the examinee to answer with his first name.

The Canadian Police College student polygraph manual lists three irrelevant questions for use in their control question test. They are: "Is your name_____?" "Do you live in_____?" and "Were you born in_____?" In their ten-question zone format, the irrelevant questions are placed at questions one and eight (Koppang, 1985).

Harrelson (1964), writing about the relevant-irrelevant Keeler technique, ascribes four uses for irrelevant questions: to reduce the excitement level; to assist in returning the tracing to or toward the proper baseline in preference to a mechanical adjustment; to serve as an aid to interpretation of specific reactions; and to conserve the subject's ability to react. Harrelson gives ten examples of irrelevant questions:

1. "Is your first name_____?"

2. "Do you live in_____?"

3. "Do you drink_____?"

4. "Are you married?"

5. "Have you had anything to eat today?"

6. "Do you smoke?"

7. "Are you a citizen of the United States?"

8. "Were you born in_____?"

9. "Are you wearing_____?"

10. "Are you_____ years old?"

This is a mix of identification and obvious irrelevant questions. Harrelson has some observations about what might cause specific reactions to these questions. He cautions not to use the topics of race, religion, politics, any topic subject to opinion or interpretation, or a condition subject to change like the weather. Harrelson encourages examiners to determine the cause of reactions to irrelevant questions.

Clarence D. Lee, an authority on relevant-irrelevant testing, wrote two texts on the topic (1943, 1953), but said little about irrelevant questions. His sample formats in both publications show irrelevant questions only as the first and second questions, and not elsewhere. They were worded:

1. "Is your name_____?"

2. "Do you live in_____?"

Weir (1974), writing about the relevant-irrelevant technique, gives four examples of suitable irrelevant questions that relate to identification and appear relevant to the examinee:

1. "Is your first name_____?"

2. "Is your middle name_____?"

3. "Were you born in the month of February?"

4. "Do you live in the City of Boston?"

Consistent, significant reactions to one or more of the irrelevant questions results in interrogation, the reactions making the question relevant. Obvious irrelevants such as "Do you smoke?" or "Are you now in the City of St. Louis?" are not used. Weir observes that these questions appear ridiculous, seem like a game, and do not pose a threat to the subject.

Paul K. Minor, writing about a modified relevant-irrelevant technique in 1989, offered examples from some of his cases. In a 12-question test he opened with two irrelevants, and had a third one at position 11. In a 13-question test he had irrelevants in positions 1, 2, 5, 8, and 12. In a 12-question test Minor had irrelevant questions in positions 1, 2, 4, 8 and 11. Some of his irrelevant questions were identification types about citizenship, residence, age, name and employment. Others were obvious, such as "Are you now in _____?" or "Is today Tuesday?" (Minor, 1989).

The Army General Question Test (GQT) began in 1951 as an RI format. It remained as such for many years, but was supplemented then replaced by control question tests, notably zone and MGQT. An Army lesson plan of 1976 states the correct terminology for this technique as any of the following: 1. Relevant-Irrelevant Technique, 2. General Question Technique, or, 3. the Keeler Technique. A Master Question List gave ten irrelevant questions, of which four were of the obvious type. The number and location of irrelevant questions in the RI format was not given. On a date unknown, the Army changed the GQT to a control question test employing disguised controls in positions 3 and 9, and irrelevant questions in positions 1, 2, 6, and 10. (USAMPS 1976, Crowe, Chimarys & Schwartz, 1995; Matte, 1996)

Formats and Usage Summary

Among the items presented in Formats and Usage, we note that:

1. All standard control question and relevant-irrelevant test formats open with an irrelevant question.

2. Some formats open with two irrelevant questions.

3. Some techniques leave the placement of irrelevants to the examiner.

4. Some fixed-sequence formats anticipate the need in placing irrelevants.

5. Two formats have irrelevants in the first and last position, and the irrelevant questions in the last position are answered "No."

6. Most irrelevant questions are designed to be answered "Yes."

7. There are two types of irrelevant questions, identity and obvious.

8. Indications of deception to identity questions result in an interrogation.

9. In some relevant-irrelevant formats the irrelevants are comparison questions.

Research

Kircher and Raskin (1986) were interested in how examinees perceived their reactions to control, relevant, and irrelevant questions. One hundred men served as subjects, with 50 stealing a ring in a mock theft. All examinees were promised $25.00 if they could produce a truthful outcome. After the test, each subject was presented with a sequence of all possible pairings of the questions and asked to choose from each pair the question he felt had produced the largest physiological response. Programmed guilty subjects reported having reacted most strongly to relevant questions; programmed innocent subjects reported that the control questions produced their strongest reactions; and both groups reported that irrelevant questions produced their weakest reactions.

Bob Roy Frisby (1979) was interested in whether the obvious or the identity irrelevants produced the least reactions. His subjects were 24 men and 24 women in police classes at Washington State University. He asked the same questions on each of two charts but with a different sequence. Each series contained half obvious and half identity questions. The first series asked about: first name, city of birth, smoking, year of birth, driving a car, breakfast today, campus address, and color of hair. The identity irrelevant questions produced fewer responses, 604, than did the obvious irrelevant questions, which produced 732.

Research Summary

From these two research projects we note that:

1. Examinees perceived their reactions to irrelevant questions were weaker than their reactions to control and relevant questions.

2. Identity irrelevant questions produce fewer responses than obvious irrelevant questions.

Conclusion

An irrelevant question is the opening question on each chart in all standard control/comparison and relevant-irrelevant test formats. It is there because the reaction caused by the opening question is not scored. Some formats include irrelevant questions within the pre-ordered list of questions. Other techniques put them in as needed. Within the test format they separate reactions to relevant questions, provide relief from relevant and control questions, and reestablish baselines. In some techniques irrelevant questions are also comparison questions. The answers to obvious irrelevant questions are almost always true, and the base rate of lying to identity irrelevants is low. The identity irrelevants are introduced as relevant, and if there are consistent significant reactions, the questions become relevant. Most irrelevant questions are worded to be truthfully answered "Yes," but some are worded to be truthfully answered "No." The significance of the difference, if any, is unknown.

With Weir's argument that all questions on a test should appear relevant, and the result of Frisby's research showing fewer reactions to identity irrelevant questions, the use of identity irrelevants in place of obvious irrelevants may be justified. Given the important functions of the irrelevant question, it is surprising that it has received such little serious attention.

References

Abrams, S. (1989). *The complete polygraph handbook*. Lexington, MA: Lexington Books.

Abrams, S. (1977). *A polygraph handbook for attorneys*. Lexington, MA: Lexington Books.

Arther, R. 0. (1987). Irrelevant questions. *Journal of Police Science*, 22(2) 1-4.

Backster, C. (1969). *Standardization notepack and technique guide, Backster Zone Comparison guide*, Rev. ed., San Diego, CA: Backster Associates.

Crowe, M. J., Chimarys, M. & Schwartz, J. (1995). The GQT polygraph test: Scoring and validity. *Polygraph*, 24(3) 234-236

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

Department of Defense Polygraph Institute (1993). *Test Question Construction*.

Department of Defense Polygraph Institute. (1991). *Zone Comparison Technique lesson plan*.

Forman, R. F. & McCauley, C. (1986). Validity of the positive control polygraph test using the field practice model. *Journal of Applied Psychology*, 71(4) 691-698. Reprinted in *Polygraph*, 16(2) 145-160.

Harrelson, L. H. (1964). *Keeler Polygraph Institute training guide*. Chicago, IL: The Keeler Polygraph Institute.

Houts, M. (1972). *King's X, common law and the death of Sir Harry Oakes*. New York, NY: William Morrow & Co.

Inbau, F. E. (1942). *Lie detection and criminal interrogation*. Baltimore, MD: Williams & Wilkins.

Inbau, F. E. (1948). *Lie detection and criminal interrogation*, second edition, rev., Baltimore, MD: Williams & Wilkins.

Inbau, F. E. (1953). *Lie detection and criminal interrogation*, third edition, rev., Baltimore, MD: Williams & Wilkins.

Kirwan, W. E. (1952), Letter to N. Ansley, October 16, on the letterhead of the New York Troopers Scientific Laboratory.

Koll, M. (1987). Question formulation. Handout at a meeting of the Wisconsin Polygraph Association.

Koppang, C. E. (1985). *Polygraph examination student manual.* Ottawa: Canadian Police College.

Krapohl, D. & Sturm, S. (1997). *Terminology reference for the science of psychophysiological detection of deception*. Chattanooga, TN: American Polygraph Association.

Leasor, J. (1983). *Who killed Sir Harry Oakes?*, Boston, MA: Houghton Mifflin Co.

Lee, C. D. (1943). *Instruction manual for the Berkeley psychograph*. Sacramento, CA: Lee & Sons.

Lee, C. D. (1953). *The instrumental detection of deception, the lie test*. Springfield, IL: Charles C. Thomas.

Marston, W. M. (1938). *The lie detector test*. New York, NY: Richard R. Smith. Reprinted by the American Polygraph Association, 1989.

Matte, J. A. (1996). *Forensic psychophysiology using a polygraph*. Williamsville, NY: J.A.M Publications.

Minor, P. K. (1989) "The relevant-irrelevant technique," in S. Abrams *The complete polygraph handbook*. Lexington, MA: Lexington Books.

Reid, J. E. & Inbau, F. E. (1966). *Truth and deception: The polygraph ("lie detector") technique*. Baltimore, MD: Williams & Wilkins.

Reid, J. E. & Inbau, F. E. (1977). *Truth and deception: The polygraph ("lie detector") technique*, second edition, rev. Baltimore, MD: Williams & Wilkins.

Smith, M. L. (moderator) (1993). "Polygraph test question formulation...". A panel discussion. Handout at the Seminar of the American Polygraph Association, Newport Beach, CA.

Summers, W. G. (1939). Science can get the confession. *Fordham Law Review*, 8, 335-354.

United States Army Military Police School (July 1976). Lesson plan "General Question Test (GQT) Construction."

Weir, R. J., Jr. (1974). In defense of the relevant-irrelevant polygraph test. *Polygraph*, 3(2) 119166.

Question Formulation

Norman Ansley

Abstract

This paper contains observations about question formulation for polygraph testing followed by specific guidelines. Applicable to the frequently used testing formats, the guidelines cover relevant questions, probable lie control/comparison questions, irrelevant questions, and wording of peak of tension and guilty knowledge tests. The paper does not offer guidelines for technical questions used by only one test format. There are abstracts of three legal cases where question formulation was an issue. The references cited are included in a larger reference section.

Keywords: Control question, comparison question, guilty knowledge test, irrelevant question, peak of tension, question formulation, relevant question, semantics.

General Observations

One of the clinical aspects of polygraph testing is the formulation of questions. Some guidelines are suggested in this paper.

A word of caution about prepared question lists and notepacks. Blind adherence may result in the examinee not understanding one or more questions, causing problems in testing. Written questions are a good guide to policy, but the suggested words may not be in the examinee's vocabulary. The advantage of prepared lists and notepacks is better compliance with policy, regulations and law.

When working with investigators or attorneys who do not understand the limits of polygraph testing, you may be presented with a lengthy list of poorly worded questions that you cannot use. Ask them to describe the case and suggest one question, answered "yes" or "no" that will solve it. Try to conduct your tests with a single-issue test format, as they tend to be more accurate than multiple issue formats. More issues create more opportunities for error.

Be wary of precisely worded relevant questions proffered by the examinee or his attorney. The question may avoid the issue or be part of an effort to rationalize.

An examinee will not readily admit he does not understand a question. The lack of understanding shows up when the examinee is asked to explain why the question is being asked and what it means.

When the questions are agreed upon, and they exclude details or the wording is a bit unusual, be sure the missing details and a discussion of the development of the relevant questions are in the report. Details that were agreed upon, but were deleted from the question, must be in the report. Persons who were not present may criticize the relevant question wording because the report does not adequately describe the question development.

In screening applicants, keep in mind that EEOC and ADA rules on job interviewing apply to polygraph testing. For example, under ADA you cannot ask medical questions until a bona fide offer of employment is made, and the questions you usually ask to determine fitness for testing are considered medical, you either don't ask, or have the polygraph tests performed after the offer stage. In addition to the Federal limits, there are state laws and city ordinances that further limit what you can say.

The technical questions that are designed to appear as relevant questions must

This article was originally published in *Polygraph*, 1998, Volume 27(3).

be treated with the same thoroughness as the relevant questions. Included are the control/ comparison questions (except in PCQT and DLC), sacrifice relevants, and the identity irrelevants in some RI tests.

While keeping a question short is often desirable for clarity, it is not essential. I have seen long and complex questions used in contract fraud, and the tests were successful.

Some technical questions such as the sacrifice relevant (Capps, 1991; Horvath, 1994) and the symptomatic (Capps, Knill & Evans, 1993) have been the topic of specific papers. Much has been written about techniques and questions for disclosure and maintenance tests in sex offender tests. It is too early to comment on those questions or suggest guidelines. Wording of relevant and control/comparison questions in certain types of crimes suggest the need for expert advice. Examples are arson, bomb cases, contract fraud, and insider trading of stocks or commodities.

Relevant Questions – Guidelines

The relevant question must solve a vital problem.

The issue covered by the relevant question must be of vital importance to the examinee.

The question must pose a dichotomy, answerable by "yes" or "no."

The question must be fully understood and mean the same thing to the examiner and examinee.

When possible, a relevant question should not use legal or technical terms.

The question must not contain obscene, profane, racial, derogatory, degrading, or insulting words or phrases.

Qualifiers, such as "Other than ..." are placed at the beginning of the question.

There should be enough facts in the question to avoid outside issues.

There should be no more facts in the question than necessary.

The facts in the question should not only be correct, but would be recognized as correct by the perpetrator.

The question must not imply or assume guilt.

The question must not imply disbelief by the examiner.

The sentence must be a question (POT/ GKT exception).

It is preferable to use the action (verb) rather than the result.

The question must not ask for an opinion.

The question should not give away facts you plan to use in a POT/ GKT.

It is generally held that you cannot test on the issue of intent.

When testing victims, the issue is truthfulness, not rape, robbery, or some other crime.

Be wary of using specific amounts of money stolen in the question.

Avoid words that are emotional, and likely to cause a response.

Separate relevants are asked about direct involvement, secondary involvement, guilty knowledge, and evidence connecting facts.

Control/Comparison Questions

The final written descriptions of control/comparison questions are by Summers (1939) whose "emotional standards" questions were paired with relevant questions. Examples he gave included, "Were you ever arrested?" "Are you living with your wife?" and "Do you own a revolver?" From the text and examples it appears that Summers used probable lies, embarrassing, evidence connecting, and other questions. Inbau & Reid (1948) introduced a test format in the 1940s which

included a probable lie and a guilt complex question for comparison purposes. The guilt complex was later dropped for a second probable lie. The Reid control question may include the offense at issue. However, the Backster (1969) technique and DoDPI control/comparison question techniques do not permit the control/comparison questions to include the offense. They separate relevant and control/comparison coverage and offense by date or location.

The guilt complex question is a known-truth answer to what appears to be a relevant question about a crime. Other comparison questions include the yes answer to the relevant question in the PCQT format, a directed lie to a trivial matter, and the situational control where the examinee confirms an inculpatory fact with a yes answer. In one relevant-irrelevant screening test format, a relevant question with a low base rate of deception, such as terrorism, may serve as a probable truth (guilt complex) comparison question.

Probable Lie Control/Comparison Question Guidelines

The control/comparison question must be treated as a relevant question.

It is broader in scope than a relevant in order to be more likely applicable.

It is usually on the same topic as the case issue, but slightly lesser in severity of offense.

It should not mention or imply sex, except where sexual behavior is the issue.

Qualifiers such as "OT" from admissions should be at the beginning of the question.

Time bars should be used or not used, depending on the rules for the format.

It is usually worded to be answered "No."

When possible it should use the same verb as is in the relevant question.

The topic of the question should be one the examinee is likely to lie about or have serious doubt regarding the truthfulness of the reply.

Do not use a control/comparison question on race, religion, or politics, or that will humiliate or embarrass the examinee.

The question must be fully discussed with the examinee.

Irrelevant Questions

Almost all test formats open with an irrelevant question. Some formats anticipate additional need for an irrelevant question and fix its place in the format, while other techniques allow the examiner to insert them as needed. Irrelevant questions allow the orienting and other reactions to return to baseline, establish a norm level, reduce general nervous tension, provide relief from a previous reaction, separate reactions to relevant questions, and confirm the identity of the person being tested.

There are two types of irrelevant questions. One is the obvious irrelevant question, such as "Are you wearing brown shoes?" The other type of irrelevant involves identity questions, and is disguised as a relevant. These involve name, date and place of birth, residence, etc. Both of these types of irrelevant questions have a place in testing, and the type is sometimes prescribed. However, favoring the identity questions, Weir (1974) notes that the obvious irrelevants appear ridiculous, seem like a game, and do not pose a threat to the examinee.

Regarding research, Kircher and Raskin (1986) found that examinees were aware that the irrelevant questions produced their weakest reactions; and Frisby (1979) found that identity irrelevants produced fewer responses than did obvious irrelevants.

Irrelevant Questions – Guidelines

Identifying irrelevants are treated as relevants and thoroughly discussed.

Consistent significant reactions to identify irrelevants warrant interrogation.

Irrelevants must pose a dichotomy, answerable by "yes" or "no."

Irrelevants, obvious or identity, must be discussed with the examinee.

Answered truthfully, an irrelevant should not provoke emotions.

The proposed irrelevant is not a question you expect an examinee to lie to.

An obvious irrelevant is not related to the topic at issue.

Irrelevant questions are usually worded to be answered "yes." However, the Marcy and Arther CQT formats have obvious irrelevants answered "no." Several irrelevants must be reviewed before the test if the examiner is allowed by the technique to insert irrelevants as needed.

Most test formats open with one irrelevant, and some open with two.

Peak of Tension and Guilty Knowledge Tests

The peak of tension group includes the known solution peak (Type A) in which the investigator and the perpetrator know some specific item of information which would not be known to someone who was not involved in the offense. There is a variant called the guilty knowledge test (GKT). The primary difference between the POT and GKT is that in the latter the key item is placed by chance in the list anywhere except the first position, while in the POT the key is in or near the middle. There is a searching peak of tension (SPOT) in which the examiner is seeking to locate evidence from a subject who may possess information he refuses to divulge, such as the location of loot, or location of the victim of a kidnapping. The stim or acquaintance test is in the peak of tension group. There are many variations of the stim, with a number described in a special issue of Polygraph (1978) 7(3) 173-215. Stim tests differ from most POT formats in that the series is asked only once, where most POT, GKT, and SPOT tests employ three series, often varying the sequence in each presentation.

Known Solution (Type A) and Guilty Knowledge Test – Guidelines

Place the key item in or near the middle of the list. In the GKT the key is to be placed by chance, but not at the beginning.

Be certain the key is the correct item.

Be certain that other items in the list cannot possibly be correct.

Be certain the guilty or involved would recognize the correct item.

Be certain the innocent would not know the correct item.

Be certain that concealing recognition of the key is important.

Try to keep all the items of similar length, one word, two words, etc.

Try to keep all the items of similar emotional content.

Do not include an absurd or illogical item.

You may use a logical sequence to the items, if the key is not first.

Five, six, or seven items are ideal, but more may be used if logical.

The examinee may be given the order, or a list posted.

To avoid dissociation have the examinee repeat the item before saying "No."

If you plan to give a POT after an RI or CQT, be sure the key item(s) are not given away in the questions or pretest.

All items must be discussed with the examinee.

If the list includes guns or cars, be certain the examinee is sufficiently knowledgeable to recognize calibers, makes, and models.

If you use Arther's false key, place it at number 2, and the key at 4 or later.

Searching Peak of Tension Tests (Type B, SPOT)

The most probable item should be in the middle of the list during the first of three presentations.

The least probable item should be at the beginning of the list during the first of three presentations.

Use a question about other possibilities as the last item on each chart.

The order should be varied with each presentation.

The order of items may be announced or posted.

Concealing the correct item must pose an obvious threat.

The items in the list should be discussed in detail.

When maps or diagrams are used, they must have clearly marked boundaries, numbers, letters, and names for each area.

Question Formulation - Legal Opinions

In *United States v. Lech*, 94 Cr. 285, 895 F.Supp. 582 (USDC SD NY 1995), a bribery case before federal trial Judge Sonia Sotomayer, defendant Wlodek Jan Lech, attempted to enter into evidence the results of a polygraph examination in which he answered such questions as, "Did you try to bribe any Board of Education officials to obtain an asbestos removal contract?" and "Did you take part in trying to bribe Board of Education officials to obtain an asbestos removal contract?" Lech sought admissibility in light of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 113 S.Ct. 2786. Judge Sotomayer did not address *Daubert* and Rule 702. She applied Rule 403 of the Federal Rules of Evidence and found Lech's polygraph evidence precluded because "its probative value is substantially outweighed by the danger of unfair prejudice, confusion, or misleading of the jury." She explained that "Each of the questions Lech seeks to introduce calls for his belief about the legal implications of his actions, without setting forth the factual circumstances underlying such conclusion." In other words, she wrote, "the jury would receive evidence showing Lech's personal belief that he did not violate any federal criminal statute, but would not receive any information that would assist its inquiry to find facts." In a footnote, the Judge indicated the outcome may be different if a defendant sought to introduce answers "to an exam where he or she completely denied any connection or involvement" with the alleged crime. [*New York Law Journal*, 28 July 1995]

In *Hester v. Milledgeville*, 777 F.2d 1492 (11th Cir. 1985) the Eleventh Circuit overruled a trial court's conclusion that the use of control questions was a violation of the Constitutional right to privacy. The appellate court said the City's interest in using control questions to improve the accuracy of the polygraph test is an important one ... and the specific control questions at issue constituted only a limited intrusion into the sphere of confidentiality. The Court noted that the questions were general in nature, were asked for a specific, limited purpose, and, although potentially embarrassing, avoided issues such as those related to marriage, family and sexual relations generally considered to be the most personal. The Eleventh Circuit issued a word of caution, saying they would have reservations if any governmental unit were to use a subject's response to a control question for any purpose other than comparing the polygraph reading for the control question to the same subject's reaction to a relevant question. The Court added there might well be a point at which a control question is so embarrassing or specific, or concerns so personal a matter, as to render the question unconstitutional even when asked for the proper purpose.

In *State v. Stowers*, 580 S.W.2d 516 (Mo.App. 1979) the defendant was appealing conviction for forcible rape. The results of a stipulated polygraph examination had been admitted, and on appeal defendant said one of the questions asked during the test was factually inaccurate. The question at issue was "Did you rape ... on Route FF?" Defense stated that the prosecutrix testified that the rape was along a gravel road just off Route FF, and that inaccuracy should cast doubt over the reliability of the whole test, thus rendering it inadmissible. The Missouri Court of

Appeals said that the reference to geographic area was sufficiently proximate to the crime site not to invalidate the test results.

