Polygraph

VOLUME 21 1992 NUMBER 3

Contents

PUBLISHED QUARTERLY

P.O. Box 1061, Severna Park, Maryland 21146

PERSPECTIVES ON POLYGRAPH: A GUIDE TO SURVIVAL

By

Ronald M. Furgerson

Keynote Address to the 18th Annual
Federal Interagency Polygraph Seminar

Welcome

The FBI is fortunate and proud to be a part of the Federal Polygraph Community. We're also delighted to again, for the 17th consecutive year, serve as institutional host for this 18th annual interagency course. We trust you will find this a valuable forum for your continuing education and for the exchange of information concerning developments in polygraph.

Since 1977 the FBI has enjoyed the benefit of polygraph training provided by the Department of Defense. I was privileged to be a member of the first class of FBI Agents trained at Fort McClellan 15 years ago. Since that time over 115 of our Special Agents have followed, accounting for over 40,000 polygraph examinations conducted throughout the United States and in various countries in support of the FBI's mission.

What I would like to discuss with you today is change (not the kind that you probably have in your pocket--but the kind that either frustrates you and makes your life miserable or provides spice and adventure and opportunity for growth to your life--depending largely on how you react to it--how you react to changing conditions and situations. More specifically, I want to address the change we're experiencing in three areas of polygraph:

- Changes in polygraph education
- Developments in polygraph related-research, and
- Changes in polygraph operations, both for national security and criminal investigation/law enforcement applications.

But before I start, I want to acknowledge the input from a few colleagues who helped shape my thinking this morning. Of course it is possible that after listening to my remarks you won't have detected any evidence of thinking. If that is the case, I am

---

This address by Ronald M. Furgerson, Special Agent, Federal Bureau of Investigation, was delivered to the 18th annual Federal Interagency Polygraph Seminar, Federal Bureau of Investigation Academy, Quantico, Virginia on June 8, 1992. S/A Furgerson has contributed previous articles to the journal. For reprints of this paper write to the author at the FBI Laboratory, Federal Bureau of Investigation, 10th & Pennsylvania Avenue, N.W., Washington, D.C. 20535.

responsible. But if you detect something of value, it probably came from Bill Yankee, Norm Ansley, Jim Murphy, the FBI's Polygraph Program Manager, Drew Richardson, John Podlesny, Dick Keifer, or Jay Miller, the Special assistant to the Assistant Director in Charge of the FBI Laboratory. Jay is largely responsible for the "marketing" of DNA technology within the crime laboratory community. Drew and John are on our program agenda.

Changes in Polygraph Education

Two years ago, in 1990, here at the Academy, Dr. Bill Yankee delivered the most insightful, honest, statement I had ever heard on the subject of polygraph, and the state of polygraph in the Federal Government. His speech was a blend of good news and bad news, and I believe was received by those of us here with rather mixed emotions.

He noted that our occupation has changed little over the years. We define control questions a little differently now, minor instrumentation changes have occurred, and we score charts more objectively. This is little change. Drew Richardson expressed it this way. He observed that Robert Goddard would likely be amazed to see modern space flight. And Alexander Fleming would be affected similarly with modern medicine. But, were Dr. William Marston (of 1923 Frye case fame), their most famous contemporary in lie detection, to return from the grave today, he would hardly feel he had skipped a beat. Amidst the dizzying explosion of technology, we have been status quo'ers, circling the wagons in a defensive posture.

Dr. Yankee's remarks were disquieting in other ways too. He even questioned the legitimacy of our claims to be a profession--and backed up his misgivings by describing the characteristics of a profession and our deficiencies.

The good news is that we are now witnessing dynamic change, although the full fruits of this change may not be readily apparent as yet. Most significant, perhaps, are the major changes in the philosophy and direction of polygraph education--particularly at the Department of Defense Polygraph Institute (DoDPI).