References

Abrams, S. *The complete polygraph handbook*. Lexington, MA: Lexington Books, 1989.

Arther, R.O. (1982). How to word peak questions. *Journal of Polygraph Science*, 17 (2), 1-4.

Arther, R.O. (1987). Irrelevant questions. *Journal of Polygraph Science*, 22 (2), 1-4.

Backster, C. (1996). The Backster zone comparison technique, an update from the source. Paper presented at the American Polygraph Association Seminar, New Orleans, LA.

Backster, C. (1974). Reducing inconclusive results by pre-testing relevant questions. *Polygraph*, 3 (2), 216-222.

Backster, C. (1969). Standardized polygraph examination notepack, 5th ed. rev. San Diego, CA: Backster Associates.

Capps, M.H. (1991). Predictive value of the sacrifice relevant. *Polygraph*, 20 (1), 1-6.

Capps, M.H., Knill, B.L., & Evans, R.K. (1993). Effectiveness of the symptomatic questions. *Polygraph*, 22 (4), 285-298.

Department of Defense Polygraph Institute (1989). *Modified general question technique summary*.

Department of Defense Polygraph Institute (1991). *Technical correction of Curriculum [regarding the peak of tension test.]*

Department of Defense Polygraph Institute (1993). *Test question construction*.

Frisby, B.R. (1979). Research into the semantics of the irrelevant question. *Polygraph*, 8 (3), 205-223.

Harman, G.W. & Reid, J.E. (1982). The selection and phrasing of lie-detector test control questions. *Polygraph*, 11 (1), 82-86.

Harrelson, L.H. (1964). *Keeler Polygraph Institute training guide*. Chicago, IL: The Keeler Polygraph Institute.

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

Horvath, F. (1994). The value and effectiveness of the sacrifice relevant question: An empirical assessment. *Polygraph*, 23 (4), 261-279.

Inbau, F.E. (1942). *Lie detection and criminal interrogation*. Baltimore, MD: Williams &Wilkins.

Inbau, F.E. (1948). *Lie detection and criminal interrogation*, second ed., rev. Baltimore, MD: Williams & Wilkins.

Inbau, F.E. & Reid, J.E. (1953). *Lie detection and criminal interrogation*, third ed., rev. Baltimore, MD: Williams & Wilkins.

Kircher, J.C. & Raskin, D.C. (1986). Visceral perception in the detection of deception. SPR Abstracts, *Psychophysiology*, 23 (4), 446.

Koppang, C.E. (1985). *Polygraph examination student manual*. Ottawa: Canadian Police College.

Lee, C.D. (1953). *The instrumental detection of deception, the lie test*. Springfield, IL: Charles C Thomas.

Matte, J.A. (1996). *Forensic psychophysiology using the polygraph*. Williamsville, NY: J.A.M. Publications.

Minor, P.K. (1989). 'The relevant-irrelevant technique," chapter 10 in S. Abrams, *The complete polygraph handbook*. Lexington, MA: Lexington Books.

Reid, J.E. & Inbau, F.E. (1966). *Truth and deception: The polygraph (lie detector) technique*. Baltimore, MD: Williams & Wilkins.

Reid, J.E. & Inbau, F.E. (1977). *Truth and deception: The polygraph (lie detector) technique*, second ed., rev. Baltimore, MD: Williams & Wilkins.

Smith, M.L. (Moderator). (July 1993). Polygraph test question formulation ..." A panel discussion. Handout at the Seminar of the American Polygraph Association, Newport, CA.

Special issue on stimulation tests (1978). *Polygraph*, 7 (3), 173-214.

Summers, W.G. (1939). Science can get the confession. *Fordham Law Review*, 8, 335-354.

United States Army Military Police School. (July 1976). Lesson plan "General question test (GQT) construction.

Weir, R.J., Jr. (1974). In defense of the relevant-irrelevant polygraph test. *Polygraph*, 3 (2), 119-166.

Weir, R.J., Jr. (1976). Some principles of question selection and sequencing for relevant-irrelevant testing. *Polygraph*, 5 (3), 207-222.

Wilkerson, O.W. (1978). The peak of tension tests utilized in the Ashmore kidnapping. *Polygraph*, 7 (1), 16-20.

Wygant, J. (1980). Hypothetical controls. *Polygraph*, 9 (1), 45-48.

Wygant, J. (1979). The relevant-connected control. *Polygraph*, 8 (1), 64-67.

The History and Accuracy of Guilty Knowledge and Peak of Tension Tests

Norman Ansley

Introduction

There are several types of test formats that involve recording of physiological reactions while attempting to detect deception or support truth. One of the major classes of tests are those which detect concealed knowledge, known variously as guilty knowledge tests, peak of tension tests, stimulation (stim) tests, and concealed knowledge tests. Indeed, the once popular word association test is related in principle. Other test formats include varieties of control question tests and relevant/irrelevant tests.

Detection of Concealed Knowledge

Since the 1930's, polygraph examiners have used three versions of concealed knowledge tests with some frequency. They have used the peak of tension (POT) in which the solution or key item in the list is known to the examiner and perpetrators, but not to innocent subjects. In a second version of POT, called a searching peak, the examiner does not know the key word (name, location, amount) and presumes that a person involved in a crime does know, and by reactions will disclose the key. In a third version, a stimulation test, the examiner presents the subject with a simple number selection test and compares the subject's reaction to the chosen number to his lack of reaction to other numbers. In the POT structure a common practice is to display to the subject the list of choices and the sequence in which they will be asked to achieve a peaking effect of reactions at the key item, if deceptive, followed by patterns of relief. The POT practice also calls for putting the key item near the middle of the list. In the searching peak of tension test the examiner can only put the more likely item(s) in the middle of the list on the first presentation, and rotate positions so no item remains in the first position in more than one presentation, as that first item often gets an orienting response which is discounted. In the known solution POT, the first item is considered a buffer, and reactions to it are not considered in the analysis. The Guilty Knowledge Test (GKT) format is a test in which the key item is placed anywhere in the list, by chance, except in the first position because of the need for a buffer. The sequence of the items is unknown to the subject of the test. If the list is used more than once, or there is more than one list, the sequence for each list is varied by chance selection, excluding the buffer. The name Guilty Knowledge Test suggests a use in which there will be some emotional involvement by the subject. The term concealed information or concealed knowledge may be more appropriate for those tests and laboratory simulations where the subject is not seriously involved or concerned about the outcome.

All field examiners would probably refer to guilty knowledge and concealed knowledge formats as peak of tension (POT) tests, from long and frequent usage. However, Dr. Gershon Ben-Shakhar makes a good argument for referring to POT (known solution and searching) as special cases of the GKT which may be used for different purposes. I am inclined to agree with Ben-Shakhar's broad view that GKT "refers to a set of procedures which are constructed like a multiple choice test such that the one alternative (the relevant alternative) is related to a specific event (assumed to be known to any individual who participated in that event, or has knowledge of the event), where as all other alternatives (the control alternatives) are unrelated to the event, but are equivalent to the relevant one in the sense that an individual who has no knowledge of the event cannot discriminate between the relevant and the control alternatives (i.e., cannot guess at a better than chance rate which alternative is the relevant one)." (Ben-Shakhar, 1992).

This article was originally published in *Polygraph*, 1992, Volume 21(3).

This definition of GKT is better than the older term POT, because POT suggests a format in which we want the reaction patterns to behave in a predictable way. GKT can include the POT formats, and much more. Dr. David T. Lykken, who has popularized the GKT among scientists as the only sound test for the detection of deception, states, "I developed what I called the Guilty Knowledge Test as a young psychology professor who routinely used multiple-choice test questions both in the classroom and in constructing research instruments, personality questionnaires and the like. It was natural to think of using this same format to determine whether a subject possessed guilty knowledge, i.e., whether he could identify the correct alternative to several equally plausible alternative answers to questions about the crime. Since a guilty suspect would be unlikely to answer such questions truthfully, it was natural to think of letting his involuntary, autonomic nervous system answer for him." (Lykken, 1992).

Lykken's definition of the GKT is similar to Ben-Shakhar's. Lykken writes, "I consider a GKT to be any procedure that uses some involuntary physiological response to indicate whether the subject identifies the 'correct' or crime-related alternative as distinctive or different from a set of control alternatives that are not in fact crime-related but chosen to seem equally plausible to an innocent suspect. And the crucial thing about the procedure is that, in contrast with the CQT, the incorrect alternatives provide genuine controls in the scientific sense of that term. That is, the subject's mean response to the incorrect alternatives provide an estimate of how this person ought to react to the correct alternative if he is innocent and does not recognize the correct alternative as being crime-related." (Lykken, 1992). This definition can also include all of the present POT and stimulation formats.

Lykken notes that "the physiological variable used does not define the GKT," a view shared by Dr. John J. Furedy, who is also a proponent of GKT formats. Furedy has written, "it does not matter what involuntary response or responses are measured." (Furedy, 1992). Thus, field polygraph examiners may administer a GKT and utilize the three standard channels, some other autonomic variable, or even a CNS function such as evoked potentials. Furedy, after discussing the matter with his colleague Ben-Shakhar, describes a GKT in these terms: "In our opinion the GKT is the general form of the procedure where a set of questions are generated about which the innocent have no crime-related information and which, in terms of eliciting involuntary responses, are equivalent. For the guilty, the same set of questions has a subset of questions (usually a quarter or less than the total set) about which the suspect has information, and this (concealed) information is indexed by bigger responding to this subset of questions. So for the innocent, all questions are control questions, whereas for the guilty the crime-related subset are experimental, relevant, or critical questions." (Furedy, 1992). Furedy added two observations; one that the scoring system needs to be objective and the other that serial position differences need to be ruled out. He also noted the necessity for ruling out confounding factors such as the innocent subject obtaining relevant information without having been involved in the crime.

All of this suggests that the GKT is broad in definition, and includes all of what we now consider POT. We know of course that POT and GKT are test formats long in use, with reports of regular use going back to the 1930's and a suggestion of the test format by Munsterberg appearing as early as 1907. Indeed the formats of many of the tests we have called POT in the past do not create the peaking effect. Furedy (1992) calls the POT, "a special case of the GKT in which the position of the critical item is always central in the list." He notes the "assumption of an underlying continuum is much stronger than the general GKT case."

Although it makes sense to include the POT as a specific format within the broader GKT framework, no doubt the 'POT' will be used in reference to such tests for a generation, and will not disappear before 'electrodermal' replaces 'GSR,' if then. This paper, however, is about test formats. It is about their origin, similarities and dissimilarities, sequencing, scoring, and accuracy.

Origins of the GKT/POT

In 1904 Max Wertheimer and Julius Klein published a paper entitled, "Psychologische Tatbestandsdiagnostik" in which they said, "Isn't it possible to diagnose in a perpetrator the concealed knowledge of his criminal action independent of his statements?" (Tr. by Herbold-Wooten, 1982). By 1935 the followers of Wertheimer had developed elaborate word-association tests coupled with reaction time, but Wertheimer's views on tests in general remained relevant. Some of these word association tests also employed galvanometers or motor movement recordings. Wertheimer said that for identification of 'critical symptoms' two things are necessary: "a comparison with the reaction pattern of a non-involved person by identical experimental setting and ... a comparison of reaction patterns in the same person to critical and irrelevant stimuli." (Wertheimer, 1935). The specific concept and term 'guilty knowledge' was recognized by Crane (1919) who conducted research with word-association tests.

In 1907, at Harvard, Hugo Munsterberg wrote about the problem of testing the nervous innocent man and said the "real use of the experimental emotion method is therefore so far probably confined to those cases in which it is to be found out whether a suspected person knows anything about a certain place or man or thing. Thus if a new name, for instance, is brought in, the method is reliable; the innocent, who never heard the name before, will not be more excited if he hears that one among a dozen others; the criminal, who knows the name as that of a witness of the crime, will show the emotional symptoms." He added, "And yet, it may be rash to propose narrow limits for the practical use, as the rapid progress of experimental crimino-psychology may solve tomorrow those difficulties which seem still to stand in the way today." (Munsterberg, 1907). For an early application of this method, see Gina Lombroso Ferrero's biography of Cesare Lombroso (1911). In the same chapter on 'Traces of Emotion,' Munsterberg wrote about the case of the pneumograph, sphygmograph, galvanometer, and other measures of emotion. The equipment was there to build a modern polygraph. For illustrations of the apparatus see MacDonald (1905).

Development by Practitioners

In the 1920's and 1930's practical application of lie detection methods was limited to a few researchers and practitioners in universities and law enforcement agencies. In the East, Dr. William Moulton Marston, J.D., Ph.D., who studied under Munsterberg developed a technique for use in practical cases with the assistance of his wife Elizabeth Holloway Marston and Olive Richard. He was apparently acquainted with the principle of a GKT in describing an "elimination test." He said that in this examination "another series of critical questions may be asked, and another polygraphic record run. These questions are designed to reveal the testee's knowledge of other suspects connected with the case. For example, if the testee is known to be a member of a certain gang, and the examiner wishes to identify other members of the same mob, a series of this sort is asked: 'Was Jones with you on the night of the murder?', 'Was Smith with you?', 'Was Doe with you?', and so on. The testee in such cases usually answers 'no' to all of the identification questions, but his uncontrollable b.p. responses reveal which individuals were present in the murder gang. Other types of questions may be arranged similarly in groups, and further b.p. records may be taken as desired." (Marston, 1938). Today we would call that a searching peak of tension test. Note that more than one item (person) may be correct.

In 1936, Professor John E. Winter of West Virginia University successfully found the thief among 25 women who lived in a dormitory using a cardio-pneumo psychograph and a relevant/irrelevant test format. He also tested all the suspects with the word-association and reaction-time method but it produced one false positive and no useful results (Winter, 1936). Wertheimer's test was not often used in real cases, and Winter may be among the last to have used it in a criminal investigation.

On the West Coast lie detection development was the product of Chief August Vollmer of the Berkeley Police Department who directed the work of John A. Larson, C.D. Lee, and influenced the work of Leonarde Keeler and others. Larson, a patrolman in the Department was studying for his Ph.D. at the

University of California. His laboratory unit employed a pneumograph and Erlanger sphygmograph that recorded on a smoked drum kymograph. Used for several thousand cases in the 1920's, Larson developed technique and analytic methods. A young associate of Vollmer's was Leonarde Keeler who developed a portable polygraph which recorded with ink on a paper graph, which he patented and sold. Captain of Detectives C.D. Lee also developed a portable polygraph, which he sold. Lee and Keeler insisted that the buyer take instruction on test methods and use of the instrument before it was delivered. Keeler, who earned a B.A. in Psychology from Stanford, took his instrument to Los Angeles when Vollmer became Chief of Police, and later to Chicago where he worked for the Institute for Juvenile Research. After a trip back to California Keeler returned to Chicago to join the Crime Laboratory at Northwestern University when the university law school founded the nation's first crime laboratory (Goddard, 1954). Although Larson published a scholarly and thorough book on lie detection in 1932, he does not say much about question sequencing or test formats (Larson, 1932). However, C.D. Lee and Leonarde Keeler did write about test formats, including methods we would now call GKT or POT.

Lee wrote to John Edgar Hoover, Director of the FBI, on August 26, 1937, in reply to a letter of inquiry from Hoover. Lee took some time to explain various techniques. Lee gave an example of a test that has the elements of a GKT, but multiple key items. Describing the instructions given to the examinee, Lee wrote:

> "You are suspected of a recent crime. I am merely going to mention some of the facts connected with the crime. If innocent, they will mean nothing to you; but if you are guilty, your consciousness will associate them with your crime. You need say nothing. Just hold still and listen."

1. You were recently in Chicago.
2. San Francisco.
3. Portland.
4. An old women was clubbed and robbed.
5. A women was criminally assaulted.
6. A young boy was kidnapped.
7. The boy was kept in an apartment house in town.
8. In a barn in the hills.
9. In an old house in the country.
10. His captors demanded $10,000.
11. $20,000.
12. $50,000.

Lee added, "If the peaks in the blood pressure curve correlated with questions 3, 6, 9, and 12, there could be little doubt that the suspect was the right man. Failure to do so would surely eliminate him."

Lee also described what is now called a 'searching peak of tension,' or to Keeler graduates, 'Type B.' Lee suggested to Hoover that the suspect may reveal details of a crime not known to authorities. For a case involving the disappearance of a person for unknown reasons he suggested this format:

1. The Bank of American was robbed this morning.
2. Jones was found dead in bed. (mythical)
3. Brown has been missing for two weeks (the missing man)
4. He has lost his mind.
5. He was accidently drowned.
6. He was murdered.
7. He was shot.
8. He was poisoned.
9. He was beaten to death.
10. He was strangled.
11. He was stabbed.
12. His body was buried.
13. His body was hidden.
14. His body was thrown in the water.
15. His body was cut up or destroyed.
16. The motive was financial gain.
17. The motive was revenge.
18. The motive was jealousy or hatred.

Lee commented that "If our stimuli here is properly balanced, the consciousness of an innocent suspect should react about equally to all the suggestions, but if guilty there should be pronounced reactions at certain points which would indicate real facts of the case." (quoted in Ansley & Furgerson, 1987)

On March 4, 1935, E.P. Coffey, head of the new crime lab at the FBI wrote a lengthy

memo to Clyde Tolson, Deputy Director of the FBI, reporting on the training he received from Leonarde Keeler in Chicago during the period February 25th to March 3rd, 1935. Coffey observed and conducted cases with Keeler, including a number of cases involving banks. In regard to test methods, after describing a relevant/irrelevant test, he noted another method called the "amounts test." Coffey said that this test is used when prior test indicated some guilt. "The subject is asked whether his thefts from the bank exceed any of a series of amounts which are called off to him which generally range from a nominal sum to $20,000. Invariably the charts would indicate relief in emotion as the amounts passed into larger sums and according to Keeler the amount of the theft on the mind of the subject is accurately indicated on the charts." Later confessions seemed to bear him out on this statement. (quoted in Ansley & Furgerson, 1987)

There is a brief description of a GKT test by Thomas Hayes Jaycox writing in The Scientific American in 1937. Jaycox was the police examiner for Wichita. In describing interesting cases he mentioned one in which a highway patrolman took into custody a man who might know who committed the "one-way-ride gang murders" of a rum runner. The man refused to talk but agreed to a test. Jaycox gave him a "name" test which he described as a group of names of men who might have committed the crime. Jaycox said the examinee gave "little or no apparent response, except to one name at which his blood pressure and respiration became abnormal. He confessed."

Most of the many illustrative cases described by John A. Larson in his classic 1932 work Lying and its Detection were tested with the relevant/irrelevant method. However, in describing a 1928 case conducted by himself and Leonarde Keeler, the latter a collaborator in writing the book, Larson described tests that took place over several days and included the use of maps of California, Oregon and Washington, then just Western Washington, to locate a victim's body. Then the test results narrowed it to King County (Seattle). The method was to point to parts of the map and asking, "Is it here?" When they used large scale maps of subdivisions of Seattle, the reactions were to an area called Bothel. There were two cemeteries in that section so a plat was made of every grave in the more likely one, the Swedish Cemetery. The suspect refused to look at the map, then jumped up and smashed the polygraph instrument! Before the map tests, the suspect had been tested with another searching peak of tension test, and some of the questions are listed in the account:

Q. Did you stab Bassett with a knife?
A. No, sir.
Q. Did you poison Bassett?
A. No, sir.
Q. Did you dope Bassett?
A. No, sir.
Q. Did you shoot Bassett?
A. No, sir.
Q. Did you strangle Bassett?
A. No, sir.
Q. Did you destroy the body?
A. No, sir.
Q. Did you burn the body?
A. No, sir.
Q. Did you cut up the body?
A. No, sir.
Q. Did you destroy the remains with a chemical?
A. No, sir.
Q. Did you scatter the remains?
A. No, sir.
Q. Did you bury the body?
A. No, sir.
Q. Did you get rid of the remains near Clark's 'Little Brown House'?
A. No, sir.
Q. Near 'The Little White House' near Bothel?
A. No, sir.
Q. Near one of the two houses in Cathcart?
A. No, sir.
Q. Did you drop the body in the well?
A. No, sir.
Q. Did you drop a concrete slab on top of the body?
A. No, sir.

Larson said that he and Keeler believed the reactions to specific question in that test proved that Mayer, the suspect, shot Bassett, did so at the 'Little White House' near Bothel, and that he buried his body under a concrete slab. After repairing the instrument and before further tests could be completed,

an injunction to stop testing was obtained by Mayer's attorney.

A chapter by Leonarde Keeler on "The Detection of Deception" in Keeler, et.al. (1938) includes a description of a "Peak of Tension test" to be used in criminal and personnel cases. Keeler said "this test may be used particularly in cases in which common facts are known to the suspect." He said the common uses were a name test, amounts test, object test, map test, age test, and type of crime test. Keeler said "the test procedure is explained to the subject and instructions to remain quiet as possible are given. In the usual experimental test a 'normal' of the subject is obtained, 1½ - 2 minutes depending on extent and frequency of normal variations. The subject is then instructed to answer all questions by 'yes' or 'no' or to refrain from giving verbal responses." The time between questions was ten to twenty seconds and a 'normal' was again obtained of 30 seconds duration following the last question. Keeler noted that the list might be repeated once or twice for verification. Keeler does not comment on placement of the key item, nor did he say anything about a review of the items in the list beforehand. Keeler did give specific instructions for reading these charts (as opposed to the instructions for the "Specific Response Test" which we would now call relevant/irrelevant). The instructions were:

"One or more of the following factors are indicative of point of deception:

(Blood pressure pulse)

1. Peak of tension (highest point on blood pressure curve).
2. Decrease in pulse frequency usually followed by increase.
3. Greatest variation in blood pressure curve immediately following stimulus.
4. General irregularity of blood pressure curve preceding point of deception followed by a smoother curve.
5. General gradual rise in B.P. curve following point of deception (rate type of response).

(Respiration)

6. Regular normal respiration to point of deception, suppression (decreased amplitude and rate) during period between deception stimulus and next stimulus followed by relief (deeper and more rapid respiration).
7. Suppressed respiration during entire period preceding deception stimulus followed by deeper respiration for remainder of test.
8. Respiratory blocking (apnoea) at deception stimulus. (Subject stops breathing in expiration for one or more respiratory cycles.)
9. Regular respiration preceding and including period following deception stimulus followed by irregular respiration for remainder of test.
10. Irregular respiration preceding deception stimulus followed by regular respiration for remainder of test.

(Muscular)

11. Muscular movement after the stimulus following deception stimulus.

(Psycho-galvanic reflex)

12. Decrease in apparent skin resistance up to and including period of deception, followed by increase in resistance (peak of tension).
13. Greatest response (apparently change in skin resistance) following stimulus."

This work by Keeler may be the first time the format is described specifically as a "peak of tension test". Keeler gave a case example and instruction for making up the list. In the case, a burglar stole four diamond rings, two watches (Waltham and Elgin), and a ruby breast pin. The burglar ate some raspberry pie and drank some milk. Keeler noted that "except for the victims and the police, the only person who knew the description of the stolen property and the food consumed was, of course, the burglar himself." Of three suspects, two did not give specific reactions to the questions below. The third suspect reacted to question four, but also reacted to the correct items in a list of jewelry and to the raspberry pie in a list of food. He confessed. The first question series was:

1. Within the last two days did you steal an auto?
2. Within the last two days did you steal a bicycle?

3. Within the last two days did you hold-up someone?
4. Within the last two days did you burglarize a home?
5. Within the last two days did you pass a bad check?
6. Within the last two days did you rob a bank?

The test series was repeated twice for each suspect. The repetition, Keeler said, was to eliminate accidental responses. Keeler noted that reaction to one question, such as the burglary above, did not necessarily indicate the suspect's guilt to a particular burglary. "However," he said, "if in another test during which questions about different types of jewelry were asked the subjects responds specifically to questions pertaining to the stolen jewelry, indications of his guilt become stronger. If, in yet another test during which ten types of food are mentioned as having been eaten at the time of the burglary, the subject responds to pie and milk the operator can safely make a diagnosis of guilt in the particular case." Keeler added that, "this particular procedure (peak of tension test) is only reliable when the facts mentioned in the tests have not been divulged directly by the investigators or through the press." (The theft case was also described by H. Mulbar, Michigan State Police, in 1944.)

The principle of concealed knowledge is illustrated by another of Keeler's cases, even though the format was R/I, and quite unlike the usual POT or GKT format. In 1931, Keeler wrote the following account:

> "There was one case where a burglar was opening a safe in a second story apartment when the owner of that apartment came in. The burglar turned and fled for the window, and in his effort to get out, pulled down some heavy plush curtains. He couldn't make his exit, so he wheeled around, shot the owner of the apartment, and bolted out of the door.
>
> "Through the modus operandi system, they put their finger on four burglars the next day and brought them in, and did not tell them what they were suspected of. We put them on the machine, one at a time, and at first ran a long normal about four minutes, to ascertain their reactions or their fluctuations which are normal to that individual. Then we asked three or four questions that had nothing to do with the crime: 'Is your name Jones? Have you had breakfast? Do you own an automobile?' and other such questions, merely to find out how they respond, what fluctuations we obtain when they answer questions. Then we asked questions such as: 'Do you own an apartment on Main Street?' That was the name of the street that this burglarized apartment was on. 'Have you a second story apartment? Have you some heavy plush curtains on your windows? Have you a safe in your apartment?'
>
> "It happened that these four burglars were innocent of that job, and they thought we were crazy asking them such foolish questions. We were aware of where they lived, and why should we ask them such questions as those? But the next day a burglar was brought in. We gave him the test, and he responded violently, gave great fluctuations in blood pressure and respiration whenever we mentioned any description or any point of that apartment house. On the third test we turned him around so he could watch the machine, and suggested that he watch the needles carefully, and told him what they would do whenever he lied. In the middle of the test he confessed and said that he saw he couldn't beat it, and he told us the complete story, which was later verified."

What is interesting about this approach is that no direct reference was made to the crime, and none was needed.

In his instruction manual of 1943, C.D. Lee describes an "association method of questioning." The test was administered as an R/I sequence with many questions, but there were choices as to the method of murder, time of day, location, and what was stolen. Here is his example:

" 1. Is your name Black?

2. Do you live in Berkeley?
(3) Do you know who killed White?
(4) Did you kill White?
5. Did you shoot White?
6. Did you stab White?
(7) Did you hit White on the head?
8. Did you use a pick handle?
9. Did you use a wooden club?
(10) Did you use a piece of lead pipe?
11. Did you attack White in the morning?
12. Did you attack White in the afternoon?
(13) Did you attack White after dark?
14. Did you attack White near a lake?
15. Did you attack White near a house?
(16) Did you attack White near some bushes?
17. Did you attack White in an alley?
18. Did you attack White in the street?
(19) Did you attack White in a park?
20. Did you steal White's suitcase?
21. Did you steal White's overcoat?
(22) Did you steal White's leather wallet?
23. Did you take a $10 bill from the wallet?
24. Did you take $75 in currency from the wallet?
(25) Did you take $500 in currency from the wallet?
26. Did you steal White's cigar lighter?
27. Did you steal White's pocket knife?
(28) Did you steal White's gold Waltham watch?
29. Did you steal White's gold fountain pen?
30. Did you steal White's pearl tie pie?
(31) Did you steal White's diamond ring?"

The questions in parentheses are the crucial questions. The others are controls. The innocent suspect, Lee states, "cannot possibly associate only the crucials as distinguished from the controls with the crime, since he knows nothing concerning these details." (Lee, 1943) Lee repeated this example in his 1953 book.

Also in both of Lee's works in an example of a test of hotel employees shortly after a jewel theft from a guest who was in suite 350. The jewels had been well concealed behind a hat box in the closet. The test readily discovered the thief. As an interesting point, the crime is not mentioned in any question.

"1. Is your name Y?
2. Do you like your work at Hotel W?
3. Do you like nice clothes?
(4) Do you like jewelry?
5. Between 2 and 3 today were you on the fifth floor?
6. Between 2 and 3 today were you on the fourth floor?
(7) Between 2 and 3 today were you on the third floor?
8. Did you call at suite 370?
9. Did you call at suite 360?
(10) Did you call at suite 350?
11. Did you enter the living room?
12. Did you enter the bathroom?
(13) Did you enter the closet?"

The controls are questions 3, 5, 6, 8, 9, 11 and 12. In regard to the sequence, he said that it was not essential that there always be two control questions before the crucial question, but he observed that the crucial should never be first. He also said that any number of controls may be used with each crucial, some placed before and some after if desired, but two usually suffice. Lee also suggested that the prefix, "Do you know whether ..." may be used in place of "Did you ..." because the former does not carry the imputation of guilt. (However, "Do you know whether ... " creates a problem in that the deceptive subject is lying to all the POT/GKT question, as he does know whether.) Lee suggested the "association" method worked well with a general time of day, places, objects used or stolen, other evidence, motive, and manner of concealment. Lee's use of "association" here is in the POT/GKT sense, not the word-association format as a test.

C.D. Lee (1949), writing about "Formulating the Test Questions" described a case in which Berkeley detectives had interrogated a man for days who was suspected of raping a child in some poison oak bushes. They had some good evidence that he had a poison oak infection on his genitals. Keeler was home in California for a short visit and was asked to give the suspect a polygraph test. From the traditional who, what, where, when and how, he selected when. Knowing **when** the assault took place, Keeler used a test sequential questions beginning with, "Do you know whether it happened about one o'clock?," ending with six o'clock. The subject reacted with a peak of

tension at four o'clock. When this was pointed out as the correct time the subject confessed.

In 1942 Fred E. Inbau, professor of law at Northwestern University and former Director of the Chicago Police Scientific Crime Detection Laboratory, wrote a book, Lie Detection and Criminal Interrogation. In his book Inbau referred to the "peak of tension test," and described the procedure for developing the test. Assuming a diamond ring was stolen the examiner would, "(1) draw up a list of about seven articles of value - for example, a gold watch, a pearl necklace, etc. - including a 'diamond ring' as one of the articles; (2) show the list to the subject, with an explanation to the effect that among the articles on the list is the one which was stolen from the burglarized premises; (3) inform him that on the test he will be asked, in separate questions, if, to his knowledge, the object taken in the burglary was any of those named on the list, to all of which questions the subject will, of course, answer 'no;' (4) then obtain two or three lie detector records based upon such test questions." In a footnote Inbau said, "it was advisable to place the name of the missing article somewhere in between the first and the last on the list." Inbau's 'peak of tension' test criteria was either "the highest point in the blood pressure-pulse tracing, or a line of demarcation, so to speak, between a somewhat irregular, unsteady portion of the respiration or blood pressure-pulse tracing, and a more regular, steady recording from that point on." He noted that in many instances the 'peak of tension' would show up in both tracings. A galvanometer was not then used by Inbau. In illustrations in the book the charts showed tests on two subjects in which the reactions correctly indicated which of ten persons shot a sheriff, and which one drove the automobile occupied by the bandits at the time of the shooting. This peak of tension, said Inbau, "is attributable (1) to the guilty person's anticipation or apprehension of being asked the one question on the list to which he will lie, and (2) to the relief of tension he experiences after answering that question."

Inbau said that peak of tension tests may be used in a variety of cases, provided of course, the subject has not been informed of the essential details, such as the object stolen, the amount of missing money, or the implement used in the commission of the crime. The same instructions appeared in the second (1948) and third (1953) editions of the book, the third edition being co-authored by John E. Reid.

By 1951 at the Keeler Institute the use of one item per list was well established. The POT was taught as being more accurate than the 'relevant/irrelevant' or 'general question' test. A peak of tension 'Type A' was one where the examiner knew the key item, and 'Type B' was a searching peak where the key item was unknown to the examiner. Examiners were taught to use the POT whenever they could, and that two repetitions of the question lists could be put on one chart. The list of questions was always shown to the subject before the test to build upon the anticipation and to accentuate relief afterwards. The questions were to be worded alike except for the one variable. For example:

"Did you steal a Buick last night?"
"Did you steal a Ford last night?"
"Did you steal a Plymouth last night?"
"Did you steal a Chevrolet last night?"
"Did you steal a Pontiac last night?"

Examiners were taught to put the crucial item in the center of the list. They were also taught to use a logical progression if there was one, as might be the case with room numbers or amounts of money. Examiners were not to put an illogical item in a list. Several lists could be used, and there should be five to seven items in each list. Deception criteria were rise and drop in blood pressure, which was called ideal; or an irregular cardio pattern before the key and regular, straight or down pattern afterwards. There could also be a single rise and fall of the blood pressure in response to the key item, and irregular thereafter. In the pneumograph the pattern could be irregular to the key, regular thereafter, or the reverse. There could also be a specific reaction to the key item, between regular patterns. The galvanometer, considered the least reliable (cardio the most) would probably rise at each item, a big rise at the key, then level off or drift after the key item. However in a Type B POT the galvanometer was considered much more useful, and respiration second, although a

cardio reaction could be expected. A case conducted by Keeler in which he found the body of a Navy officer by starting with a national map and worked downward to local maps was described (Ansley, 1951).

Detective Sergeant Freeman B. Ramer of the Pennsylvania State Police sent a story to the ISDD Bulletin (International Society for the Detection of Deception) which was published in January 1949. The case involved a murder and robbery in which a man had been beaten to death with a rock that had been painted white. The rock, which had blood on it, had been thrown over the bank from the location of the body. A black man was the prime suspect because the explanations of his actions on that night were unsatisfactory, and he was nearly incoherent. There were also two other suspects who were white. The tests were peak of tension in which the kind of murder weapon and its location were the key items. The black suspect did not respond to the correct items, but both the white suspects did response specifically to those key items, and they subsequently confessed to the murder.

Colonel Ralph W. Pierce, USA (Ret.) wrote an article in 1950 on "The Peak of Tension Test." He said that when using POT type of test the questions were either shown to the subject or read to him before the test was given. Col. Pierce said, "the deception criteria in the peak of tension test may be either a peak or high point in the blood pressure-pulse recording or an irregular pattern to the point of deception, followed by relief evidenced by a regular pattern in the tracing from the point of deception to the end of the test. This criteria may be found in either or both the blood pressure-pulse and respiration recordings." Col. Pierce added that "the psychogalvanic reflex, or electrodermal response, is also very important in peak of tension tests. In fact, in some cases where little, if any, change is found in either the blood pressure-pulse or respiration recordings, it becomes the most important indication of deception."

Colonel Pierce gave as an example a case which happened in Wildburg, Germany in 1946 in which an Army Captain was shot seven times by a soldier. The only evidence at the scene was a bag of food dropped by the soldier and a German Luger pistol found later not far from the scene. The food was identifiable as coming from a particular mess hall so those who had access to the keys were given polygraph tests. The test was as follows:

"Did you shoot the Captain with an American Colt?"
"Did you shoot the Captain with an Italian Beretta?"
"Did you shoot the Captain with a German Luger?"
"Did you shoot the Captain with a Swiss Sauer?"
"Did you shoot the Captain with a German P-38?"

One suspect reacted with his blood pressure rising until the question concerning the German Luger was asked, then it declined. He showed marked irregularity in his breathing up to the question about the Luger, followed by regularity to the end of the test. The galvanometer pen rose sharply at the question concerning the Luger. Assuming that only the guilty man knew this detail, the suspect was interrogated, and he confessed. Col. Pierce was then President of Leonarde Keeler, Inc. which included the polygraph school. In 1950 there were no other polygraph courses.

Charles H. Patnode, Special Agent of the United States Secret Service, described peak of tension at the New York Conference on Criminal Interrogation and Lie Detection at New York University Law Center on November 8, 1952. He said the peak of tension test "consists of one pertinent question surrounded by **six** or seven irrelevant questions. In the case of a murder weapon, the type known only to the murderer and the investigator, questions relating to the types of weapons one might use in committing a murder would be ideal. He suggested this format: Do you know if the murder weapon was a shotgun, an ax, revolver, hatchet, hammer, knife, or a poison? The actual weapon would be placed anywhere in the list except at the very beginning or the very end. The subject is to be shown the list of questions before the test, and if guilty, the ink impressions should form a peak at the murder weapon." He suggested subsequent tests to cover the place of the murder, objects stolen from the corpse, and any other data in the investigation. Each test is repeated two or

three times to ensure the elimination of accidental responses (Patnode, 1956).

Cleve Backster, who had been an instructor at the Keeler Institute in 1951, included a POT format in each of his Standardized Polygraph Notepack booklets (Backster, 1963, 1969, 1979). The instructions in each were alike.

His instructions for a 'Known Solution Peak of Tension Test' included development of a preparatory question, a question prefix, padding choices, and a key choice. The format was as follows:

Preparatory Question
Question Prefix
1. Padding choice
2. Padding choice
3. Padding choice
4. Padding choice
5. Padding choice
6. Padding choice
7. Padding choice
8. Padding choice

The preparatory question might be, "In regard to the car used in the holdup, and the question prefix was, "was it a ...?" The padding choices were likely alternatives to the key, such as Buick, Chevrolet, Dodge, etc. The key item, such as Oldsmobile, could be placed in positions, 3, 4, 5 or 6, but not in the first two or last two positions.

Backster's searching peak of tension was called a 'Probing Peak of Tension Test,' and the format was as follows:

Preparatory Question
Question Prefix
1. Less probable choice
2. Less probable choice
3. More probable choice
4. More probable choice
5. More probable choice
6. More probable choice
7. Less probable choice
8. Less probable choice
9. All inclusive choice

These formats were, and remain, widely used. They are used by graduates of the Backster School of Lie Detection and many others who have heard Cleve Backster lecture on his several techniques at polygraph seminars since his first notepack appeared in 1963. Actually, Backster has been lecturing at polygraph seminars since 1950, and has been very influential in standardizing methods.

As a rather interesting sidelight, in 1959, Cleve Backster sent a memo to all examiners in the Academy for Scientific Interrogation (a predecessor of the APA) outlining his research results. That memo stated:

> "Research has been conducted in which a three choice peak-of-tension test, involving various amounts of money in each of three envelopes, is superimposed on the regular test. Each of the three 'money envelope' choices, including the one theoretically stolen by the subject, are placed in critical locations within the test. We now have a mild created lie, which is subject to preliminary review and conditioning effects through respiration. It is directly comparable to reactions or lack of reactions to pertinent question. This technique is usually far too subtle to stimulate deception indication in the blood pressure-pulse or breathing, but has produced very interesting results with the G.S.R. tracing."

We have no knowledge of this ever being put to use in field testing, but it is possible.

Practitioners in More Recent Times

In 1970 Richard 0. Arther defined a known-solution peak of tension test as usually containing seven questions having to do with a particular detail of a crime in which the polygraphist words the seven so that only one is true and the other six false. The true question is the key and the others irrelevants. The truthful person, he said, does not know which is the key. However the liar must recognize the key. He observed a danger, that the truthful person has learned the key but does not want to admit it perhaps because he got the information improperly, such as reading the case file when the investigators were out of the room; or the information was given to him by the perpetrator. Another

danger is that the key is relevant to the truthful subject for other reasons. The truthful may respond to '1.38 caliber revolver' because of some other crime or incident in which he used such a weapon (Arther, 1968). In a 1970 article on question formulation in peak of tension tests, Arther introduced a novel concept, the 'False Key'. He recommended that in every known solution peak of tension test there should be a false key in position two. Arther told of a case in which a prior control question test indicated truthfulness, and in the seven-item peak of tension test the person always reacted to the same irrelevant item, even though the sequence of the items was changed around for each of the three charts. The item was a rather obvious one for an innocent person. However, there was no more reaction to the key than there was to the remaining irrelevants. The subject's innocence was later verified. Following that 1960 case, Arther had another in which a suspect in a robbery of a woman who had just shopped at a grocery store did not react to the key at number four position, a hat box, but to the more logical paper bag at number two, which was irrelevant. Since then, Arther has always used a false key at number two in each peak of tension test. When possible, this has been an item that has been the most obvious item. If the obvious item is the key, then Arther stated it is necessary to subtly overemphasize an irrelevant at number two so the truthful will guess that is probably the key item. The subtle emphasis is done by saying a little more about the item, use of a gesture, possibly by reading the item a little louder.

Otherwise, Arther's peak tests followed a fixed pattern, seven items if possible with the key at number four, and seven items if possible in searching peak of tension tests, with number seven being a question about something else not mentioned (Arther, 1970).

In their 1977 textbook Truth and Deception, John E. Reid and Fred E. Inbau described a "peak of tension test" as "asking a series of questions in which only one has any bearing upon the matter under investigation. This one pertinent question refers to some detail of the incident or occurrence (e.g., the kind of object stolen, the kind of implement used in a crime, etc.) which could not have been known by an innocent person or by anyone who had not been informed previously of such detail." They said that when the item is mentioned during the test, "a peak of tension may appear in one or more of the subject's polygraph tracings."

In giving the test they said the subject was not to be told the order of the questions or articles before the first peak of tension test, and not even told what the various named articles will be. The object was to "achieve an element of surprise on the first peak of tension test, but only on this first test. Thereafter, on the subsequent peak of tension tests (of which there should be three in all), the original order of the questions should be maintained and the subject so advised prior to each test." The second peak of tension test should be given shortly after the first one and the subject told that the questions would be the same, and asked in the same order. After the second test the Reid and Inbau instructions called for the examiner to leave the room for a few minutes, and to tell the subject that when the third test is given their blood pressure may go up at the exact time the question is asked that includes the item that was actually stolen. The examiner adds, "If you're not telling the truth, of course, the next test will point to the item that was stolen, and I'll know you took it." The examiner leaves to allow the subject some time to review in his own mind the prior test. Reid and Inbau were very specific in their directions for conducting the test. The text contains 32 charts from their case files that illustrate the way in which such charts are analyzed.

Some charts show the peaking effect of cardio, respiratory and electrodermal patterns, other charts show responses that are specific to the item. Their charts included searching peak of tension tests where the examiner was asking the amount, location, or name from a list of choices (Reid & Inbau, 1977).

The Searching Peak of Tension Test at Work

On March 17, 1977 Lori Ashmore and Kathy Brown were kidnapped. The kidnapper demanded $500,000. A tap on the receiving telephone during a second call traced the origin to a trailer park, with the number listed to an ex-convict named Larry L. Chaney. A

subsequent call was traced to a telephone booth, where latent prints matched those of Chaney. On March 19th, Chaney was arrested, and he denied knowledge of the crime. Chaney and his attorney agreed to a polygraph examination concerning the location of the missing women, utilizing a 'searching peak of tension' technique. It was agreed that the examiner would ask only geographic locations and Chaney would answer "no." Owen W. Wilkerson, an examiner from the Oklahoma Bureau of Investigation, conducted the examinations. Tests had been prepared by counties, in lists of five each, with the county in the middle being the more likely, and the unlikely counties as padding, presumably first and last on the lists. However, the likely counties constituted a land mass as large as Massachusetts. The test first amounted to a list of five counties printed in dark letters placed on the wall in front of Chaney. There was no "coverall" question at the end, a question about a place not mentioned. The first three lists did not produce significant responses except a spot response to Cherokee County, which is next to Sequoyah County. On the fourth list the cardio tracing built up to Sequoyah County, and dropped dramatically after Chaney replied "no." His pulse rate had gone from 96 to 120 during the first three lists, now it went to 144. County lists were mixed up and shown again in lists of ten with Sequoyah County omitted the first time, resulting in no responses. The second mixed list, also of ten counties, included Sequoyah in the sixth position. The cardio on this chart built up to Sequoyah, then fell. With an altered list, that test was conducted again, with the same results.

Maps of each county had been prepared with lines dividing them into four quarters, A, B, C and D. Sequoyah County was shown and Chaney reacted to section C. It was already known that Chaney had property there, and it had been searched with negative results. Plat maps were used for section C, and Chaney reacted to plat A. Was this just a reaction to his property being there on the map? Upon being asked to do so, he pointed to the location of his property. He was asked if there was a pond on the property. Chaney's reply was that it was "too shallow to put anything in." Between two and three hundred law enforcement officers with airplanes, helicopters, dirt bikes and dogs converged on the densely wooded property. A shallow grave with the bodies of the two women was found west of the pond, within the area circled on the plat map. Chaney was found guilty of murder and sentenced to death (Wilkerson, 1977).

In April 1977 a woman was reported missing to the City of Starkville Police Department in Mississippi. When those who knew the missing person were questioned it was determined that a certain male was the last person seen with the missing female. A routine background check found that he had been the suspect in an unsolved murder investigation in Alabama. When questioned about the disappearance of the missing woman this suspect gave such an outrageous reason why he could not have been involved in her disappearance that his statement along with the background check influenced officers to request that he take a polygraph examination.

The suspect submitted to two standard zone comparison polygraph examinations, conducted by Detective Edward P. Brennan. The first examinations centered on the issue of being involved in the disappearance of the missing woman, the second on causing her death. The examiner's opinion was that deception was indicated in both situations. However, the suspect continued to deny involvement.