My observation is that the need for a truly more professional education is the single greatest need for us at this time. One could argue convincingly that advances in research or in application of the technique are the most important. But a quality, professional education builds the intellectual foundation which drives research and operational excellence.

We operate in an environment which is essentially scientific. Investigative experience is obviously important in the polygraph suite, but the examination process consists primarily of a clinical interview, the focus of which just happens to be investigative or personnel security concerns. While it is true that some charlatans masquerading as examiners use the polygraph as nothing more than an electronic crutch to wrench confessions out of subjects, the true professional examiner is operating in the realm of practical psychophysiology. The professional examiner must thoroughly understanding the scientific disciplines which are the underpinnings of polygraphy.

To impart the needed depth of scientific knowledge and understanding to investigators, most of whom do not have a scientific background, is not a trivial task. It suggests the possibility that future generations of examiners should be selected from the ranks of seasoned investigators having solid scientific credentials or who, through testing, can demonstrate an aptitude for scientific work.

Please do not misconstrue my remarks. I am not advocating the use of polygraph examiners in the investigative arena who are not investigators. To the contrary, I have become convinced by 22 years of service in the FBI Laboratory that the best forensic scientists, those who are the most reliable and effective in furnishing usable support to investigations, are experienced investigators. We require our examiners to have a minimum of three years' investigative experience before they can be selected to serve in the Lab. Depending upon their scientific discipline, they then receive from one to two years of specialized training to equip them for their new responsibilities. This would be an excellent model for polygraph examiners.

Under Dr. Yankee's leadership, DoDPI is working toward a curriculum which will furnish students information and concepts with transfer value. The goal is to develop in them the ability and desire to learn independently outside the classroom. Our students need this to prepare for future changes and enhance their professionalism. To accomplish this, DoDPI is improving the quality of its faculty and working to integrate the headier academics, the physiology and psychology, into the curriculum in a way that will have relevance to the examiners' operational practice. Full integration will take time and creativity.

By 1995 DoDPI hopes to have a fully implemented Master of Science Program with authorization to issue the M.S. degree and be accredited by the Southern Association of Colleges and Universities.

We should applaud Dr. Yankee's foresight and leadership and the work of his staff for their efforts in transforming DoDPI from a military-type training program, as excellent as it was, into a first rate academic institution befitting the master's degree level of professionalism.

Education is the sine qua non of professionalism.

Developments in Polygraph Research

Even with the improvements in polygraph education, polygraph remains an inviting target for those who oppose us; those in the scientific community, the American Medical Association, the American Psychological Association, the Society for Psychophysiological Research, as well as powerful opponents in Congress and in the media, who were not fully satisfied with the passage of the Employee Protection Act of 1988, because the Act left a breath of life in polygraphy. When I speak of those who oppose polygraph, I'm including those who generally support law enforcement and personnel security measures, but who are honestly and thoughtfully convinced that important decisions should not be based on polygraph findings.

I believe the time will soon arrive when, if there isn't a dramatic showing of the scientific validity of polygraph testing, we will lose its availability. Convincing demonstrations of the utility and effectiveness of polygraph will not be sufficient, alone, to save it. Likewise, the mantle of law enforcement and the intelligence community and the needs of the Federal bureaucracy alone will not shield us from extinction.

For this reason, I believe the greatest threat to polygraph is the shakiness of the scientific theory upon which it rests. The Congressional Office of Technology Assessment, in their 1983 technical memorandum, Scientific Validity of Polygraph Testing, A Research Review and Evaluation, concluded that: "The basic theory of polygraph testing is only partially developed and researched." In essence, OTA found that, while polygraph seems to work, at least in criminal investigations they could find no underlying, commonly accepted or proven theory to support polygraph. Also, they could find no credible studies validating--or offer any scientific justification for--polygraph usage in screening type examinations, which, incidentally, must account for well over half the polygraph examinations given in the Federal government. Finally, OTA concluded that a stronger theoretical base is needed for the entire range of polygraph applications. All-in-all, a less than ringing endorsement of polygraph.