Brennan decided to conduct a searching peak of tension test in an attempt to locate the body. He divided a map of Mississippi into counties assigning each a letter identifier. He conducted a nine question test, the first two being buffers, the next four the most probable counties from the Mississippi map, question seven an area not mentioned, questions eight and nine as buffers. The suspect consistently demonstrated strong physiological reactions to questions involving three counties on the map. Although the examiner was perplexed that the suspect consistently reacted to three areas rather than one; when the general position of intersection of those three counties was pointed to on a map during interrogation the suspect confessed to the abduction and murder. The multiple reactions were caused by the fact that the examinee buried the body

on County Line Road at a place where the three counties intersected. Subsequent POT testing identified other states in which the suspect committed murders. He not only admitted to the murder at issue but during subsequent interrogations admitted to six murders in other states. According to Detective Brennan the suspect has now admitted to ten murders in all. Brennan also conducted a searching peak of tension test in an effort to determine the number of deaths in which the suspect had been involved, and he believes that number to be sixteen (Brennan, 1992).

FBI Format

A five-page Federal Bureau of Investigation handout, distributed in 1985 at a seminar of the American Polygraph Association described their version of peak of tension tests. It was much like other POT instruction at the time. The paper listed four principles: examinee is placed under tension, tension is increased to the crucial point, tension is decreased after the crucial point, and a truthful opinion should not be given based solely on any form of POT testing. There were three types of POT tests: known solution (Keeler Type A), stimulation test, and searching (Keeler Type B). POT tests were to be supplementary tests used after a zone comparison, MGQT, or other general test. Questions were to be reviewed in sequence with the examinee, a visual list was used to reinforce the sequence, and a test consisted of three charts with the questions asked forward in the first two charts and in reverse order on the third chart. A fourth chart could be conducted if the results were inconclusive. An even number of items were to used to preclude a "middle" number, moving the position of the critical item when the reverse order was used on the third chart. The prefix phrase for each question was to be, "Do you know if it was," and there were to be between five and nine items, but six was preferred. Only one key fact was to be used in each list. If an odd number of items was used, the key should be near but not at the center of the list. There were to be at least two padding (irrelevant) questions before and after the key, and the key position was to be changed for each different list. The use of a "false key" was "optional, discouraged, and discouraged except in closely controlled circumstances." A "false key," they noted, was a padding question that has special meaning to the examinee, a meaning that may be generated by the examiner. The false key concept was developed as a control to allow an innocent person to focus on an item which is not the key item, and the presumption that the guilty person will react instead or in addition to the key item. If a false key was to be used, the instruction was to put it in position number two of a known solution test. It is not used in a searching peak of tension test. The searching peak of tension tests were to be used to locate evidence or identify accomplices. In constructing a searching peak of tension test, sometimes called a SPOT, the examiner was to cover all possibilities, and padding questions which were outside the realm of possibility were to be at the beginning and end of the list, with two at the end if possible. Visual stimuli such as maps or lists were permissible.

GKT Taught as a Technique

A DoD polygraph course (not taught at the Institute) in 1986 included instruction on POT and additional instruction on GKT. The lesson plan on GKT noted these differences from the POT: the subject does not know the sequence of the questions, the relevant question is randomly distributed, and there is no problem with spot responders. The plan said there should be at least four alternatives, one placed first as a buffer and only one correct item in each test. Ideally there should be four to ten tests, with one item in each test, and the position of the key item varying by chance except that it was not to be in the buffer (first) position. It said that more than one chart could be conducted for a list. All parameters were to be analyzed. The scoring devised by Lykken (1959) was explained, but the preference was for global analysis. Disadvantages were listed as the uncertainty about a guilty person recognizing the crucial item, and the problem of finding items not already revealed to all suspects. The plan noted that errors would probably be false negatives.

Standard Text Description

Dr. Stanley Abrams' most recent book, The Complete Polygraph Handbook (1989) has a chapter devoted to "The Guilty Knowledge or

Peak of Tension Test." Citing an Oregon State Police case as an example of a searching peak of tension test, these questions were used:

Is your wife's body in the river?
Is your wife's body by the railroad tracks?
Is your wife's body in the potato field?
Is your wife's body by the farm buildings?
Is your wife's body by the house?

The reactions were to the question about farm buildings, which generated another series of questions which isolated the shed, where the body was unearthed.

Abrams gives extensive instructions on preparing lists, pointing out errors such as the use of a two-word key like "white sweater" in a list with single-word items like "loafers," "jeans," etc. Other errors cited were a list of guns with a knife included and a cheap piece of jewelry in a list of otherwise expensive items. He noted that the key item must be something remembered by the guilty person. The more lists used, Abrams said, the more certain the examiner may be of his results. Abrams described Arther's (1970, 1982) known solution test with the false key in position two of a seven-item list. The false key was described as a "control question." Arther, he noted, reviews the questions in advance, but not in the order used during the test. The key item is at position five in a list of seven items. However, the two or three charts that follow in an Arther series would be given the same sequence used the first time, creating a "classical peak-like reaction." Abrams mentioned Lykken's preference for placing the critical item in a different position each time the test was administered. Arther, said Abrams, had also recommended that the subject should repeat the last word in each question before answering "no," saying that would increase the accuracy.

Standard Army Method for POT

For many years, most federal examiners have been trained by the U.S. Army at the Military Police School. The lesson plan for Peak of Tension Polygraph Examinations at the U.S. Army Military Police School (USAMPS) for November 1984 cited the text book by Inbau and Reid (1977) and material by L. Harrelson of the Keeler Polygraph Institute. The school used a form for known solution peak of tension tests that called for a "preparatory question" at the beginning, such as "regarding the amount of money that check was written for," followed by several questions, each with the same prefix, i.e., "Do you know if it was ...?" The illustrations they used had seven items with the relevant item in the middle. The plan did not address the analysis of chart or repetition of charts, but the list of items and sequence of questions was presented to the subject before the test. The USAMPS course later became the Department of Defense Polygraph Institute (DoDPI). Although remaining at Fort McClellan, Alabama, the Institute added instructors, research personnel and support staff from all DoD agencies and services that utilize the polygraph. New buildings were constructed and instructional material was added to the course.

DoDPI Revision of POT Format

In September 1991 the DoD Polygraph Institute made a technical correction in the construction of POT tests. In the prior USAMPS system, in use for over 20 years, they asked "Do you know if ...?" In the memo changing the prefix, DoDPI noted that the Keeler Polygraph Institute Training Guide (Harrelson, 1964) specifically warned against the use of "Do you know ...?" The DoDPI observed that when the preface is "Do you know if ..." it requires the guilty examinee to lie to all the possible choices not just the key. That is so because the guilty subject does "know," and lies when he says "no" to each choice on the list. Now the methodology is more direct, as it asks only "Is it ...?" or "Was it ...?"

.45 caliber?

An example of a current DoDPI searching POT is:

Regarding the location of that bomb,
Is is located in:
Atlanta?
Birmingham?
Area A?
Area B?
Area C?
Area D?
In an area I have not mentioned?
Taledega?
Huntsville?

An example of a current DoDPI known POT is:

> Regarding the caliber of the pistol used to shoot that man,
> Was it a:
> .22 caliber?
> .25 caliber?
> .32 caliber?
> .38 caliber? (key)
> .44 caliber?

The DoDPI memo (Yankee, 1991) observed that "care must be taken during the pretest to establish the question format so that 'no' answers can logically be given. No format should allow an 'I don't know' answer."

Laboratory Research with GKT and POT Formats

In the early years many psychologists experimented with word association, and there were occasional reports of its use in criminal cases, sometimes successfully, and sometimes not (Herbold-Wootten, 1982; Winter, 1936). Laboratories often had galvanometers, of varying quality, and these were the instruments of choice for much of their research on lie detection. Not only were laboratory galvanometers occasionally used for the solution of criminal investigations, two electrodermal units have been marketed and sold to law enforcement agencies for lie detection. The first was the Fordham Pathometer, designed, and sold by Father Walter G. Summers, S.J., Ph.D., a professor of psychology at Fordham University. His was a recording galvanometer. However, he did not teach a peak of tension test. Rather, he had a sophisticated control question test in which control and relevant questions were paired and the reactions to each were compared (Summers, 1936, 1934, 1938, 1939). His test format was similar to the central part of the modern zone comparison. Another GSR instrument sold for police use was a visual meter, the B&W. It was widely used from 1944 until the 1970's. A peak of tension or a relevant/irrelevant test format was recommended (B&W Associates, 1960; Guertin & Wilhelm, 1954; Wilhelm & Burns, 1951, 1954). Pathometers and B&W galvanometers were also used in laboratory research, as the equipment was reliable. The B&W, however, did not have a chart recording capability, but later models had a tape playback feature. In the research reported in this study, a variety of laboratory instruments have been used, including units manufactured by Beckman (including Offner and Sensormedics), Brush, Grass, Lafayette, Narco-Bio-Systems, Sanborn, and Stoelting. Even a toy lie detector has been marketed (c. 1973). The "Super Sooth," at $20, came complete with meter, electrodes, and a detailed instruction book on searching and known peak of tension tests, worked into games.

Perhaps it was the preoccupation with word association that kept psychologists from taking much notice of the peak of tension test or other varieties of GKT in the 1930's. In 1947, two professors who were trying to solve a theft at Cornell University, successfully supplemented their relevant/irrelevant tests with peak of tension tests, and correctly concluded that the thief was not among the 81 men they tested (Bitterman & Marcuse, 1947). This was their only attempt at real lie detection, but in 1954 Marcuse was the co-author of an article in which they performed peak of tension tests, to detect a playing card, with a cardiopneumo polygraph and an electrodermal meter. Their detection rate was well above chance (VanBuskirk & Marcuse, 1954).

The first real laboratory study involving a peak of tension test as we know it now is probably the work of Christian A. Ruckmick (1938). Using an electrodermal meter with a 30,000 ohm range (that range is quite limited, a modern Lafayette has a range of one million ohms, up from 500,000 ohms in 1979), Ruckmick tested 89 students. In his first experiment Ruckmick tried to detect with meter deflection the number the subject chose from a pile of cards. That didn't work very well so he changed to ten cards with three letter words such as "nor," "and," "can," etc. There were buffer words at the beginning and end of the list which were not written on cards. The question prefix was, "Is it ...?" and the answer was "no" to all words. The detection rate was 78%. The number of judges was not listed but the removal of an undergraduate student's work raised the detection rate to 83%. The experimenter had an additional phase in which a half a dozen students who "got excited" about the wrong word were generally successful in "throwing

the examiner off." That was the first and one of the few POT projects that mentions application of a countermeasure.

Edward W. Geldreich conducted two experiments on lie detection using peak of tension tests in which 50 college students picked one of five cards (1941, 1942). His instrument was a wheatstone bridge with a calibrated potentiometer to balance in the subject and a visual meter. The first test for each subject was asked about five cards, but not one of them was the card the subject picked out. Subjects were all truthful in this test, a test to "condition" them. The next test had the card the subject picked among the five cards that were turned face up. Geldreich picked the right card in 37 of the 50 trials for a detection rate of 74%. This first combination was interesting because of Geldreich's use of a truthful series to begin with. Truthful subjects have not always been used in POT/GKT research (Timm, 1989). Unfortunately, Geldreich did not report on false positive errors, if there were any. A second part of his 1941 study was also novel. In that research there was a series of tests in which he prolonged the conditioning test to develop fatigue, with 25 to 50 irrelevant cards presented to each subject, until habituation was so complete that there was no response at all to five successive cards. Then the five cards were presented that included the card the subject had mentally selected. There was an odd result of this repetition to extinction with truthful responses, prior to the test with a lie. The detection rate was 100%.

In 1942, Geldreich decided to study the effect of fear on detection. Using his first study in 1941 as a control group, with detection at 74%, he gave his experimental subjects the same instructions and test as those in the control group except that each subject was told they would be given an electric shock if the GSR gave away their selection of a card. Before the test, each subject was given an electric shock so severe it made them jump. In fact, no shocks were given during the tests, but that was not what subjects believed. The detection rate for the experimental group was 43 of 50, or 86% correct. The average electrodermal response for the irrelevant responses in the control group was 3.6 mm, while the aroused experimental group averaged 4.4 mm. The average response to lying by those in the control group was 13.9 mm, while the shocked experimental group averaged 16.8 mm. Many of the laboratory research projects that followed Geldreich lacked the useful data he included.

In 1948, Baesen, Chung and Yang published an experiment in which they reported on a peak of tension test which appeared to have been mixed in with another test format, both relating to a mock crime. The problem was to separate perpetrators from witnesses. Their format of relevant questions was described as:

Set 1
(4) "Does [amount stolen] have particular significance to you?"
(8) "Did [name of accomplice] steal the money?"
(10) "Did you steal the money?"

Set 2
(3) "Did you steal the money?"
(9) "Does [amount stolen] have particular significance to you?"
(12) "Did you watch [name of accomplice] steal the money?"
(16) "Did [name of accomplice] watch you steal the money?"

The authors said directly after the list above, "The peak of tension on the stolen sum was brought about by arranging the questions in consecutive order beginning with two amounts not stolen and then the third question as the critical sum followed by the last sum known not to be critical. With the exception of the peak of tension series of questions, the relevant questions were adequately separated by irrelevant and control questions." The instrument recorded cardio and respiratory functions. It is not clear from the description as to whether the amounts were consecutive or spread out among the irrelevant, other relevant, and control questions. It does appear that both test methods appeared together on one chart. That they were correct in 86% of their trials is remarkable, considering the mixed format.

In 1952, D.G. Ellison at Indiana University conducted several lie detection studies for the U.S. Navy. One was a simple test with a B&W meter and ten college

students. The students were given a sheet of paper and told to circle any one of the six months listed. The list was the first six months of the year. The questions by the experimenter were, "Is it January?" "Is it March?" and so forth. Each question was answered "no," producing five truthful answers and one lie in each series. The question interval was 20 seconds. Each question was asked six times in an order which was semi-random, in that no question was repeated before all questions had been asked once. After all this was done, the procedure was repeated with a month from the last six months of the year. The experimenter computed the mean meter deflection rate for each month from the two runs. The month with the largest mean deflection from each run was considered the "detected" (circled) month. The "detected" month was the correct month for the first experiment with eight of the ten persons tested. They missed once because there was a tie in mean deflection between two months, although one of the pair was the examinee's selection. In the second series the "detected" month was correct with seven of the subjects. The results were significantly above chance. The semi-random distribution was an interesting feature of the experiment as it cancelled out any serial effect. Also, the examinee was blind to the sequence for each series.

In a second experiment, Ellison used 23 students, 11 in one group and 12 in a second group. The experimental method was identical to the prior project except that after the first phase the subjects of group one were told the month the experimenter believed was correct, based on the mean meter deflection, and the subjects of group two were told a month that was probably wrong, as it was the month with the least deflection. As in the first experiment, the 23 subjects were tested again, on which month they circled on a list of the last six months of the year. For group one, the detection rate on the first phase was nine of 11 (82%) and was 3 of 11 (27%) on the second run, after being correctly informed of the first test results. The two failures on the first run were also failures on the second, and the three successes on the second run were also successes on the first run. For group two, who were misinformed of the first test results, the initial detection rate was nine of 12 (75%), and was ten of 12 (83%) on the second run. Eight of the nine correct decisions on the first run were persons who were among the ten of 12 correct decisions on the second run. One of the failures on the first run was among the two failures on the second run. The novel aspect of this project was informing one group correctly of their decision in the first series and misinforming a matching group, and assuming the difference in results was related to the differing instructions. The results, however, defy conventional wisdom, as one would expect the misinformed to be detected at a lower rate or at the same rate. Saxe (1988), a polygraph critic, has insisted that belief in the validity of testing was necessary for it to work. Using a zone comparison format and a mock crime, Yankee and Grimsley (1986) found a trend in which accurate feedback was 94%, inaccurate was 86%, and 79% for no feedback. However the differences did not reach statistical significance ($p < .05$). Barland and Raskin (1972) used a peak of tension stimulus test with a Backster zone, a test in which one group was shown a polygraph chart which correctly indicated the card picked, another group was shown a polygraph chart which depicted an incorrect selection, and a third group did not receive a stimulus test. The manipulation of these stimulus test results did not produce any significant effect on the detection of guilt of innocence for the mock crime. Diaz (1985) found that of those told they were detected after the first card test, the subsequent detection was 27 of 40 (68%), while those who were told they were not detected by the card test were subsequently detected in 28 of 40 (70%) tests. Elaad (nd) reported no change in detection rates for positive feedback in GKT tests, and a modest decrease in the detection rates of those given no feedback. There were no numerical data in the paper. Regardless of the outcome of Ellison's research, he was the first to explore the effect of positive feedback and false feedback on subsequent tests.

In another experiment, using a different galvanometer, Ellison tested eight students to determine the month of their birth. Each test list contained four months of the year, and each series was tested twice with the four months asked twice, each time in a different sequence. The eight students were tested three times in this manner, once

with instructions to say "no," once with instructions to say "yes," and once with instructions to remain mute. The sequence of these conditions was varied so as to offset the serial effect. Detection for the "no" tests was four of eight, two of eight for "yes" answers, and one of eight from the mute tests. The idea of "no," "yes," and mute has since been tested by many others, but this may be the first research on this topic.

In a 1955 review of the accuracy and status of lie detection, Bejamin Burack said the "disguised questions test," "when used for a person who could not reasonably be expected to be familiar with certain details of the offense, has logical validity." As an example of a "disguised questions test" Burack considered a burglary in which a gold watch was taken. These questions would be asked:

> Do you know whether a pearl necklace was stolen?
> Do you know whether a diamond ring was stolen?
> Do you know whether a gold watch was stolen?
> Do you know whether a fur coat was stolen?
> Do you know whether a silver bracelet was stolen?

Although this is the classic five question peak of tension test with the key item in the middle, Burack suggested a variation in which no answer is given, and another variation in which only key words in each question were asked, such as "pearl necklace?," "diamond ring?," etc. Burack observed that "some examiners permit the person to see the list of questions before asking them, on the theory that knowing what will be asked serves to stimulate (in guilty persons) greater emotional response to the one relevant question. Because the guilty person builds up tension as the examiner approaches the anticipated relevant question, this variation of the disguised question test is sometimes called the 'peak of tension test'."

Use of GKT and POT in Foreign Nations

Although polygraph tests are given in many foreign nations, the volume and/or research is sufficient for comment on only Israel, Canada, Germany, Japan, United Kingdom, and India.

Israel began its police polygraph program with Backster and Reid techniques, and both methods include POT formats (Ansley, 1973; Backster, 1963; Ben-Ishai, 1961; Elaad & Kleiner, 1986, 1990; Inbau & Reid, 1953; Reid & Inbau, 1977). There has, however, been a reported increase in the use of GKT formats by Israeli police (Ben-Shakhar & Furedy, 1990). Israel, like India and some other nations, has a great diversity of cultures within its borders but applies polygraph testing to all of them (Cohen, 1976; Kugelmass & Lieblich, 1968; Kugelmass, Lieblich & Ben-Shakhar, 1973).

There has been a great preference for POT and GKT formats in the academic research in Israel (Ben-Shakhar, Bar-Hillel & Lieblich, 1986; Ben-Shakhar & Furedy, 1990; Ben-Shakhar, Lieblich & Kugelmass, 1975; Kugelmass, Lieblich & Bergman, 1967; Lieblich, 1974; Lieblich, Ben-Shakhar & Kugelmass, 1975), although the research has also included CQT formats (Ginton, Netzer, Elaad & Ben-Shakhar, 1982; Shterzer & Elaad, 1984).

The results of polygraph tests are inadmissible as evidence in criminal trials in Israel (Harnon, 1982; Kugelmass, 1976), however, prosecutors may be influenced by favorable test results (Cohen, 1976). In civil trials, Ben-Shakhar & Furedy (1990) report that the results of tests are admissible under stipulation. Israel is one of the few nations that has had the benefit of formal training of polygraph examiners (Ansley, 1973; Cohen, 1976). The others are the United States, Canada, Japan, and Turkey.

Canada is a bilingual nation, and many polygraph examinations are conducted in French. The Canadians have a basic polygraph training course at the Canadian Police College in Ottawa. All Canadian law enforcement examiners are trained there, and through Canadian generosity, many law enforcement examiners from the United States have received their basic training at the Canadian Police College. The course teaches a control question test method similar to zone comparison and peak of tension test formats

(Canadian Police College, 1985; Desroches & Thomas, 1984).

There has been some academic Canadian research on control question techniques (Bradley, 1988, 1989; Patrick & Iacono, 1989, 1991; Raskin & Hare, 1978), but there has been much more research on peak of tension and guilty knowledge test formats (Bradley, 1988; Bradley & Rottinger, 1992; Davidson, 1968; Day & Rouke, 1974; Forth, Hart, Hare & Harper, 1988; Forth, Stratchan & Hare, 1989; Furedy & Ben-Shakhar, 1991; Furedy, Davis & Gurevich, 1988; Iacono, Boisvenu & Fleming, 1984; Iacono, Cerri, Patrick & Fleming, 1992; Janisse & Bradley, 1980) and two studies comparing CQT and GKT formats (Bradley & Ainsworth, 1984; Bradley & Janisse, 1981). Polygraph results are not admissible in Canada (Canadian Supreme Court, 1987). However, they play a significant role in investigations.

Germany does not permit polygraph testing for law enforcement or business under any circumstances. German interest in polygraph testing was developed by observing its use by U.S. military forces in Germany. But it will never be used in Germany because their courts take the view it is contrary to their Constitution (Kaganiec, 1956; Schwabe, 1982).

In the very early years of lie detection, publication of the word association concept by Wertheimer (1906) and Wertheimer and Klein (1904) was followed by an extensive body of German publications building on Wertheimer's concept (Binswanger, 1908; Heilbronner, 1907; Hoegel, 1907; Kramer & Stern, 1906; Lederer, 1906, etc.). This work on the conceptual framework of Tatbestandsdiagnostik continued well into the 1930's (Herbold-Wootten, 1982).

Because of the lack of application in Germany, there are only a few post-war articles on lie detection (Curio & Scholz, 1991; Steller, Haenert & Eiselt, 1987; and Undeutsch, 1977). Of these, only Steller et.al. employed a GKT format in research involving the relationship of extraversion and the detection of simple deception. Using skin conductance as a measure, they found statistically significant ($p < .05$) higher scores for guilty subjects who were extraverts than the scores of guilty introverts. The detection rates for high extraversion was 100%, medium extraversion 87%, and low extraversion 67%.