I am unaware of any seminal research since the OTA report which advances a new or sounder scientific theory for polygraph. And yet, there is no doubt that deception is a cognitive activity. It has an emotional, psychological component which affects our autonomic nervous system. So the process is a phenomenon from which we should be able to draw inferences through physiological measurement. But it is all very complex and there are a tremendous number of variables to consider. There is nothing simple about polygraph or about polygraph research.

At the risk of totally alienating you, let me ask you a question. If you were testing an innocent suspect in a murder investigation, how confident would you be that the suspect would focus their psychological set on a control question such as, "Prior to 1992 did you ever intentionally hurt someone physically?" In a rape investigation, what level of confidence would you have that an innocent subject would focus on a control question such as, "Have you ever engaged in any sexual activity while you were by yourself?" It is not difficult to understand why much of the scientific community has serious reservations about polygraph.

Actually, I have every confidence in control question/psychological set theory--confidence, that is, that it works very well with some subjects, less well with others, and not at all with still others. It would be very useful if we could determine which subjects belong to each group.

Is it too lofty an ambition to want something better than control question theory and methodology as presently practiced? I have far greater confidence in the theory of guilty knowledge tests (GKTs). The GKT is intellectually more satisfying as it is an evidence-connecting examination which is not dependant on the hope that the pretest interview will be successful in properly directing the guilty or innocent examinee's attention to the relevant or control questions respectively. It also avoids ethical dilemmas which accompany the introduction and setting of control questions, where it is frequently necessary to actually encourage an examinee to lie with statements such as, "Surely you have now told me everything about (such-an-such control question activity). You don't want me to think you're a thief, or liar, or rapist, or whatever, do you?" Finally, the GKT carries with it the possibility of making meaningful statements concerning the likelihood that the examinee was truthful or deceptive.

The rub with the Guilty Knowledge Test technique is that it may not be applicable to all investigative situations--although I suspect that it has far greater applicability than most of us realize. And even when the technique has potential application,

our investigators will have to make fundamental adjustments to the way they do business, or else the key information will not be properly protected for polygraph use.

I am happy to report that DoDPI has conducted Guilty Knowledge Test research and is planning on giving it added emphasis in their training program. The Bureau (FBI) has also commenced GKT field research. It should go a long way toward determining its applicability in cases where investigators have been thoroughly indoctrinated and work closely with polygraph examiners right from the very outset of investigations.

Please permit me a brief excursion away from polygraph to discuss deoxyribonucleic acid (DNA). I believe the Bureau's (FBI's) experience in pushing the forensic use of DNA to scientific acceptability will be instructive for us.

Starting in 1987, when DNA testing began for biological evidence, defense experts (drawn from academia and the broader scientific community), attempted to disqualify DNA evidence in a series of attacks. Attacks on the scientific basis for forensic DNA testing were brief and unsuccessful because basic DNA research had been conducted at leading medical centers. By the time crime laboratories, led by the FBI Laboratory, started applying DNA technology to forensic evidence, the underlying principles were established and accepted within the scientific community.

Next, defense experts challenged the FBI's laboratory procedure, or protocol, for using DNA. Critics then attacked the qualifications of individual DNA examiners. Finally, defense experts sought to convince the courts to exclude DNA evidence by arguing that the population studies, on which the statistical estimates of DNA matches are based, have fatal design flaws. But now, both the Office of Technology Assessment (1990) and the National Academy of Sciences (1992) have confirmed the validity and scientific basis of DNA use in forensic science. Both acknowledge that the FBI's protocol, in the hands of qualified and experienced DNA examiners, produces reliable and accurate results.

While a few isolated challenges have been successful, the basic science underlying DNA, the DNA testing protocols, the qualifications of individual examiners, and the population studies have withstood the scrutiny of the scientific community and the courts. In my estimation, the science of DNA has been established and is firmly rooted. Oh, there will be continued challenges to DNA's use in court--but the challenges will be at the fringes, i.e., did the examiner use the proper procedures, or was the examiner properly trained and competent in application of DNA

technology, etc. But, in my view, the basic validity of DNA technology will never again be seriously challenged.