Japan emphasizes the use of peak of tension and guilty knowledge test formats in criminal investigation cases (Ben-Shakhar & Furedy, 1990; Fukomoto, 1980, 1982; Nakayama & Yamamura, 1990; Nepote, 1966; Widacki, 1986; Yamamura & Miyake, 1978). The Japanese police are able to use these techniques with greater frequency than police in North America and Europe because they have complete control of the crime scene. In Japan, results of polygraph tests are of great importance as they are admissible in evidence in criminal trials (Abrams, 1973; Mito, 1969; Nepote, 1966; Takahashi, 1958, 1976; Tamiya, 1971; Yamamura & Miyati, 1990). In one case, polygraph results were the only evidence in a successful criminal prosecution (Fukumoto, 1980).

The Japanese National Police use control question test formats when necessary, and have done so for many years (Aobayashi, 1979; Hikita & Suzuki, 1963; Sagae, 1979; Suzuki, 1979; Yamamura & Miyata, 1990). In a 1975 report Suzuki said that of 2,749 cases, 1,082 (38%) were tested with known solution peak of tension, 706 (26%) were tested with a searching peak of tension test, and 961 (35%) were tested with control question tests.

For more than 30 years, the National Police have conducted polygraph research through their Laboratory, and the quality has been outstanding. Also, the training of their examiners is conducted at the Laboratory. Interesting, too, is the requirement that all examiners complete a research project before achieving senior status.

Actually, Japan's lie detection program began with galvanometers and peak of tension tests (Akamatsu, Ochida & Togawa, 1937; Imamura, 1952; Takei & Co., Ltd., nd; Togawa, Somia & Mochizak, 1950; Ureno, 1953). It is possible that the activities of the U.S. Army Crime Laboratory in Tokyo during the post-war occupation influenced the Japanese toward the use of multi-channel testing (Goddard, 1954). Familiar with the American method in using searching POTs to find evidence, the Japanese used the

technique in one case to lead them to a spot where they unearthed the victim's body (Takahashi, 1976).

The Japanese method of conducting peak of tension tests has been described by Jan Widacki (1986), a Polish examiner who visited Japan. Widacki said the test usually contains five questions, of which one is critical. As a rule it is administered four times, the first time with a one-to-five sequence of questions, the second with a five-to-one sequence, the third a mixed sequence, and a fourth with another one-to-five sequence. They try to use three or four topics, so there may be as many as 20 charts, but they are short charts.

The United Kingdom of Great Britain and Northern Ireland does not use polygraph examinations in the investigation of crime. Although they have known about tests for a long time, police have not adopted it. When a commercial company opened and offered preemployment tests, the House of Commons held hearings (Carroll, 1984; Great Britain, 1985). However the company failed and no legislation was introduced.

There was a trial program in which polygraph testing was used by the government for national security, but that has ceased (Cunningham, 1988; Jones, 1988; Norton-Taylor, 1983). At the present time there is no polygraph operation in the U.K.

The British Psychological Society is opposed to polygraph testing (Bull, 1983; Dowler, 1987; Gale, 1988), but they have no practical experience and little laboratory expertise. The only polygraph research performed in England in recent years has been the work of an Icelandic scientist, Gisli H. Gudjonsson. He has published several papers on the topic, and has made extensive use of POT/GKT test formats.

The only old reference to a real case in Great Britain is by H.J. Eysenck (1961). Writing about the "'peak of tension' or 'guilty knowledge' technique" he noted the utility in cases where a guilty person may possess knowledge which no innocent person would have. He said, "Any question regarding this knowledge, or any reference to it, would produce emotional reactions in the guilty person which would not be present in an innocent one." As an example, Eysenck mentioned a case of which he had personal knowledge, that concerned the mutilation of bed sheets in a hospital, and the use of a hundred words in a word-association test. He said the key words, such as bed-sheet, linen, cut, and bin, produced a very marked increase in autonomic activity for those guilty words by one nurse, who confessed. There were 12 other nurses tested. A "psychogalvanic reflex" was the measure, and the term suggests that the test predated the book by many years. The case represents an interesting combination of two techniques, POT and word association.

In research on emotion, Gudjonsson (1982) told subjects the questions beforehand, but not the sequence. During the test they actually read the questions to themselves and answered truthfully aloud. Two trials were performed, one with the list in one to seven order, and the other reversed. The skin resistance magnitude was converted into logarithms, and to avoid a logarithm of zero, a one was added to all resistance values. The test of 24 men disclosed a high correlation of the response magnitude to self-reported emotional disturbance. Questions were from the inoffensive "Are you sitting down?" to the offensive "Do you ever steal things?" Gudjonsson's work on personality (1977) used 12 cards with a different month on each, and the object was to pick the subject's month of birth. There were also cards with numbers and cards with words. All lists were read twice. He added to this a relevant/irrelevant test format with a combination of inoffensive irrelevant questions and offensive control questions. Gudjonsson found relationships between responsivity and some 16 P.F. measures and some Arrow-Dot measures that report on id and superego, but not ego. In a test of a hospitalized amnesia patient who did not remember her identity or her past, Gudjonsson (1979) used a searching peak of tension format for the month of her birth and her age. He also tested her ability to react with a straightforward number test. When she recalled a little more of her past a month later he used searching peak of tension for her school, and then the roads near the indicated school. Finally, using field data from the school, a list of pupils that attended were put in a list. The early tests for the month of her

birth only narrowed the choice to one of three, and the age test was not successful. However, later, when tested about roads, she gave consistent responses to only one road among ten, and it was subsequently confirmed that she lived there as a child. After the recognition of the road, tests were given again about the month of birth and age. Gudjonsson reported that at this second trial the month and age lists got specific and consistent responses which proved to be accurate. Using the list of pupils from the school, she reacted to one name, and it subsequently proved to have been her name as a pupil, even though she had changed her name twice since them. These searching peak of tension tests, neglected in much of the research literature, have great utility in solving real cases.

Gudjonsson (1983) used peak of tension tests of numbers, and skin resistance, to determine the effectiveness of countermeasures. Those who did not use a countermeasure were significantly harder to detect than those who did, a finding similar to that of Lykken's earlier work (1960). In another single-person research project (Powell, Gudjonsson & Mullen, 1983), a 36-year-old male, described as a classic case of pseudologia fantastica, was the subject of GKT tests to detect details of a mock crime. The subject was given GKT tests about knowledge of four critical items: time of the crime, means of entry, the room, and the object stolen. Electrodermal activity was the measure. The subject was told he might occasionally be shocked with moderate severity if he failed to deceive the operation, but no shocks were used. To each item in each list were four neutral items. Each list was presented four times with the sequence randomized for each presentation. Maximum GSR deflection was used for detection, and counting one tie as an error, the detection was 13 of 15, or 87%. The mean deflection for critical items was -2.76 and for non-critical items was -0.14, which was significant at $p < .002$. The personality variable did not prevent detection.

India uses the polygraph extensively in law enforcement, despite the 15 languages and variety of cultures involved (Ganguly, 1982, 1987). India began the use of polygraph examinations in 1948, after Puttappa Shivabasappa of the CID of India completed the six-week course at the Keeler Polygraph Institute. He was then an Inspector of CID in Bangalore. Shivabasappa said he was co-inventor of a polygraph used in India, and used it in narrowing down suspects in the Mahatma Ghandi assassination plot (Polygraph Student, 1948). However, after some research and a few cases, polygraph testing ceased until 1974 (Ganguly, 1987).

In addressing the American Polygraph Association in 1987, Dr. A.K. Ganguly said the results of tests conducted by police officers are not generally accepted by the courts, although there have been a few accepted; and the courts are more likely to accept the results if the test is by a person other than a police officer and for the benefit of the defense. He said they had conducted field research indicating a validity between 90% and 98% (Ganguly, 1982). Between 1974 and 1987, the Central Forensic Laboratory conducted over 3,000 examinations.

The POT or GKT test is known in India, and they have completed one research experiment with the method (Lahri & Ganguly, 1978). They conducted a simple test in which the subject took one of eight face-down cards. He looked at the card and wrote on a piece of paper the three-digit number and what a photograph depicted (bird, animal, fruit, etc.). Cards were shuffled and the subject was shown each card one-by-one, the examiner asking if it was the chosen card. The answer was "no," truth for seven, a lie to one. A field polygraph instrument was used. Half of the 80 male subjects were suspects in criminal cases randomly selected from those brought to the Central Forensic Laboratory in New Delhi. The other 40 men were government employees. The detection rate for the government employees was 28 correct (70%) and 12 incorrect. The detection rate for the criminals was 36 of 40 (90%), with errors in the other four. The extent of use of peak of tension tests or guilty knowledge tests as used in field practice has not been reported.

Diverse Courses of Development

The progress of the peak of tension test format in law enforcement has been the direct result of the strong influence of Keeler, his school, and schools following the methods

taught by the U.S. Army. This was reinforced by published accounts and informal discussions of cases where POT was successful. The use of the term peak of tension fostered the use of fixed sequences with the key in the middle, in which the subject knew the sequence. The expectation was that this would increase tension to the point of deception, followed by relief. Other POT formats were used, but fixed list was predominant.

The research community had a much less structured view and tried all kinds of formats. Following publication of Lykken's 1959 article on the guilty knowledge test, many researchers adopted both his test format and the scoring methodology. The scoring, interesting from a research viewpoint, has little practical value in law enforcement where second-best carries no weight, it is a miss. The GKT, which avoided the peaking effect in favor of specific responses did not seem advantageous to practitioners, as there was no evidence that it was more accurate than their POT, a method used with great confidence.

Researchers did use GKT formats to explore a variety of formats and theoretical questions. They continued their study of variations in answering, "yes" or mute, they considered repeating a word from the list with the answer, they studied serial effects, the effects of varying the number of control items in the list, the detection rates related to the personal significance of the key, visual versus aural presentation of questions, use of evoked potentials, and other physiological measures to detect deception. They studied a fundamental question of whether or not GKT tests could distinguish perpetrators from those who merely acquired a knowledge of the details of an event. There was also some evaluation of stimulation tests which are used in conjunction with other standardized polygraph test formats.

Validity of GKT/POT in the Field

In Japan, Yamamura and Miyake (1980) used peak of tension tests in the investigation of a riot case. They were able to establish independent ground truth in 95 cases. They were correct in their calls of DI or NDI in 85 decisions (89%). Of those 65 who were not deceptive, they were correct in 61 (94%). Of those 30 who were deceptive, they were correct in 24 (80%). When they polygraphed the deceptive subjects on details, as to which of five riot acts they committed, accuracy was 79%, testing made difficult because many subjects were guilty of more than one act.

In Israel, Elaad (1990) selected from the police files 98 sets of confirmed criminal polygraph cases in which the control question tests were followed by one to six guilty knowledge tests (mean 2) in all but three cases. In three cases there were no control question tests. Each key item had four to eight norms, excluding the opening buffer. The lists were repeated two to four times (mean 3). Forty-eight sets were from verified deceptive examinations and 50 sets were from verified truthful examinations. A blind global analysis of the GKT tests produced these results; of 50 truthful, one (2%) was scored deceptive, 46 (92%) were non-deceptive, and three (6%) were inconclusive. Excluding inconclusives, the decisions were correct in 46 of 47 decisions (98%). Of the 48 deceptive, 20 (42%) were scored deceptive, 20 were scored non-deceptive (42%), and eight (17%) were scored inconclusive. Inconclusives deleted, the decisions were correct in 20 of 40 cases (50)%. Employing unusual signal detection methods, Elaad's decisions were correct for 94% of the truthful and 65% of the deceptive.

POT and GKT Compared

Only two studies have compared elements of a GKT format with elements of a POT format. One did so in the context of stimulus tests rather than the use of mock crimes or real case material. In a study by Barland (1984), the research compared feedback with non-feedback, electrodermal recording in d.c. mode with the electrodermal recording in a.c. mode (self-centering), POT and GKT formats, and the value of each channel of data. The difference between the two formats was that in the peak of tension test the subject knew the question sequence in advance, and in the guilty knowledge test the subject did not know the question sequence in advance. In both cases Barland was testing for recognition of a picked number, a feature common to stimulus tests. Of 40 tests given, there were 25 correct

decisions, seven inconclusive, and eight errors. Excluding inconclusives, the overall accuracy was 76%. The component accuracy, excluding the inconclusives, was electrodermal a.c. 88%, electrodermal d.c. 87%, plethysmograph 36%, cardio 29%, and respiration 25%. The GKT test was correct in 15 decisions and wrong in two (88%) while the POT test was correct in ten decisions and wrong in six (62%), a difference that approached significance.

Dufek (1969) conducted two similar procedures (#2 and #4) in his research on POT, in which one group received the list of six items in random order and a similar group knew in advance the exact order in which the items would be asked. The detection rate for the random presentation group was 18 of 20 (90%), and inconclusive for two. The detection rate for the known sequence group was 17 of 20 (85%), and for the remaining three, there were reactions of equal magnitude to two items in the list, one of which was correct.

GKT and COT Compared

Podlesny, Raskin and Barland (1976) compared the accuracy of control question tests and guilty knowledge tests in testing 60 subjects about a mock crime. Excluding the 10% inconclusive outcomes, the CQT tests were correct in 89%, in error on 11%. There were no inconclusives on the GKT tests, and they were correct in 90%, in error on 10%. An independent evaluator who read these charts had an inconclusive rate of 10%, and was completely in agreement with the original examiner's determination in every case in which he made a decision.

Bradley and Ainsworth (1984), while testing to determine the effects of alcohol, tested all 40 male students with a GKT and a zone comparison, half took one first, half the other. Thirty-two played deceptive roles in a mock crime, and eight were innocent. Of those 32 guilty, there were 16 who committed the crime while intoxicated and 16 who committed it while sober. Half of each of these groups were tested while intoxicated, half were tested while sober. The truthful were sober when tested. The overall accuracy of the GKT was 95% (38 of 40), and 100% with the eight truthful. GKT was 94% (30 of 32) with the deceptive. The overall accuracy of the zone comparison (CQT) was 80% (32 of 40) and 86% (six correct, one error, one inconclusive) with the truthful. Zone was 79% (22 correct, six errors, one inconclusive) with the deceptive. Some caution in generalizing is necessary because of the intoxicated states of subjects. Incidentally, alcohol before the test did not alter accuracy, but alcohol before the crime created more false negative and inconclusive results.

Bradley and Janisse (1981) tested 192 male students, of which half committed the theft of a hidden dollar. Half the guilty and half the innocent were told they would receive a painful but not permanently damaging shock if adjudged guilty. No shocks were given. Prior to the tests for mock crime participation, each subject was given three trials of a rigged card stimulus test. Subjects were variously "detected" on none, one, two, or all three trials. For crime tests, a Backster zone comparison with theft controls was used followed by a fixed series GKT on the amount, the order being $10, $5, $1, $20, and $15. The test was administered once. The guilty all stole one dollar, the middle item in the test. Measures were pupillary response, heart rate, and skin resistance. The numerical analysis of the zone comparison charts was 80% correct, and the GKT was correct in 74% of the decisions. While these detection rates are lower than some comparable studies, two of the three physiological measures were uncommon.

Significance of the Items

One of the problems in comparing detection rates of various POT and GKT experiments is that the level of personal significance of key items and controls varies. Research has demonstrated that when two lists are used, one highly significant to the subject and one of low significance, the detection rate for the highly significant test will be greater than the detection rate for the low significance test (Dufek, 1969; Krapohl, 1984; Pinneo, Johnson & Mahoney, 1975; Stern, Breen, Watanabe & Perry, 1981). Gudjonsson (1982) also found a high correlation between electrodermal reactivity and self-reported emotional disturbances, with the more disturbing questions creating the greater responses.

Method of Presentation, Aural and Visual

Different methods have been used to present the items or numbers in laboratory tests. Beijk (1980) used a projector showing random numbers from one to ten (subject picked one) for a total of three repetitions of each question. Detection was 80% of 102 subjects. Eighty-six more undergraduate students took the same test but with one guilder (Dutch) reward if the number was not discovered. Detection was 76%. In a third test Beijk tested 40 graduate students, but substituted a tape recorded presentation for the screen. Detection was 87%. Chance for all tests was 10%. The results were not significantly different.

Carlton and Smith (1991) investigated the relative accuracy of peak of tension tests where one group received visual presentations on a computer screen and the other group received the presentations aurally. The overall accuracy of the examiner was 78%, 74% for a blind review of the charts by another examiner. Accuracy for the visual was 83% for the original examiner, 78% for the blind examiner. Accuracy for aural was 73% for the original examiner, 70% for the blind examiner. The mode of presentation did not produce a statistically significant difference.

Ben-Shakhar and Gati (1985) used electrodermal responses to evaluate four experiments, involving 30 subjects in each. Two were tests employing pictorial stimuli, two employing verbal stimuli; and the difference in the two groups in each mode of presentation was the number of common and distinctive features of the relevant and critical stimuli presented during detection trials. The results indicated detection efficiency was lower for pictorial than for verbal stimuli, and detection increased as a function of the number of common components shared by the critical and the relevant stimuli.

Can GKT Distinguish Knowing but Innocent Subjects from the Guilty?

The problem of using GKT with innocent persons who have knowledge of crime details has been investigated. Practicing examiners will not use a POT or GKT if the details are known to the subject. Nonetheless, it may be that merely knowing the correct items in lists does not create reactions of sufficient magnitude or duration to produce misleading results. If guilty knowledge tests can reliably differentiate those who committed an act from those who merely know the details, then the practical value of GKT is greatly expanded.

Geisen and Rollison (1980) investigated the ability of the GKT format with electrodermal recordings to differentiate 20 subjects who knew the key items from knowledge of a mock crime from those 20 subjects who knew the key items from reading about an award received for outstanding work. They were correct in classifying all the innocent, and all but one of the guilty (95%).

Stern, Breen, Watanabe and Perry (1981) also found that they could distinguish those who had innocent associations with the key words from those whose association with key words came from knowledge of details of a planned assassination. The research measured only electrodermal amplitude.

Mason, Johnson and Lauer (1982) reported on a study addressing knowledge and participation. In the first study the "guilty" subjects read a script detailing their rape of a woman and the other groups read about sexual intercourse with a consenting woman, but their script lacked the details in the guilty script. That first part of the study apparently provided the control information, and their detection of the truthful was 100%, and 86% for the guilty. In the second part, all subjects read a newspaper account of a rape, and the "guilty" subjects were instructed that they had committed the rape they read about, and the innocent were told they did not commit the rape but had only read about the details. They were given guilty knowledge tests in which skin resistance responses were scored. Eighty-nine percent of the "innocent" subjects were correctly classified, with two false positives. Fifty percent of the "guilty" were correctly classified. The false negatives and inconclusives were not given, nor was the number of participants.

Bradley and Rettinger (1992) using skin resistance, found that subjects who were simply aware of the key information did not obtain detection scores as high as those who perpetrated the mock crime; and the

innocent-but-aware subjects scored higher detection scores than those who were completely unaware of the key items.

Konieczny, Fras and Widacki (1984) also investigated the issue of knowledge compared to involvement. Their experiment employed 30 Polish college students, of which 15 watched an autopsy and 15 were told the details. Two peak of tension tests were conducted, one of five types of bodies (the subject of the autopsy) in which the critical item was in position four, and one of six types of bodies also in the room, with the critical item at position four. With chance for each person at 20%, they detected 80% (12 of 15) for both groups with routine tests, 93% (14 of 15) with GSR biofeedback, and 87% (13 of 15) with POT tests with no answer given.

Evoked Potentials

The first reference we find to electroencephalography and lie detection is by VonHeindl (1944) who in turn mentions work during World War II by Dr. Bernard and Professor Gelma, French psychiatrists. VonHeindl also mentions using an "electroscope," loaned to him by the great Professor Roentgen (c. 1909) for interrogation, but the electrodes were on the wrists which suggests an electrodermal, electrocardiograph, or electromyograph application, not EEG. VonHeindl reported he got a swinging of the pointer at every insidious question, particularly at every dishonest answer. There is no mention of a systematic test format.

The GKT format has been used successfully in research on lie detection with evoked potentials, particularly P300, and occasionally N400. The way the material is presented, the number of times items are shown, and the interstimulus interval differs considerably from the typical field polygraph test. However, the principle is the same. Results have been promising. See Boaz, Berry, Raney, Fischler and Shuman (1991), Farwell and Donchin (1986, 1988, 1989), Fischler, Bloom, Childers, Arroyo and Perry (1984), Fischler, Bloom, Childers, Roucos and Perry (1983), Forth, Hart, Hare and Harpur (1988), Forth, Strachan and Hare (1989), Neshige, et.al. (1981), Pinneo, Johnson and Mahoney (1975), Rosenfeld, Nasman, Whalen, Cantwell and Mazzeri (1987), and Voronin, Konovalov and Serikov (1970, 1972). EEG has also been a topic in Japanese research (Ohnishi, Tada & Tanaka, 1967; Miyake, Okita, Kohishi & Matsunaga, 1986a, b).

Mode of Answer

"Mode of answer" is the informal name of methodology in which an examinee repeats a word from the question before answering "no." The first use of it appears to be by Richard 0. Arther (1970) who has used it, taught it, and written about it as a method to improve peak of tension testing. He apparently does not use it in his control question tests. In the Arther version the subject answers with the essential word from each peak of tension test question before saying "no." For example, "Do you know if the gun used in the robbery was a Colt revolver?" Answer, "Colt, no."

In 1985 Grimsley and Yankee completed a research project for the Department of Defense in which the examinee answered with the last word in the question, then said "no." The research, performed jointly by the University of North Carolina at Charlotte and the A. Madley Corporation, involved mock screening examinations with the relevant/irrelevant technique. Use of the mode of answer increased the accuracy. Accordingly, the A. Madley polygraph school began to teach the method to students, and there are probably examiners who are using the method in the field.

In 1987, W. Michael Floyd published a study in which the mode of answer was used in real cases, and the results compared to cases when it wasn't used. Floyd's variation used the verb in the question as opposed to a descriptive word or the last word in the question. There was no discernible difference in inconclusive rates, admission rates, time of administration, or confusion by examinees. Accuracy, in the field, could not be measured.

In the laboratory, Balloun and Holmes (1979) conducted research involving student cheating and used a guilty knowledge test in which the last word of the question became the answer, but the subjects did not say "no." The last word was also the descriptive or essential word. For example, "Was it tobacco?" Answer, "Tobacco." Balloun and

Holmes tested their subjects twice, using heart rate, finger pulse volume, and skin resistance. They were correct in 11 of 18 cheaters (61%) and 14 of 16 truthful (87%) on the first test. Detection of cheaters fell significantly on the second test to three of 18 (17%), while truthful was 15 of 16 (94%).