It will not be so easy for polygraph to be accepted because we're dealing with a more complex area. We're confronted with the seemingly infinite variability of the human psyche. The subject of our examinations plays a dynamic role in the examination process. Examinees are not locked into a set molecular pattern or sequence, such as is the case with DNA. And polygraph examinee fear or emotional arousal are not susceptible to accurate measurement in a manner similar to the way cocaine levels are measured in parts per million in a urinalysis exam.

Based on the FBI's experience introducing DNA to forensic science, general rules for convincing the scientific community and the courts of the value of technical evidence can be summarized as follows:

We must base our practice and operations on solid scientific foundations. We must be open to working with and seeking the recognition of organizations such as the Office of Technology Assessment and the National Academy of Science.

We must develop and insist that our examiners use validated, standardized, and defensible procedures and do not deviate from ethical standards of conduct.

We must insure that our examiners are truly professional. In the future they will preferably be graduates of a rigorous post graduate level course wherein they become masters of all applicable psychological and physiological aspects of polygraphy and are competent to lay and defend the scientific foundation for the use of polygraph in the physiological diagnosis of deception. They must be sufficiently knowledgeable of research design and statistics to understand and apply scientific research findings. And finally,

We must know and acknowledge our limitations and refuse to be drawn into areas where we are incompetent. When we are unsure of our capabilities, we must err on the conservative side.

Polygraph Operations

I mentioned earlier that 40,000 plus polygraph examinations have been conducted by FBI Agents since I graduated from the Basic Course in 1977. Among those 40,000 examinations are many success stories.

For example, polygraph played a key role in the detection and conviction of members of the John Walker spy ring. It was polygraph which gave credibility to the otherwise unsubstantiated allegations of Walker's former wife when she accused him of spying against the government. It was polygraph which made it worthwhile to devote the extensive resources to the investigation necessary to develop sufficient information to show probable cause and to sustain a conviction. Walker's brother, Arthur, confessed his treachery during a polygraph examination.

William Peter Kampiles, a former Central Intelligence Agency employee, was suspected of disclosing secret information on the Early Bird reconnaissance satellite. Kampiles was found deceptive and confessed during a polygraph examination. During the trial which followed, the testimony of the examiner and the polygraph confession were instrumental in convicting Kampiles of espionage.

In the early 1960s, Joseph G. Helmich, an Army Warrant Officer, sold highly classified cryptographic information and other military secrets to the Soviet Union. He too was convicted of espionage largely on the basis of his confession during a polygraph examination.

Just a couple of weeks ago a black police officer with a community college in Los Angeles reported receiving KKK hate literature in his departmental mail slot. He had been a spokesman for other black officers and drew attention to racism in their respective departments. His case received widespread local media attention in the wake of the Rodney King atmosphere and the officer angrily demanded punishment for the perpetrators. During the civil rights investigation conducted by our field office, our local examiner pushed for a polygraph examination of the officer. The officer was indignant at having his veracity questioned, but agreed to be tested. After being found deceptive, the officer confessed that he had put the literature in his own mail slot. His motivation was to draw attention to what he perceived as racism.

Polygraph identified the arsonist who torched a hotel in San Juan which cost 96 lives. It identified leftists in El Salvador who raped and murdered the Maryknoll nuns. It was used to identify Congressmen who received bribes in the Korea Gate case involving Tong Sun Park. It has been indispensable to countless bank robbery, kidnapping, extortion, drug trafficking, and financial institution fraud investigations. I know your Agencies can cite equally compelling examples of the value and utility of polygraph in your operations. Our investigations simply would not be as effective without polygraph.