Silent Answer Tests

A silent answer test and other no-answer tests have been used some in field testing, and occasionally in research. The principal usage employs the Reid Silent Answer Test (Reid & Inbau, 1977; Suzuki & Yatsuda, 1965), but it is not part of their peak of tension procedure. In using this test method it is almost always in conjunction with a Reid Control Question Test in which verbal answers have been given in earlier charts, and it is used when the first few charts do not clearly reveal the subject's status as truthful or deceptive. It is also used when the effort of the subject to answer causes some distortions in the tracings. With the latter, the silent answer test may be used with the first chart. It may also be used when the subject is engaging in countermeasures involving respiratory distortions. CQT and RI tests have also been administered to persons who are mute, often deaf-mute, in specific issue and screening examinations. In these cases prior agreement may be sufficient, or a nod replaces the spoken response.

There is no literature on the use of a silent answer method or a no-answer method with field applications of POT or GKT formats. There are, however, research reports on this topic. Most of them have produced detection rates above chance (Ben-Shakhar, 1977; Ben-Shakhar, Lieblich & Kugelmass, 1975; Davidson, 1977; Day & Rouke, 1974; Dufek, Widacki & Valkova, 1975; Elaad & Ben-Shakhar, 1989; Gudjonsson, 1977; Gustafson & Orne, 1963, 1964, 1965; Horneman & O'Gorman, 1985; Janisse & Bradley, 1980, Minouchi & Kimura, 1965; and Stern, Breen, Watanabe & Perry, 1981).

Konieczny, Fras and Widacki (1984) gave peak of tension tests to two groups of Polish students, one group that had watched an autopsy and one group that were told all the details, including the details that would be used in the test. Three POT tests were administered to each person: routine, no answer, and with biofeedback. The detection rates for both groups were the same for each type of test: routine detection was 12 of 15 for each group (80%), no answer 13 of 15 (87%), and biofeedback 14 of 15 (93%). Stern, Breen, Watanabe and Perry (1981) had a higher detection rate for a no-answer group than the routine group, but the experiments were so dissimilar that the difference in answering may not be significant.

Ellson (1952) used a galvanometer and eight students in which he attempted to detect the month of their birth. He broke the year into three groups of four months and asked, "Were you born in _____?" twice for each month in the group in a semi-random order for each; semi-random in that no month was repeated until the four were asked once. In this experiment the subject lied during one of the three phases of four-month groups. Each of the eight subjects were given three such tests in offsetting order for sequence, with one series answered "no," one answered "yes," and one mute. Ellson's detection rate for the eight students was four of the "no" answers, two of the "yes" answers, and one from the mute tests.

Yes Answered Tests

A "yes test" is part of the Reid technique (Reid & Inbau, 1977). It is used primarily "where the subject has tried to evade detection by distortion of the tracings" on the stimulation chart or the relevant charts. The subject is instructed to say "yes" to all questions, including the relevant questions. The Reid experience has been that subjects who lied while answering relevant questions often tried to distort their responses to the yes answered questions to make their responses look like lies. Control questions are often deleted from the format when a "yes test" chart is administered. Reactions to the "yes" answers are often genuine, because the "yes" answer is disturbing. Indeed it is this very disturbance to truthful people that is the basis for the yes-no test, now known as the Positive Control Question Test (PCQT) (Driscoss, Honts & Jones, 1987; Forman & McCauley, 1986). However, there do not appear to be any "yes" answered GKT or POT formats in field use.

Although Horneman and O'Gorman (1985) found "yes" answers in GKT test produced detection rates only at chance, other researchers have found that yes answers produced detection rates above chance (Dufek, Widacki & Valkova, 1975; Elaad & Ben-Shakhar, 1989; and Gudjonsson, 1977). Answering "yes" to the critical item and "no" to the other items also produced detection rates above chance (Ohkawa, 1963).

Where there was a comparison of detection rates for "yes" answers with "no" answers, the "no" answers provided higher rates (Elaad & Ben-Shakhar, 1989; Ellson, 1952; Furedy, Davis & Gurevich, 1988; Gustafson & Orne, 1965; Horneman & O'Gorman, 1985; Janisse & Bradley, 1980; and Ohkawa, 1963). However, one study found a higher detection rate for the "yes" answers than the "no" answers (Kugelmass, Lieblich & Bergman, 1967), but the difference was not significant.

Stimulus Tests

Stimulus tests are widely used with specific issue test formats and sometimes with screening tests. Today, all such tests are a form of the peak of tension test. They may have evolved from early examiners who wanted a norm pattern to determine the general state of arousal. The tests also served to get the equipment adjusted prior to the real test, no small consideration in the 1920s to 1940s.

The purpose of stimulus tests has been widely discussed, and no consensus has formed (Marcy, Backster, Harrelson & Reid, 1975). Those who favor the tests suggest they improve the clarity of subsequent charts, possibly because the truthful are reassured that the test works, and the deceptive become more fearful of detection. Also, the examinee becomes familiar with the testing procedure. Some examiners use the results for chart interpretation, noting the patterns at truth telling and at deception. This use is more important to those who use relevant/irrelevant tests and those who are going to use a GKT or POT format as tests to solve the issue. The examiner learns something about the subject's physiological level of arousal and ability to react, particularly at the point of deception. Finally, many examiners find it useful in detecting countermeasures as deceptive subjects don't want the test to work, and don't want the examiner to see the pattern they produce when lying (Scarce, 1978). Countermeasures occur often and their detection is useful (Magiera, 1975).

There are a great many stimulus test formats, some elaborate, some simple. Many have been described in books, journals, and particularly *Polygraph* (Abrams, 1978, 1989; Barland, 1978; Bowling, 1978; Fingerhut, 1978; Hickman, 1978; Keeler, 1931; Lovvorn, 1978; Matte, 1980; Matzke, 1972; Reid, 1952; Reid & Inbau, 1977; Scarce, 1978; and Yamashita, 1974). Most of those 1978 references are in an issue of *Polygraph* devoted to stimulus tests. Not every technique includes the use of stimulus tests, and some prominent examiners do not believe them to be useful. Backster, who used them for a while, stopped in the mid-1970s; and Raymond J. Weir, Jr., a past present of the APA, takes the conservative view, stating that a multiple series of stim tests should not be used routinely in each examination. Weir said he used them only as a last resort to prevent an inconclusive examination. Weir also advised against the use of any test that gave the appearance of parlor games or trickery (Weir, 1978).

The Reid Stim test, a rigged card test, published in the Reid and Inbau textbooks, caused much controversy for a while, although the test has been used by relatively few examiners. The criticism was sufficient that in 1975 Reid said they were modifying the test so that the examiner and examinee agreed on the card selected before the stim test. The most widely used stim test is the one taught by the DoD Polygraph Institute and its predecessor, the U.S. Army polygraph course. In that method the examinee is asked to pick a number between three and seven, and write it down. The paper he wrote it on is hung in front of him on the wall during the test. He is told to deny having picked the number in front of him. If the examinee has picked 3, 4, 6, or 7, a buffer of two numbers is placed next to the chosen number. A series of seven numbers may be used instead of five. The series is given once, in sequence, with fifteen second intervals. The test is normally given after the first relevant CQT chart. If

there is a reaction to the chosen number and another number, the other one is also discussed, as it may have been an attempt to get a reaction to the wrong number. The question wording is simple. The preparatory question is, "Regarding the number you wrote," followed by the questions in a series, "Did you write number three?", etc. The selected number should be the middle one (Decker, 1978).

Matte (1980) does not use a buffer, is blind to the number picked, and the numbers are 3, 5, 8, 10, 12, and 15. The subject picks one of the cards (blank on the back) and keeps it during the test. Afterwards he looks at the others to be sure there was a variety. The test is simply, "Did you pick card number 3?", etc., in sequence. Subject choice provides a random distribution of key numbers over a large number of cases. Matte avoided numbers 7 and 13 because they sometimes have a special meaning. Matte first tells the subject the number he reacted to, then asks to see the card.

Abrams (1989) uses the stimulation test after the first CQT chart. Abrams describes both a blind test, where the examiner truly doesn't know the number picked, and a test where the subject picks a numbered card and turns it over so they can both see it. Two padding numbers are added to the beginning of the sequence. They are numbers not represented in the deck. If there are distortions suggesting countermeasures in the first chart, Abrams uses a series of seven numbers, if not, the series is only five numbers. Padding numbers, numbers that could not have been picked, are not only at the beginning, but interspersed among the possible choices. In the sequence, where P = padding and C = possible choice, Abrams' long series is P, P, C, P, C, P, C.

Hickman (1978) uses a list of either a series of even numbers: 10, 12, 14, 16, 18, 20, or a series of colors: white, blue, orange, yellow, red; and has the subject circle any one of the items. Called a "control" chart by his students, Hickman has the test given before the relevant charts. His description does not mention any alteration of the sequence, so the item covertly picked and written down, may be in the first or last position, or anywhere else. The examiner is blind to the item until after the test. The instructions have an unusual feature. It includes, "It will be most interesting to see if you are mentally capable of defeating me during this preliminary test. What I would like you to do is envision another one of the numbers (colors) written on that piece of paper and see if you can concentrate on it to the extent that I will not know at which the actual lie took place. The reason I offer you this challenge is because I know you cannot do it. The harder you try not to think of the number (color) you actually circled, the more your thoughts are directed to that very number. I will tell you this, however, if you are capable of defeating me on this preliminary test, we will not bother with the rest of the examination. Now, do you remember the number (color) you circled? Is it clear in you mind that you are to answer 'no' to all of the questions during this test, even when you know that one of those 'no' answers is a deliberate lie?"

Lovvorn (1978) also uses a stim test in which he is blind to the number chosen. Using a list of numbers such as 31, 32, 33, 34, and 35 (avoiding those numbers if it includes the examinee's age) he asks the subject to write down one of them and not show the number to the examiner. Using a 30 and 36 to pad the beginning and end, he first runs a series asking the subject to say "no" to all, then with the same instrument still in operation he instructs the subject to answer the questions "truthfully" during the following questions. Lovvorn starts that second list with the number with the largest reaction and if the subject answers "yes," he stops there.

There has been some research on the utility of stimulation tests. Senese (1976) used polygraph charts from 30 investigations and had them reviewed by seven staff examiners at John E. Reid and Associates. Fifteen sets were from verified truthful subjects (someone else confessed) and 15 sets were from verified deceptive subjects (they subsequently confessed). The reviewing examiners who made 210 decisions did not know whether a stim chart had been given or not. Actually, all had been administered a stim chart after the first chart. First, the examiners made a determination of truth or deception solely from the first charts of those 30 sets. Their accuracy was 55.7%. A month later they evaluated the third chart in each

set, that was the chart immediately after the stimulation chart. Their accuracy was 71.4%. The inconclusive rate on the first charts was 20.5%, and 14.3% on the third charts. In addition to inconclusive calls from erratic and inconsistent responses, there was another class of "unresponsive" subjects. They were 10.5% of the first charts and 5.3% on the third. While reading the third chart yielded a greater accuracy than reading the first chart, we do not know if the stimulation chart influenced the third chart, or if the third chart would have been just as good if there had not been a stimulation chart.

Elaad and Kleiner (1986) had access to the charts involving the investigation of multiple arson in warehouses. There were 223 suspects examined, all subsequently verified innocent by the confession of a person not tested. All were control question tests and in 116 cases (51.8%), a stimulation chart was given between the first and second relevant charts. In 107 cases, 48.2% the stimulation test was omitted. Fifty sets of charts from each group were selected at random to study. At issue was whether those charts that followed a stimulation test had greater clarity than the second and following charts where the stimulation was omitted. There was no significant effect attributable to the stimulation test, or lack of the test. There was no highly significant difference in scores for any of the three indices, but the electrodermal scores were somewhat lower following the stimulation tests than when a stimulation test wasn't conducted. However, there was a slight but not significant increase in scores in the respiration measures for those who received stimulation tests.

Kirby (1981) compared the effect of two groups of stimulation tests on real cases. All tests were conducted with the Reid Control Question Test, but half were tested with the Reid card test in which the examinee does not know that the examiner knows the number on the card he selected, and a known card test in which the examinee reveals the card to the examiner prior to the stimulation chart administration. Kirby used 40 sets of confession verified specific issue charts, of which half were from deceptive and half were from truthful subjects. Half of each of these groups had received Reid stimulation tests and half received known card tests. Ten examiners were asked to make determinations of truth or deception from the first charts on each set. A month later those ten examiners read the chart after the stimulus chart, the third chart, and made a determination of truth of deception. Finally, those ten examiners, along with two more, read the stimulus charts for the 40 sets and were asked to classify their reaction to the chosen card as: 1) significant to moderate, 2) minimal/erratic, and 3) disturbed. The results of the latter were:

	Significant/ Moderate	Minimal/ Erratic	Distorted
Known Card Test - All	47%	41%	12%
Standard Card Test – All	56%	40%	4%
Known Card Test - Truthful	55%	43%	2%
Standard Card Test – Truthful	56%	44%	0
Known Card Test - Deceptive	39%	39%	22%
Standard Card Test – Deceptive	56%	36%	8%

There was a significant difference in that persons more often distorted their charts when a known card test was given than when a standard (Reid) card test was given.

Excluding inconclusive results, examiners were correct in reading the first chart at 79.5% for the known card test and 72.9% for the standard card test. The examiner accuracy for the third chart was 68.6% for the known card test and 66.1% for the standard (Reid) card test.

Except for the greater distortions following the known card test, there were no significant differences attributable to the different stimulus tests.

Horowitz, Kircher and Raskin (1986) used 100 mock crime tests to determine whether or not stimulation tests improve the accuracy of the following CQT in a laboratory setting. Using a stim test before the first relevant chart, the examinee picked a number from three to six, and was questioned about numbers one through seven, in sequence. Skin conductance amplitude was the measure from which one of the four numbers was picked. The accuracy of the stimulation test was compared to the accuracy of the control question test that followed. CQT polygraph tests were 86% correct when they followed a correct outcome of the stimulus test and 89% correct when they followed an erroneous outcome of the stimulus test. The stimulus tests were correct in 51% of the cases, which is significantly above the chance rate of 25%. The outcome of the stimulus tests did not appear to have an effect on the accuracy of the following CQT tests.

Whether or not field stimulus tests improve the accuracy of tests or reduce the inconclusive rate remains unknown. If the evidence that they improve test results is problematic, no one has introduced evidence to suggest they are counter-productive. Although there is a considerable variation in the details of their presentation, all stimulation tests represent one form or another of searching or known solution peak of tension or guilty knowledge tests. Some tests provide for precise placement of the selected number, others leave it to chance, whatever the subject chooses. In some tests the examiner knows, openly or covertly, the number chosen, in other tests the examiner is blind to the test. In one widely used test the subject sees his chosen number in a list on the wall and the sequence is known. In others the sequence is unknown to the subject. In all, stimulus tests represent a wide variety of POT/GKT formats.

Validity and Reliability of GKT/POT Test Formats

Only two studies exist that describe the field accuracy of POT and GKT tests. They are quite different. The Japanese study by Yamamura and Miyake (1978) involved known solution peak of tension tests, and for those who were deceptive, searching peak of tension tests on the specific acts suspects committed during a riot that included arson and murder. Their accuracy is based on those cases for which there were eventual verification. The results are well above chance (see Table 1). The other study is a reliability study in which the researcher in Israel drew confirmed deceptive and truthful sets of charts from police file in which one or more GKT test followed control question tests. Analyzing those GKT charts globally, blind to the status of the cases, the independent reviewer was quite accurate with the truthful but only right on half of the deceptive cases (see Table 1). There isn't enough information on these disparate research projects to arrive at a generalization.

Table 2 represents the accuracy of peak of tension tests and guilty knowledge tests conducted in a laboratory setting. While they are all placed on one table, they are so different that the totals are of little value. Whether they were POT or GKT was based on what the author called them, or if not called, what they appeared to be.

If the reader is interested in totals, despite the varied nature of the projects, see Table 3.

TABLE 1

Field Validity of POT and GKT Testing
Where Results Were Confirmed

Author	Year	% Overall Accuracy	Number of Subjects	% DI Accuracy	Number of DI Subjects	% NDI Accuracy	Number of NDI Subjects	Technique	Issues	Notes
Yamamura & Miyake	1978	89%	95	80%	30	94%	65	POT	arson & murder	one riot case
Elaad	1990	76%	87	50%	40	98%	47	GKT	criminal	blind analysis of GKTs

Notes: No inconclusives in Yamamura & Miyake
Inconclusives excluded from Elaad

TABLE 2

Peak of Tension (POT) and Guilty Knowledge Tests (GKT)

Author	Year	% Overall Accuracy	Number of Subjects	% DI Accuracy	Number of DI Subjects	% NDI Accuracy	Number of NDI Subjects	Subject Type	Incentive	Technique	Chance	Notes
Balloun & Holmes #1	79	73%	34	61%	18	87%	16	S	who cheated	GKT	17%	GSR, HR, FPV
Balloun & Holmes #2	79	53%	34	17%	18	94%	16	S	who cheated	GKT	17%	no answer
Barber	64	26%	60	26%	60	--	--	S	--	POT A	17%	field instruments
Barland	84	75%	20	75%	20	--	--	P	$15	POT	20%	SRR
Barland	84	95%	20	95%	20	--	--	P	$15	GKT	20%	SRR
Beijk	80	80%	228	80%	228	--	--	S	--	POT	10%	SCR
Ben-Shakhar, et.al.	70	77%	27	77%	27	--	--	S	--	GKT	20%	SRR
Davidson	68	87%	48	50%	12	100%	36	S	--	GKT	25%	SRR
Day & Rouke	74	44%	80	44%	80	--	--	S	--	POT	20%	SRR; 'no' answer
Diaz	85	64%	120	64%	120	--	--	P/S	$3	POT	20%	SRR
Dufek #1	69	83%	30	83%	30	--	--	P	--	POT	11%	GSR + odd
Dufek #2	69	90%	20	90%	20	--	--	P	--	POT	17%	GSR + odd
Dufek #3	69	100%	10	100%	10	--	--	P	--	POT	10%	GSR + odd
Dufek #4	69	85%	20	85%	20	--	--	P	--	POT	17%	GSR + odd
Dufek, et.al.	75	73%	30	73%	30	--	--	S	keep wine	POT	10%	field
Forman & McCauley	89	72%	40	45%	20	100%	20	S	$2 - $10	GKT		
Frese	78	51%	75	51%	30	--	--	S	--	POT	20%	field
Furedy & Ben-Shakhar	91	86%	21	86%	21	--	--	S	$ .75 + ego	GKT	20%	SCR; 'no' answer; low motivation
Furedy & Ben-Shakhar	91	48%	21	48%	21	--	--	S	$ .75 + ego	GKT	20%	SCR; 'yes' answer; low motivation

Furedy & Ben-Shakhar	91	55%	20	55%	20	--	--	S	$ .75 + ego	GKT	20%	SCR; mute answer; low motivation
Furedy & Ben-Shakhar	91	62%	21	62%	21	--	--	S	$ .75 + ego	GKT	20%	SCR; 'no' answer; high motivation
Furedy & Ben-Shakhar	91	45%	20	45%	20	--	--	S	$ .75 + ego	GKT	20%	SCR; 'yes' answer; high motivation
Furedy & Ben-Shakhar	91	55%	20	55%	20	--	--	S	$ .75 + ego	GKT	20%	SCR; mute answer; high motivation
Geldreich	41	74%	50	74%	50	--	--	S	--	POT	20%	GSR meter
Geldreich (fatigued)	41	100%	50	100%	50	--	--	S	--	POT	20%	GSR meter
Geldreich	42	86%	50	86%	50	--	--	S	electric shock	POT	20%	GSR meter
Giesen & Rollison	80	97%	40	95%	20	100%	20	S	--	GKT	20%	SRR
Gudjonsson	77	85%	123	85%	123	--	--	P	--	POT	14%	GSR meter
Gustafson & Orne	63	64%	18	64%	18	--	--	S	$1 + ego	POT	20%	SRR
Gustafson & Orne	63	28%	18	28%	18	--	--	S	--	POT	20%	SRR
Gustafson & Orne	64	48%	47	48%	47	--	--	S	--	POT	20%	SRR; no answer
Gustafson & Orne	64	69%	49	69%	49	--	--	S	--	POT	20%	SRR
Gustafsor & Orne	64	79%	24	79%	24	--	--	S	--	POT	20%	SRR; guilty person
Gustafson & Orne	64	33%	24	33%	24	--	--	S	--	POT	20%	SRR; guilty information
Gustafson & Orne	64	75%	24	75%	24	--	--	S	--	"RI" GKT	20%	SRR; guilty person
Gustafson & Orne	64	62%	24	62%	24	--	--	S	--	"RI" GKT	20%	SRR; guilty information
Gustafson & Orne	65	54%	50	54%	50	--	--	S	--	POT & GKT	20%	SRR; no answer
Gustafson & Orne	65	69%	42	69%	42	--	--	S	--	POT & GKT	20%	SRR
Horneman & O'Gorman	85	54%	121	29%	78	100%	43	S	--	POT	20%	SCR; no answer
Horneman & O'Gorman	85	64%	121	44%	78	100%	43	S	--	POT	20%	SCR
Horneman & O'Gorman	85	50%	121	22%	78	100%	43	S	--	POT	20%	SCR; answer "yes"
Horvath	78	69%	20	69%	20	--	--	S	--	POT	20%	SRR; with cuff pressure
Horvath	78	42%	20	42%	20	--	--	S	--	POT	20%	SRR; no cuff pressure
Horvath	79	52%	64	52%	64	--	--	S	—	POT	50%	SRR & PSE
Iacono, et.al.	84	91%	60	88%	45	100%	15	S	$5	GKT	20%	SCR & HR; drug no effect included
Jones & Salter	89	100%	8	100%	3	100%	3	P	--	GKT		