Polygraph in the hands of competent, well-trained, ethical examiners has identified the guilty, exonerated the innocent and provided investigative direction. When used properly it is a highly efficient and effective way of doing business. But for us to truly realize its worth and take advantage of what it can provide, we have to understand and accept what it will not do for us. If we fail to recognize its limitations and do not work within its boundaries, polygraph can confuse, mislead and eventually bring us to disaster.

I have learned that our most successful examiners, and by successful I mean those agents who consistently furnish quality information of investigative value and in whom I have the greatest confidence, operate within the established guidelines. They have learned that some people are not proper candidates for testing, irrespective of how important the case is or how much they would like to do the examination. They have learned that in some cases they cannot draw a conclusion and that in other cases polygraph testing may not be timely, appropriate or well-advised.

Our greatest problems arise when examiners rush into cases as if polygraph were an investigative panacea. They attempt to solve every facet of every case. They are quick to offer opinions based upon little more than a compulsion to satisfy their own egos.

Our organizations have the duty to create an atmosphere conducive to examiner creativity and innovation and in which initiative is encouraged. But they also have the responsibility to insure that boundaries are set and guidelines are observed.

In conclusion, one final observation: In this audience are those who function in the realms of operations, training, and research. Much of what is amiss with polygraph is a direct result of these groups acting independently and each sharing culpability for the misdirection and stagnation that exist in some quarters of polygraph.

Those who conduct research have often forgotten that the end of their research is not to perform better simulated crimes, but to produce methodology and technology to advance field examinations.

Those who educate often appeal to authority or "the rule" as opposed to logic in regard to what they teach. We need to ensure the relevancy and full integration of what is taught with what goes on in the polygraph suite.

And we who make up the majority of this audience, the operational types, are no less guilty. In addition, to self

satisfaction and complacency with the status quo, even our focus is subject to question. We all proclaim that we are truth seekers. Yet, much of what drives us is an inordinate interest in the number of confessions we chalk up. Is not a confession merely a by-product of correctly assessing truth about guilty subjects (only a subset of our total examinee population) and not that which should be our guiding light? Our goal should be to furnish our customers the most thorough and accurate information possible. Some of that information may include confessions.

Because of the adverse consequences of not doing so, some of my comments today were critical of polygraph and focused on needed change. We all share responsibility for the status quo. We can all play an important role in advancing our profession. Every action we take in our respective roles affects the ability of others with different roles to function effectively. Not training, not research, not even operational practice, which serves as the vehicle for producing a useable product, can function as an independent agent. We must integrate these functions and honestly examine our goals and every method we use to reach those goals.

Thank you for your kind attention. Again, welcome to the FBI Academy. I hope your attendance here will be rewarding to you personally, and to your agencies.

End Notes:

1. Scientific Validity of Polygraph Testing: A Research Review and Evaluation -- A Technical Memorandum (Washington, D.C.: U.S. Congress, Office of Technology Assessment, OTA-TM-H-15, November 1983). According to the OTA report, "The most commonly accepted theory at present is that, when the person being examined fears detection, that fear produces a measurable physiological reaction when the person responds deceptively. Thus, in this theory, the polygraph instrument is measuring the fear of detection rather than deception per se. And the examiner infers deception when the physiological response to questions about the crime or unauthorized activity is greater than the response to other questions." (p. 101) OTA pointed out that various factors -- e.g., the examinee's intelligence level, psychological health, emotional stability, and belief in the "machine" -- may, at least theoretically, affect the physiological response. We could of course add cultural and language differences to OTA's list.

2. U.S. Congress, Office of Technology Assessment, Genetic Witness: Forensic Uses of DNA Tests, OTA-BA-438. (Washington, D.C.: U.S. Government Printing Office, July 1990).

3. National Research Council, DNA Technology In Forensic Science (Washington, D.C.: National Academy Press, April 1992).

* * * * * *

THE HISTORY AND ACCURACY OF GUILTY KNOWLEDGE AND PEAK OF TENSION TESTS

By

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

---

The author is a Life Member of the APA and the Editor-in-Chief of APA Publications. For reprints write to P.O. Box 794, Severna Park, MD 21146.