Keeler #1	30	95%	75	95%	75	--	--	P	--	POT	10%	BP & Pneumo; no GSR
Keeler #2	30	93%	30	93%	30	--	--	S	--	POT	33%	BP & Pneumo; no GSR
Kizaki, et.al.	76	53%	40	53%	40	--	--	S	--	POT	20%	SRR; "no" to the word
Kizaki, et.al.	76	65%	40	65%	40	--	--	S	--	POT	20%	SRR; "no" to an associated word
Konieczny, et.al.	84	80%	30	80%	30	--	--	S	--	POT	20%	normal test; autopsy details
Konieczny, et.al.	84	93%	30	93%	30	--	--	S	--	POT +	20%	personality & bio-feedback by GSR
Konieczny, et.al.	84	87%	30	87%	30	--	--	S	--	POT	20%	no answer
Krapohl	84	60%	60	60%	60	--	--	P	$5	GKT	20%	field instrument
Krapohl	84	20%	60	20%	60	--	--	P	--	GKT	20%	field instrument
Krenbergerova & Dufek	69	97%	10	97%	10	--	--	P	--	POT	6%	SRR; random "yes"s with "no"s
Kugelmas, et.al.	67	59%	27	59%	27	--	--	P/S	--	POT	17%	SRR; answer "no"
Kugelmas, et.al.	67	70%	27	70%	27	--	--	P/S	--	POT	17%	SRR; answer "yes"
Lahri & Ganguly	78	90%	40	90%	40	--	--	P	--	POT	12%	criminal suspects; field instrument
Lahri & Ganguly	78	70%	40	70%	40	--	--	P	--	POT	12%	office workers; field instrument
Lieblich, et.al.	76	62%	39	62%	39	--	--	Pr	$5 + cigarettes	POT	20%	SRR; prison inmates
Lieblich, et.al.	70	70%	44	70%	44	--	--	S	--	POT	50%	SRR; two cards
Lieblich, et.al.	70	61%	44	61%	44	--	--	S	--	POT	25%	SRR; four cards
Lieblich, et.al.	70	52%	44	52%	44	--	--	S	--	POT	12%	SRR; eight cards
Lieblich, et.al.	74	50%	8	50%	8	--	--	S	Pride	POT	20%	SCR; high motivation; intelligent can
Lieblich, et.al.	74	48%	28	48%	28	--	--	S	--	POT	20%	SCR; low motivation; ten series
Lieblich, et.al.	74	42%	20	42%	20	--	--	S	Pride	POT	20%	SCR; high motivation; + countermeasures
Lykken	59	96%	49	100%	35	86%	14	S	shock	GKT	20%	SRR
Lykken	60	100%	20	100%	20	--	--	S/P	$10	GKT	17%	SRR; medical students, psychologists; CMs ineffective
Miyake	78	63%	20	63%	20	--	--	S	--	POT	20%	SRR; eye movement, 43%; vasomotor, 47%
Moroney	72	25%	26	25%	26	--	--	S	pride	POT	10%	SRR
Ohkawa #1	63	87%	40	87%	40	--	--	?	--	POT	12%	SRR; answered "no" to theft item
Ohkawa #2	63	87%	40	87%	40	--	--	?	--	POT	12%	silent

Ohkawa #3	63	75%	40	75%	40	--	--	?	--	POT	12%	"no" answer; "yes" to correct item
Pennebaker & Chew #1	85	65%	10	65%	10	--	--	S	--	POT	20%	SRR; normal test
Pennebaker & Chew #2	85	72%	30	72%	30	--	--	S	--	POT	20%	SRR; closely watched to inhibit
Podlesny, et.al.	76	90%	60	90%	30	90%	30	P	$10	GKT	20%	research instrument
Ralloff & Johnson	88	71%	28	71%	28	--	--	?	--	GKT	10%	SCR; motor response, push a button
Ralloff & Johnson	88	86%	28	86%	28	--	--	?	--	GKT	10%	SCR; no motor response
Richardson, et.al.	90	82%	70	82%	70	--	--	M	--	POT	17%	SCR
Ruckmick	38	78%	89	78%	89	--	--	S	--	POT	10%	SRR; meter
Steller, et.al.	87	92%	87	85%	47	100%	40	Po	--	GKT	17%	SCR
Stern, et.al. #1	81	50%	48	50%	48	--	--	S	--	GKT	20%	SRR; with feedback of GSR, geometric figure
Stern, et.al. #2	81	67%	48	67%	48	--	--	S	--	GKT	20%	SRR
Stern, et.al. #3	81	88%	52	88%	52	--	--	S	--	GKT	25%	SRR; hostage/murder plot; no feedback
Stern, et.al. #4	81	96%	52	96%	52	--	--	S	--	GKT	25%	SRR; hostage/murder plot; GSR-tone feedb.
Suzuki	80	49%	24	49%	24	--	--	S	--	GKT	20%	SRR (?)
Suzuki, et.al. #1	69	60%	10	60%	10	--	--	S	--	POT	20%	SPR; no feedback
Suzuki, et.al. #2	69	70%	10	70%	10	--	--	S	--	POT	20%	SPR; feedback
Suzuki, et.al. #3	69	80%	10	80%	10	--	--	S	--	POT	20%	SPR; feedback + fake to first item
Timm	82	82%	270	82%	270	--	--	S	--	GKT	20%	SRR + respiration
Timm	89	87%	61	100%	5	86%	56	S	course credit	GKT	20%	SRR + respiration
VanBuskirk & Marcuse	54	72%	50	72%	50	--	--	S	--	POT	12%	cardio & pneumo only
Voronin, et.al. #1	72	26%	22	26%	22	--	--	Ch	--	POT	20%	8/9-yr-old children; first test
Voronin, et.al. #2	72	44%	22	24%	22	--	--	Ch	--	POT	20%	second test; HR & GSR
Voronin, et.al. #3	72	86%	22	86%	22	--	--	Ch	threat of pain	POT	20%	third test; 30 days later; + threat
Waid, et.al.	78	77%	34	79%	23	72%	11	?	--	GKT	17%	SRR
Waid, et.al.	78	71%	28	61%	18	90%	10	?	shock	GKT	25%	SRR
Waid, et.al.	78	76%	30	73%	15	80%	15	?	--	GKT	17%	SRR
Waid, et.al.	81	86%	44	82%	33	100%	11	?	pride	POT	25%	w/o meprobamate and placebo groups

Waid, et.al.	81	65%	74	55%	40	76%	30	?	pride	GKT	25%	SCR
Wakamatsu	76	60%	20	60%	20	--	--	P	1000 yen or shock	POT	20%	field instrument; w/o CM & "carefree" groups
Wakamatsu	76	55%	20	55%	20	--	--	P	pride	POT	20%	3 tests with field instrument
Wakamatsu	76	35%	20	35%	20	--	--	P	pride	POT	20%	3 tests with field instrument
Yamaoka & Suzuki	73	77%	13	77%	13	--	--	?	--	POT	20%	SPR; skin blood flow, 33%; SRR, 15%
Yamaoka & Suzuki	73	55%	31	55%	31	--	--	?	--	POT	20%	skin potential - numbers
Yamaoka & Suzuki	73	48%	31	48%	31	--	--	?	--	POT	20%	skin resistance - numbers
Yamaoka & Suzuki	73	45%	31	45%	31	--	--	?	--	POT	20%	pulse rate - numbers
Yamaoka & Suzuki	73	35%	31	35%	31	--	--	?	--	POT	20%	breathing amplitude - numbers
Yamaoka & Suzuki	73	29%	31	29%	31	--	--	?	--	POT	20%	breathing cycle time - numbers
Yamaoka & Suzuki	73	77%	31	77%	31	--	--	?	--	POT	17%	skin potential - name
Yamaoka & Suzuki	73	81%	31	81%	31	--	--	?	--	POT	17%	skin resistance - name
Yamaoka & Suzuki	73	62%	31	62%	31	--	--	?	--	POT	17%	pulse rate - name
Yamaoka & Suzuki	73	32%	31	32%	31	--	--	?	--	POT	17%	breathing amplitude - name
Yamaoka & Suzuki	73	29%	31	29%	31	--	--	?	--	POT	17%	breathing cycle time name

Abbreviations on Table 2

-- = no data
Population: S = student
P = general population
M = military
Pr = prisoners
Po = police
Ch = children
? = unstated
Notes: GSR = galvanic skin response
HR = heart rate
FPV = finger pulse volume
SRR = skin resistance response
SCR = skin conductance response
SPR = skin potential response
meter = no strip chart recording
field instrument = cardio, respiratory and electrodermal recordings
PSE = psychological stress evaluator
BP = blood pressure
pneumo = respiration
CM = countermeasure group deleted

TABLE 3

Cumulative Table

	Overall Accuracy	No. of Subjects	DI Accuracy	No. of Subjects	NDI Accuracy	No. of Subjects
All Tests	68%	4,874	65%	4,396	93%	478
Labeled GKT	76%	1,519	72%	1,181	91%	338
Other POTs	66%	3,355	65%	3,215	100%	140

Note: The only generalization one might be tempted to make from this is that POT/GKT formats may be better at detecting or supporting truthfulness than they are at detecting deception.

References

Abrams, Stanley (1973). The polygraph in Japan. *Polygraph*, 2(1), 36-41.

Abrams, Stanley (1978). The utilization and effectiveness of stimulation tests. *Polygraph*, 7(3), 178-181.

Abrams, Stanley (1989). *The complete polygraph handbook* (esp. pp. 120-122, 127). Lexington, MA: D.C. Heath & Company.

Akamatsu, P., Uchida, Y. & Togawa, Y. (1937). The measurement of the PGR (2), *Phylosophia*, 7, 171-204.

Aleksic, Zivojin L. (1972). *Naucho otkrivanje zlocina* (Scientific crime detection). Zagreb: Yugoslav League of Bar Associations.

Ansley, Norman (1951). *Notes from the Keeler Polygraph Institute*. Unpublished volume, 405 pp.

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

Aobayashi, T. (1979, Sep 20-21). *Reexamination of the Reid Control Question Test*. Paper presented at the National Convention of Police Psychology, National Research Institute of Police Science.

Arther, Richard 0. (1968). Peak of tension: Dangers. *Journal of Polygraph Studies*, 2(5), 1-4.

Arther, Richard 0. (1970). Peak of tension: Question formulation. *Journal of Polygraph Studies*, 4(5), 1-4.

Arther, Richard 0. (1982). How to word peak questions. *Journal of Polygraph Science*, 17, 1-4.

B&W Associates (1960). *B&W lie detector operator requirements for certificate of competency*. Michigan City, IN: B&W Associates.

Bach, Gavriel (1976). Admissibility of polygraph findings as legal evidence. *Crime and Social Deviance*, 6(1,2) (special issue).

Backster, Cleve (1959, Aug 16). *Interim Research Report*. Memorandum to "All Polygraph Examiners" from the Chairman of the Research and Instrumentation Committee of the Academy for Scientific Interrogation.

Backster, Cleve (1963). *Standardized polygraph notepack and technique guide: Backster zone comparison technique.* New York: Backster Institute of Lie Detection.

Baesen, Henry V., Chung, Chia-Mou & Yang, Chen-ya (1948). A lie detector experiment. *Journal of Criminal Law, Criminology and Police Science*, 39, 532-537.

Balloun, Kristen D. & Holmes, David S. (1979). Effects of repeated examinations on the ability to detect guilt with a real polygraphic examination: A laboratory experience with real crime. *Journal of Applied Psychology*, 64(3), 316-322.

Barland, Gordon H. (1978). A fail-proof blind numbers test. *Polygraph*, 7(3), 203-208.

Barland, Gordon H. (1984). *Research in electrodermal biofeedback with stimulus tests*. Contract MDA904-83-M-1150. Salt Lake City, UT: Barland & Associates.

Barland, Gordon H. & Raskin, David C. (1972). An experimental study of field techniques in "lie detection." *Polygraph*, 1(1), 22-26.

Beijk, J. (1980). Experimental and procedural influences on differential electrodermal activity. Psychophysiology, 17, 274-278.

Ben-Ishai, Akiva (1961, Sep). *Some remarks on polygraph examinations in Israel*. Paper presented at the annual meeting of the American Academy of Polygraph Examiners, Washington DC.

Ben-Shakhar, Gershon (1977). A further study of dichotomization. *Psychophysiology*, 14(4), 408-413.

Ben-Shakhar, Gershon (1992, Feb 2). Letter to the author.

Ben-Shakhar, G., Bar-Hillel, M. & Lieblich, I. (1986). Trial by polygraph: Scientific and juridical issues in lie detection. *Behavioral Sciences and the Law*, 4(4), 459-479.

Ben-Shakhar, Gershon & Gati, Itamar (1985). The effect of similarity on psychophysiological responsivity to pictorial and verbal stimuli. *Anti-Terrorism; Forensic Science; Psychology in Police Investigations* (p. 194). Jerusalem: The International Congress on Techniques for Criminal Identification.

Ben-Shakhar, Gershon & Gati, Itamar (1987). Common and distinctive features of verbal and pictorial stimuli as determinants of psychophysiological responsivity. *Journal of Experimental Psychology*, 116, 91-105.

Ben-Shakhar, G., Lieblich, I. & Kugelmass, S. (1975). Detection of information and GSR habituation: An attempt to derive detection efficiency from two habituation curves. *Psychophysiology*, 12(3), 283-288.

Binswanger, L. (1908) . Ueber das verhalten des psychogalvanischen phaenomens beim assoziationsexperiment. *Diagnostische Assoziationsstudien*, 2, 113-195. *Reprinted in Journal of Psychol. Neurol.*, 10, 11+.

Bitterman, M.E. & Marcuse, F.L. (1947). Cardiovascular responses of innocent persons to criminal interrogation. *American Journal of Psychology*, 60, 407-412.

Boaz, T.L., Perry, Jr., N.W., Raney, G., Fischler, I.S. & Shuman, D. (1991). Detection of guilty knowledge with event-related potentials. *Journal of Applied Psychology*, 76(6), 788-795.

Bowling, Melberth (1978). Comparative analysis of responses in unknown and known solution stimulus tests. *Polygraph*, 7(4), 263-265.

Bradley, M.T. (1988). Choice and the detection of deception. *Perceptual and Motor Skills*, 66, 43-48.

Bradley, M.T. (1989). Monetary incentives and the detection of deception. *Canadian Psychology*, 30(2a), 450.

Bradley, M.T. & Ainsworth, D. (1984). Alcohol and the psychophysiological detection of deception. *Psychophysiology*, 21(1), 63-71.

Bradley, M.T. & Janisse, M.P. (1981). Accuracy demonstrations, threat, and detection of deception: Cardiovascular, electrodermal, and pupillary measures. *Psychophysiology*, 18(3), 307-315. Reprinted in *Polygraph*, 10(2), 77-91.

Bradley, M.T. & Rettinger, J. (1992). Awareness of crime-relevant information and the guilty knowledge test. *Journal of Applied Psychology*, 77(1), 55-59.

Brennan, Edward P. (1992, Apr). Personal communication with Michael H. Capps.

Bull, Ray H. (1983, Jun). The truth about lie detection. *Police Review*, pp. 1190-1191 and 1246-1248.

Burack, Benjamin (1955). A critical analysis of the theory, method, and limitations of the 'lie detector'. *Journal of Criminal Law, Criminology and Police Science*, 46, 414-426.

Canadian Police College (1985). *Polygraph examination student manual*. Ottawa, Canada: Canadian Police College.

Canadian Supreme Court decides against polygraph evidence (1987, Nov/Dec). *APA Newsletter*, 20(6), 12.

Carlton, Barbara L. & Smith, Brenda J. (1991). *The effect of aural versus visual presentations of questions during a detection of deception task.* Fort McClellan, AL: Department of Defense Polygraph Institute.

Carroll, Douglas (1984). Testimony before the House of Commons Employment Committee, 1983/84 Session. In *The implications for industrial relations and employment of the introduction of the polygraph* (pp. 55-139). London: Her Majesty's Stationery Office.

Cohen, H. Victor (1976). The polygraph and research in Israel. *Polygraph*, 5(3), 235-243.

Cunningham, Cyril (1988). Vetting, investigation and interrogation. In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 40-52), Sage Publications.

Curio, I. & Scholz, O.B. (1991). Glaubhaftigkeitbeurteilung von kurzen zeugenaussagen mittels behavioraler and psychophysiologisher parameter. *Zeitschrift fur Experimentelle and Angewandte Psychologie*, 38(2), 188-200.

Davidson, P.O. (1968). Validity of the guilty knowledge technique: The effects of motivation. *Journal of Applied Psychology*, 52(1), 62-65.

Day, David A. & Rouke, Byron P. (1974). The role of attention in lie detection. *Canadian Journal of Behavior Science*, 6(3), 270-276.

Decker, Ronald E. (1978). The Army stimulation test: A controlled procedure. *Polygraph*, 7(3), 176-178.

DesRoches, Frederick J. & Thomas, Alfred S. (1984). The police use of the polygraph in criminal investigations. *Canadian Journal of Criminology*, 26(1), 43-66.

Diaz, Raul (1985). *The effects of pre-experimental expectancy, opinion, and demonstration of accuracy on the physiological detection of information*. Doctoral dissertation, Bowling Green State University.

Dowler, A. (1987). Do you believe in Santa Claus, fairies, or lie detectors? *Bulletin of the British Psychological Society*, 40, 65.

Driscoll, Laurence N., Honts, Charles R. & Jones, David (1987). The validity of the positive control physiological detection of deception technique. *Journal of Police Science and Administration*, 15(1), 46-50.

Dufek, Miroslav (1969, Feb). A contribution on the problem of polygraph examinations. *Czechoslovak Criminalistics*. Also discussed in Emotion and the polygraph, *Prokuratura* (1970, Mar).

Dufek, Miroslav, Widacki, Jan & Valkova, Vera (1975). Experimental studies of the use of the polygraph for a house search. *Archiv Med. Sad. i Krym.*, 25(2), 163-166.

Elaad, Eitan (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychology*, 75(5), 521-529.

Elaad, Eitan & Ben-Shakhar, Gershon (1989). Effects of motivation and verbal response type on psychophysiological detection of deception. *Psychophysiology*, 26(4), 422-451.

Elaad, Eitan & Kleiner, Murray, (1986). The stimulation test in field examinations: A case study. *Journal of Police Science and Administration*, 14(4), 328-333.

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

Ellson, D.G. (1952, Sep 15). A report on research on detection of deception; Office of Naval Research Contract No. N6-ONR-18011. Lafayette, IN: Indiana University.

Eysenck, H.J. (1961). *Sense and nonsense in psychology* (pp. 81-105). Baltimore, MD; Penguin Books.

Farwell, L.A. & Donchin, E. (1986). The 'brain detector': P300 in the detection of deception. *Psychophysiology*, 23(4), 434 [SPR Abstracts].

Farwell, L.A. & Donchin, E. (1988). Event-related potentials in interrogative polygraphy: Analysis using bootstrapping. *Psychophysiology*, 25(4), 445 [SPR Abstracts].

Farwell, L.A. & Donchin, E. (1989). Detection of guilty knowledge with ERPs. *Psychophysiology*, 26(4A), S8 [SPR Abstracts].

Federal Bureau of Investigation (1985). *Peak of tension (POT) test construction*. Paper distributed at the annual seminar of the American Polygraph Association, Reno, NV.

Ferrero, Gina Lombroso (1911). *The criminal man according to the classification of Cesare Lombroso.* New York: G.P. Putnam's Sons.

Fingerhut, Keith Robert (1978). Use of the stimulus test in preemployment testing. *Polygraph*, 7(3), 185-187.

Fischler, Ira, Bloom, Paul A., Childers, Donald G., Arroya, A. Antonio & Perry, Jr., Nathan W. (1984). Brain potentials during sentence verification: Late negativity and long term memory strength. *Neuropsychologia*, 22(5), 559-569.

Fischler, I., Bloom, P.A., Childers, D.G., Roucos, S.E. & Perry, Jr., N.W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, 20(4), 400-409.

Floyd, W. Michael (1987). Comparison of standard 1 yes/no 1 response and 'keyword' response ina counterintelligence-suitability polygraph examination. *Polygraph*, 16, 97-105.

Forman, Robert F. & McCauley, Clark (1986). Validity of the positive control polygraph test using the field practice model. *Journal of Applied Psychology*, 71(4), 691-698.

Forth, A.E., Hart, S.D., Hare, R.D. & Harper, T.J. (1988). Event-related brain potentials and detection of deception. *Psychophysiology*, 25(4), 446 [abstract].

Forth, A.E, Stratchan, K.E. & Hare, R.D. (1989). The use of event-related brain potentials for detection of guilty knowledge. Paper presented at the annual meeting of the Society for Psychophysiological Research, New Orleans, LA. *Psychophysiology*, 26(4A), 526 [abstract].

Frese, III, Frederick J. (1978). *General reactivity and stereotypy in detection of deception*. Doctoral dissertation, Ohio University.

Fukumoto, Junichi (1980). A case in which the polygraph was the sole evidence for conviction. *Polygraph*, 9(1), 42-44.

Fukumoto, Junichi (1982). Psychophysiological detection of deception in Japan: The past and present. *Polygraph*, 11(3), 234-238.

Furedy, John J. (1992, Jan 14). Letter to the author.

Furedy, John J. & Ben-Shakhar, Gershon (1991). The roles of deception, intention to deceive, and motivation to avoid detection in the psychophysiological detection of guilty knowledge. *Psychophysiology*, 28(2), 163-171.

Furedy, John J., Davis, Caroline & Gurevich, Maria (1988). Differentiation of deception as a psychophysiological process: A psychophysiological approach. *Psychophysiology*, 25(6), 683-688.

Gale, Anthony (Ed.) (1988). *The polygraph test: Lies, truth and science*. Sage Publications.

Ganguly, A.K. (1982). Application of polygraph in India. *Indian Journal of Criminology and Criminalistics*, 2(1), 31-32.

Ganguly, A.K. (1987). The polygraph in India. *Polygraph*, 16(3), 173-187.

Geldreich, Edward W. (1941). Studies of the galvanic skin response as a deception indicator. *Transactions of the Kansas Academy of Science*, 44, 346-351.

Geldreich, Edward W. (1941). Further studies of the use of galvanic skin response as a deception indicator. *Transactions of the Kansas Academy of Science*, 44, 279-284.

Giesen, Martin & Rollison, Michael A. (1980). Guilty knowledge versus innocent associations: Effects of trait anxiety and stimulus context on skin conductance. *Journal of Research in Personality*, 14, 7-11.

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

Goddard, Calvin H. (Col.) (1954). Some reminiscences on early days of lie detection. *Bulletin of the Academy for Scientific Interrogation*, 7(3), 11-21. Reprinted in *Polygraph*, 5(3), 252-265.

Great Britain, House of Commons, Employment Committee (1985). *Third report: The implications for industrial relations and employment of the introduction of the polygraph: Together with the proceedings for the committee*. London: Her Majesty's Stationery Office.

Grimsley, Douglas L. & Yankee, William J. (1985). *The effect of a response type in a relevant/irrelevant polygraph examination*. Contract MDA-84-R-2284, University of North Carolina at Charlotte and A. Madley Corporation.

Gudjonsson, Gisli H. (1977). *The efficacy of the galvanic skin response in experimental lie detection: Some personality variables*. Unpublished Master of Science Thesis, University of Surrey.

Gudjonsson, Gisli H. (1979). The use of electrodermal response in a case of amnesia. *Medicine, Science and the Law*, 19(2), 138-140.

Gudjonsson, Gisli H. (1982). Electrodermal responsivity to interrogation questions and its relation to self-reported emotional disturbance. *Biological Psychology*, 14, 213-218.

Gudjonsson, Gisli H. (1982). Some psychological determinants of electrodermal responses to deception. *Personality and Individual Differences*, 3, 381-391.

Gudjonsson, Gisli H. (1983). Lie detection: Techniques and countermeasures. In S.M.A. Lloyd-Bostock & B.R. Clifford (Eds.), *Evaluating witness evidence* (Chapter 8, pp. 137-155), John Wiley & Sons, Ltd.

Guertin, Wilson H. & Wilhelm, Paul L. (1954). A statistical analysis of the electrodermal response employed in lie detection. *Journal of General Psychology*, 51, 153-160.

Gustafson, Laurence A. & Orne, Martin T. (1963). Effects of heightened motivation on the detection of deception. *Journal of Applied Psychology*, 47(6) 408-411.

Gustafson, Laurence A. & Orne, Martin T. (1964). *The effect of 'lying' on 'lie detection' studies*. Paper presented at the 35th annual meeting of the Eastern Psychological Association, Philadelphia, PA.

Gustafson, Laurence A. & Orne, Martin T. (1965). Effects of perceived role and role success on detection of deception. *Journal of Applied Psychology*, 49(6), 412-417.

Gustafson, Laurence A. & Orne, Martin T. (1965). The effects of verbal response on the laboratory detection of deception. *Psychophysiology*, 2(1), 10-13.

Harnon, E. (1982, Jun). Evidence obtained by polygraph: An Israel perspective. *The Criminal Law Review,* 340-349.

Harrelson, Leonard H. (1964). *Keeler Polygraph Institute Training Guide*. Chicago: Keeler Polygraph Institute.

Heilbronner, K. (1907). Die grundlagen des "psychologischen tatbestandsdiagnostik" nebst inem praktischen fall. *Zeitschrift fuer strafrechtswissenschaft*, 27, 601-656.

Herbold-Wootten, Heidi (1982). The German Tatbestandsdiagnostik: A historical review of the beginnings of scientific lie detection in Germany. *Polygraph*, 11(3), 246-257.

Hickman, Richard C. (1978). Usefulness and theory of the stimulus test. *Polygraph*, 7(3), 182-184.

Hikita, Y. & Suzuki, A. (1963). A study of the CQT technique of the polygraph examination. *Reports of the National Institute of Police Science*, 16, 257-262.

Hoegel (1907). Die "Tatbestandsdiagnostik" im strafuerfahren. *Monatschrift fuer Kriminal Psychologie and Strafrechtsreform*, 4, 26-31.

Horowitz, Steven W., Kircher, John C. & Raskin, David C. (1986). Does stimulation test accuracy predict accuracy of polygraph tests? *Psychophysiology*, 23(4), 23 [abstract].

Horneman, Christopher & O'Gorman, J.G. (1985). Detectability in the card test as a function of the subject's verbal response. *Psychophysiology*, 22(3), 330-333. Reprinted in *Polygraph*, 15(4), 261-270.

Iacono, William G., Boisvenu, Guy A. & Fleming, Jonathan A. (1984). Effects of diazepam and methylphenidate on the electrodermal detection of guilty knowledge. *Journal of Applied Psychology*, 69(2), 289-299. Reprinted in *Polygraph*, 13, 297-312.

Iacono, William G., Cerri, Anna M., Patrick, Christopher J. & Fleming, Jonathan A. (1992). Use of antianxiety drugs as countermeasures in detection of guilty knowledge. *Journal of Applied Psychology*, 77(10), 60-64.

Imamura, Y. (1952). The investigation of techniques using psychogalvanometry. *Reports of the National Institute of Police Science*, 19, 124-132.

Inbau, Fred E. (1942/1948/1953). *Lie detection and criminal interrogation* (1st/2nd/3rd eds.). Baltimore: Williams & Wilkins.

Irving, B. & Hilgendorf, L. (1981). Police interrogation: The psychological approach. *In Royal Commission on Criminal Procedure, Research Study No. 1*. London: Her Majesty's Stationery Office.

Janisse, Michael P. & Bradley, Michael T. (1980). Deception information and the pupillary response. *Perceptual and Motor Skills*, 50, 748-750.

Jones, Ken (1988). A British trade union view. In A. Gale (Ed.) *The polygraph test: Lies, truth and science* (pp. 188-199). Sage Publications.

Kaganiec, H.J. (1956). Lie detector tests and freedom of will in Germany. *Northwestern University Law Review*, 51, 446-460.

Keeler, Leonarde (1931). Lie detector applications. In *Proceedings of the International Association of Chiefs of Police* (p. 184), Washington, DC.

Keeler, Leonarde, Muehlberger, C.W., Wilson, Charles M., Inbau, Fred E., Keeler, Katherine & O'Neill, M. Edwin (1936). *Outline of Scientific Criminal Investigation*. Ann Arbor, MI: Edwards Brothers, Inc.

Kirby, Steven R. (1981). The comparison of two stimulus tests and their effect on the polygraph technique. *Polygraph*, 10(2), 63-76.

Kizaki, Hisakazu & Yamaoka, Kazunobu (1976). Effects of preceding stimulus on physiological responses on the polygraph test. *Reports of the National Institute of Police Science*, 29(2), 33-39.

Konieczny, Jerzy, Fras, Miroslav & Widacki, Jan (1984). The specificity of so-called emotional traces and certain features of personality in the polygraph examination. *Journal of Forensic Medicine and Criminology*, 34(1), 25-30.

Kramer, F. & Stern, W. (1906). Selbstverrat durch assoziation. Experimentelle untersuchungen. *Beitrage zur Psychologie der Aussage*, 2, 457-488.

Krapohl, Donald (1984). *The detection of information of high or low personal significance using a polygraph: The effects of motivation.* Unpublished manuscript.

Krenbergerova, Jana & Dufek, Miroslav (1969, May/Jun). Experimental experiences with the use of the polygraph. *Socialist Legality*.

Kugelmass, Shiomo (1976). Admissibility of polygraph findings as legal evidence. *Crime and Social Deviance*, 6(1-2)(special issue).

Kugelmass, Sol & Lieblich, Israel (1968). Relation between ethnic origin and GSR reactivity in psychophysiological detection. *Journal of Applied Psychology*, 52, 158-162.

Kugelmass, Sol, Lieblich, Israel, Ben-Ishai, Akiva, Opatowski, Abraham & Kaplan, Maier (1968). Experimental evaluation of galvanic skin response and blood pressure change indices during criminal investigation. *Journal of Criminal Law, Criminology and Police Science*, 59, 632-635.

Kugelmass, Sol, Lieblich, Israel & Ben-Shakhar, Gershon (1973). Information detection through differential GSRs in Bedouins in the Israeli desert. *Journal of Cross-Cultural Psychology*, 4, 481-492.

Kugelmass, Sol, Lieblich, Israel & Bergman, Zev (1967). The role of 'lying' in psychophysiological detection. *Psychophysiology*, 3(3), 312-315.

Lahri, S.K. & Ganguly, A.K. (1978). An experimental study of the accuracy of polygraph technique in diagnosis of deception with volunteer and criminal suspects. *Polygraph*, 7(2), 89-94.

Law, Jr., Joseph C. (1977). Report on a new stimulation test. *Polygraph*, 6(2), 132-148.

Lederer, M. (1906). Zer frage der pathologischen tatbestandsdiagnostik. *Zeitschrift fur die Gesamte Strafrechtswissenschaft*, 26, 488-506.

Lee, Clarence D. (1943). *Instruction manual for the Berkeley Psychograph*. Sacramento, CA: Lee & Sons.

Lee, C.D. (1949). Formulating the test questions. *ISDD Bulletin*, 2(4), 1-3.

Lee, Clarence D. (1953). *The instrumental detection of deception: The lie test*. Springfield, IL: Charles C Thomas.

Lieblich, Israel (1974). Efficiency of GSR detection of information with repeated presentation of a series of stimuli in two motivational states. *Journal of Applied Psychology*, 59, 113-115.

Lieblich, Israel, Ben-Shakhar, Gershon & Kugelmass, Sol (1976). Validity of the guilty knowledge technique in a prisoner's sample. *Journal of Applied Psychology*, 61, 89-92.

Lovvorn, Donald J. (1978). A modified controlled stimulation test technique. *Polygraph*, 7(3), 188-193.

Lykken, David T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43(6), 383-388. Reprinted in *Polygraph*, 7(2), 123-128.

Lykken, David T. (1960). The validity of the guilty knowledge technique: The effect of faking. *Journal of Applied Psychology*, 44(4), 258-262. Reprinted in *Polygraph*, 7(1), 42-48.

Lykken, David T. (1992, Jan 8). Letter to the author.

Machida, K. (1949). A case report on lie detection by psychogalvanometry. *Scientific Crime Detection*, 2, 53-57.

Magiera, Alex C. (1975). Patterns of purposeful distortion. In N. Ansley (Ed.) *Legal Admissibility of the Polygraph* (pp. 199-109). Springfield, IL, Charles C Thomas.

Marcy, Lynn, Backster, Cleve, Harrelson, Leonard H. & Reid, John E. (1975). Technique panel. In N. Ansley (Ed.) *Legal Admissibility of the Polygraph* (pp. 220-254), Springfield, IL: Charles C Thomas.

Maroney, William F. & Zenhausern, Robert J. (1972). Detection of deception as a function of galvanic skin response recording methodology. *Journal of Psychology*, 80, 255-262.

Marston, William Moulton (1938). *The lie detector test*. New York: Richard R. Smith. Reprinted by the American Polygraph Association, 1989.

Mason, P.J., Johnson, H.J. & Lauer, K. (1982). Sexual stimuli and the guilty knowledge paradigm. *Psychophysiology*, 19(5), 574-575 [abstract].

Matte, James Allan (1980). *The art and science of the polygraph technique* (pp. 135-140). Springfield, IL: Charles C Thomas.

Matzke, Norman A. (1972). Sensitivity level test v. card test. *Polygraph*, 1(4), 238-240.

Minouchi, M. & Kimura, T. (1965). Response of silence and forced to answer during peak of tension tests. Research materials No. 35, *Polygraph Reports*, National Institute of Police Science, pp. 25-29.

Mito, K. (1969). Legal problems of the polygraph test. *Horitsu Jiho*, 39(2).

Miyake, Y., Okita, T., Konishi, K. & Matsunaga, I. (1986). Event-related brain potentials as an index of detection of deception. *Reports of the National Institute of Police Science*, 39(3), 19-24.

Miyake, Y., Okita, T., Konishi, K. & Matsunaga, I. (1986). Effects of self-relevancy of stimuli on event-related brain potentials as an index of detection of deception. *Reports of the National Institute of Police Science*, 39(3), 14-18.

Mulbar, H. (1944). Technique of criminal interrogation. In LeMoyne Snyder (Ed.) *Homicide Investigation* (Chapter 5). Springfield, IL: Charles C Thomas.

Munsterberg, Hugo (1907). *On the witness stand.* S.S. McClure Company. Republished by Bobbs-Merril in 1908 and Doubleday, Page & Company in 1912.

Nakayama, Majoto & Yamamura, Takehiko (1990). Changes of respiration pattern to the critical question on guilty knowledge technique. *Polygraph*, 19(3), 188-198. Abstract in *Psychophysiology*, 26(4A), 45.

Nepote, J. (1966). The use of the lie detector (polygraph test) by the Japanese police. *International Criminal Police Review*, 196, 62-66.

Neshige, R., Kuroda, Y., Kakigi, R., Fujiyama, F., Matoba, R., Yarita, M., Luders, H. & Shibasaki, H. (1991). Event-related brain potentials as indicators of visual recognition and detection of criminals by their use. *Forensic Science International*, 51, 95-103.

Norton-Taylor, Richard (1983, Sep 19). Secrets chiefs to take lie test. *Manchester Guardian*.

Ohkawa, Hisatsugi (1963). Comparison of physiological response of 'yes,' 'no,' and 'mute' conditions in peak of tension tests. *Reports of the National Institute of Police Science*, 21, 1-4.

Ohnishi, K., Tada, T. & Tanaka, Y. (1967). Study on EEG as an index of lie detector. *Reports of the National Institute of Police Science*, 20, 42-45.

Patnode, Charles H. (1956). *The lie detector in the law enforcement field.* In N. Ansley and R.J. Weir, Jr. (Eds.) Selected papers on the polygraph. Washington, DC: Board of Polygraph Examiners.

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

Patrick, Christopher J. & Iacono, William G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, 76(2), 229-238.

Pennebaker, James W. & Chew, Carol H. (1985). Behavioral inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology*, 49, 1427-1433. Reprinted in *Polygraph*, 15(4), 255-263.

Pierce, Ralph W. (1950). The peak of tension test. *ISDD Bulletin*, 3(3), 1, 4-5.

Pinneo, L.R., Johnson, P.O. & Mahoney, R.A. (1975). Biocybernetic approach to the detection of deception: A feasibility study. *Polygraph*, 4(4), 311-313.

Podlesny, John A., Raskin, David C. & Barland, Gordon H. (1976, Aug). *Effectiveness of techniques and physiological measures in the detection of deception.* Report No. 76-5 for contract 75-NI-99-0001, by the Department of Psychology, University of Utah, Salt Lake City, UT.

Polygraph student braves Winter (1948, Oct). *ISDD Bulletin*, 1(3), 2.

Rallof, D.A. & Johnson, H.J. (1988). Effects of voluntary motor response on skin conductance activity within the guilty knowledge paradigm. *Psychophysiology*, 25(4), 474 [abstract].

Ramer, Freeman B. (1949, Jan). Polygraph frees innocent man of murder charge and points out the guilty. *ISDD Bulletin*, 2(1), 2.

Raskin, David C. & Hare, R.D. (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology*, 15, 126-136.

Reid, John E. (1952, Mar 6). Psychological advantages of the card control test in lie detector examinations. In *Proceedings of the American Academy of Forensic Sciences*, Atlanta, GA.

Reid, John E. & Inbau, Fred E. (1977). *Truth and deception: The polygraph ("lie-detector") technique* (2nd ed.). Baltimore, MD: Williams & Wilkens.

Richardson, Drew C., Carlton, Barbara L. & Dutton, Donnie W. (1990). Blind analysis of skin conductance response recordings from a numbers test. *Polygraph*, 19(1), 9-20.

Rosenfeld, J. Peter, Nasman, Victoria T., Whalen, Richard, Cantwell, Brad & Mazzeri, Lisa (1987). Late vertex positivity in event-related potentials as a guilty knowledge indicator: A new method of lie detection. *International Journal of Neuroscience*, 43, 125-129. Reprinted in *Polygraph*, 16, 258-263.

Ruckmick, Christian A. (1938). The truth about the lie detector. *Journal of Applied Psychology*, 22(1), 50-58.

Sagae, Masao (1979). Effects of instruction on the physiological responses in a polygraph test. *Reports of the National Institute of Police Science*, 33(3), 164-167.

Saxe, L., Schmitz, M. & Zaichkowsky, L. (1986). *Polygraph examinations as placebos: A test of the fallibility of lie detection*. Unpublished manuscript.

Scarce, Kenneth W. (1978). The true blue control test. *Polygraph*, 7(3), 194-198.

Schwabe, Juergen (1982). Der "lugendetektor" vor dem Bundesverfassungsgericht. *Neue Juristische Wochenschrift*, 8, 367-368.

Senese, L. (1976). Accuracy of polygraph techniques with and without card test stimulation. *Journal of Police Science and Administration*, 4(3), 274-276. Reprinted in *Polygraph*, 7(3), 199-202.

Shterzer, Gideon & Elaad, Eitan (1985). Validity of the control question test in two levels of the severity of crimes. *Anti-Terrorism; Forensic Science; Psychology in Police Investigations* (pp. 155-166). Jerusalem: International Congress on Techniques for Criminal Identification.

Steller, Max, Haenert, Petra & Eiselt, Wolfgang (1987). Extraversion and the detection of information. *Journal of Research in Personality*, 21, 334-342.

Stern, Robert M., Breen, John P., Watanabe, Takami & Perry, Bradley S. (1981). Effect of feedback on physiological information on response to innocent associations and guilty knowledge. *Journal of Applied Psychology*, 66(6), 677-681.

Summers, Walter G. (Rev., S.J.) (1936). A recording psychogalvanometer. *Bulletin of the American Association of Jesuit Scientists*, 14(2), 50-56.

Summers, Walter G. (1937). A new psychogalvanometric technique in criminal investigation. *Psychological Bulletin*, 34, 551-552.

Summers, Walter G. (1938). The electric pathometer. In *Proceedings of the International Association of Chiefs of Police* (p. 142), Washington, DC.

Summers, Walter G. (1939). Science can get the confession. *Fordham Law Review*, 8, 334-354.

Suzuki, Akihiro (1975). Field polygraph examination condition and analysis of its effective procedures. *Reports of the National Institute of Police Science*, 28, 15-22.

Suzuki, Akihiro (1978). A survey of polygraph examinations in Japan. *Polygraph*, 7(4), 295-308.

Suzuki, Akihiro (1980). Effects of anxiety and perceived task difficulty on experimental detection of deception. *Reports of the National Institute of Police Science*, 33(4), 57-62.

Suzuki, Akihiro, Watanabe, Takami & Shimizu, Koichi (1969). Effects of visual feedback on the detection of deception measured by skin potential response. *Japanese Journal of Psychology*, 40(2), 58-67.

Suzuki, Akihiro & Yatsuda, J. (1965). Case study of silent answer of murder suspect in polygraph test. Research Materials No. 35, *Polygraph Reports*, pp. 17-29, National Institute of Police Science.

Takahashi, K. (1958). Legal status of polygraph results. *Science and Criminal Detection*, 11, 205-215.

Takahashi, Minoru (1976). The police operation of polygraph detection and its assessment from a judicial standpoint in Japan. *Polygraph*, 5, 223-234.

Tamiya, H. (1971). Evidentiary capacity of polygraph examination. Jurist (supplement), pp. 178-179.

Timm, Howard W. (1982). Effect of altered outcome expectancies stemming from placebo and feedback treatments on the validity of the guilty knowledge technique. *Journal of Applied Psychology*, 67(4), 391-400.

Timm, Howard W. (1989). Methodological considerations affecting the utility of incorporating innocent subjects into the design of guilty knowledge polygraph experiments. *Polygraph*, 18(3), 143-157.

Togawa, Y., Sonia, H. & Mochizok, K. (1950). A study on lie detection by PGR. *Journal of Japanese Psychology* (extra issue).

Undeutsch, Udo (1977). Die untersuchung mit dem polygraphen ('Iluegendetektor'). *Kriminalistik*, 31(5), 193-193.

Urono, F. (1953). Detection of deception on the statement using the PGR method. *Proceedings of the 16th Annual Meeting on Applied Psychology*, 68.

VanBuskirk, D. & Marcuse, F.L. (1954). The nature of errors in experimental lie detection. *Journal of Experimental Psychology*, 47(3), 187-190.

VonHeindl (1944). A device for revealing lie during questioning. *Archie fur Kriminologie*, 114, 100-101. Translated and reprinted in *Polygraph*, 8(2)(1970), 173-174.

Voronin, L.G., Konovalov, V.F. & Serikov, I.S. (1972, Aug). Physiological analysis of the determination of imprinted information in children. *Doklady Adedemii Nauk SSSR*, 205(6), 1494-1497.

Waid, William M., Orne, Emily Carota, Cook, Mary R. & Orne, Martin T. (1978). Effects of attention as indexed by subsequent memory on electrodermal detection of deception. *Journal of Applied Psychology*, 63(6), 728-733.

Waid, William M., Orne, Emily C., Cook, Mary R. & Orne, Martin T. (1981). Meprobamate reduces accuracy of physiological detection of deception. *Science*, 212, 71-73.

Waid, William M., Orne, Emily C. & Orne, Martin T. (1981). Selective memory for social information, alertness, and physiological arousal in the detection of deception. *Journal of Applied Psychology*, 66(2), 224-232.

Wakamatsu, Takeshi (1976). Effects of motivating the suspect to deceive the polygraph test. *Reports of the National Institute of Police Science*, 29(2), 99-106. Translated and reprinted in *Polygraph*, 16(2), 129-144.

Weir, Jr., Raymond J. (1978). Stimulation procedures: A conservative view. *Polygraph*, 7(3), 209-214.

Wertheimer, M. (1911). Uber die assoziationsmethoden. *Archiv fur Kriminal-Anthropologie und Kriminalistik*, 22, 293-319.

Wertheimer, M. (1935). Tatbestandsdiagnostik. In E. Abderhalden (Ed.) *Handbuch der Biologischen Arveitsmethoden*, Abt. VI, Teil, C.II., 2 Berlin, Wein.

Wertheimer, M. & Klein, J. (1904). Psychologische Tatbestandsdiagnostik. *Archiv fur Kriminal-Anthropologie und Kriminalistik*, 15, 78-113.

Widacki, Jan (1986). Polygraph testing in Japan. *Archiwum Medycyny Sadowej i Kriminologii*, 36(4), 229-233.

Wilhelm, P. & Burns, F.D., Jr. (1951). *Manual of operation for the B&W lie detector*, 3rd ed. Michigan City, IN: B&W Associates.

Wilhelm, Paul & Burns, Frank D. (1954) *Lie detection with electrodermal response*, 5th ed. Michigan City, IN: B&W Associates.

Wilkerson, Owen M. (1977). The peak of tension tests utilized in the Ashmore kidnapping. *Polygraph*, 7(1), 16-20.

Winter, John E. (1936). A comparison of the cardio-pneumopsychograph and association methods in the detection of lying in cases of thefts among college students. *Journal of Applied Psychology*, 20, 243-248.

Yamamura, Takehiko & Miyata, Yo (1990). Development of the polygraph technique in Japan for detection of deception. *Forensic Science International*, 44, 257-271.

Yamamura, Takehiko & Miyake, Yoichi (1978). Psychophysiological evaluation of detection of deception in a riot case involving arson and murder. *Reports of the National Institute of Police Science*, 31, 121-127. Translated and reprinted in *Polygraph*, 9(3), 170-181.

Yamaoka, Kazunobo & Suzuki, Akihiro (1973). Studies on skin blood flow as an index of lie detection. *Reports of the National Institute of Police Science*, 26, 206-209. Translated and published in *Polygraph*, 9(4), 232-237.

Yamaoka, Kazunobo & Suzuki, Akihiro (1973). A study of physiological measurements as indices of lie detection. *Reports of the National Institute of Police Science*, 26, 185-190.

Yamashita, S. (1974). Effects of card test on control question techniques. *Reports of the National Institute of Police Science*, 27, 258-261.

Yankee, William J. (1991, Sep 24). Memorandum: Technical correction of curriculum. Department of Defense Polygraph Institute.

Yankee, William J. & Grimsley, Douglas L. (1986). The effect of a prior polygraph test on a subsequent polygraph test. *Psychophysiology*, 24(5), 621-622 [abstract].