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 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, 1 1/2 - 2 minutes depending on extent and frequency of normal variations. The subject is then instructed to answer all question 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 was 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 O. 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 which 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 '.38 calibre 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 them, 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 trap 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 examiner 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 reaction instead or in addition to the key item. If a false key was to 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 cardio-pneumo 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 30,00 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 was 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 test, 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 CQT 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 measure 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 then 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 negative 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. These 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 O. 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 become 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 affect 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 & Dayger	88	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 Cited

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 O. (1968). Peak of tension: Dangers. Journal of Polygraph Studies, 2(5), 1-4.

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

Arther, Richard O. (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 und psychophysiologisher parameter. Zeitschrift fur Experimentelle und 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 'yes/no' response and 'keyword' response in a 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 und 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, Shlomo (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 ('luegendetektor'). 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. Archiv 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-pneumo-psychograph 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].

A RESPONSE TO WYGANT'S CRITIQUE OF THE DISCOVERY TEST

By

Stan Abrams, Ph.D.

In a recent article in this journal, James Wygant (1992) discussed a series of what he considered to be ethical issues related to the testing of sexual abusers with the discovery test. Similar criticisms related to civil rights were raised at the Oregon Polygraph Licensing Board Hearings (Abrams 1991). Of the fifty to sixty psychologists, attorneys, judges, probation officers, and polygraphists who testified, he was the only one who said anything negative about these approaches. His concern was that the civil rights of sex abusers were being violated by the use of the discovery test. Later, Judge Hap Leonard of Eugene, Oregon, stated that " ... people talk about the polygraph being an invasion of an individual's rights, that it's overly intrusive. What is forgotten is that these people are convicted people. They've been convicted of felony sex crimes, and by being convicted of felony sex crimes, these people have lost a good number of the rights that insulated them before they came into the courtroom and were found guilty ... I certainly feel comfortable as a judge in deciding whether or not somebody's going to take a polygraph examination ... to control their behavior so that those people who are threatened ... can feel a little bit safer ..."

Wygant, in his paper, took issue with the polygraph being used as a method of controlling pedophiles. " ... to proclaim 'threat' or 'control' as the purpose of a polygraph examination may be contrary to the concept of a search for truth ..." He ignored the fact that it is in learning the truth that promotes the control. If in the process of testing pedophiles to determine if they have re-offended, it also serves as a deterrent to their molesting another child, and that answers a very great need in the community. Moreover, in the process, the control also increases the likelihood of a treatment success and serves to protect society.

Wygant stated that we do not treat murderers in this manner, that is, force them to take polygraph tests, get treatment, and provide a history. He does not seem to understand that murderers generally only kill one time in the heat of anger, while pedophilia is a compulsive disorder. These people will reoffend over and over leaving children with scars that they will bear for life. The research indicates that multiple personalities, drug abusers, eating disorders, and prostitutes to name only a few have often been sexually abused as children (Finkelhor, 1986). It is because of the damage that they do and because of the compulsive nature of

---

For reprints contact the author at 1618 S.W. First, Suite 401, Portland, OR 97201.

their disorder that demands that sex offenders obtain help. If they do not, and perhaps even if they do, they will act out again. Therefore, every means in our armamentarium must be used to reduce this possibility. Recognizing this, and at the request of the Governor's Advisory Committee on Corrections, the Oregon Legislature will consider mandatory polygraph testing throughout the probationary period of sexual abusers and in those instances in which these abusers are viewed as dangerous, they will consider testing them for life.

If the deterrent value of polygraph testing is not enough, consider too the value of disclosure testing in obtaining the names of the victims so that their parents can be contacted to determine if treatment is necessary. If the juvenile offender under treatment has been abused, then charges can be brought against that perpetrator. Through this technique, a determination of how much threat the individual is to society can be made, and probation officers and therapists can adjust placement, supervision, and other controls as needed. Moreover, the specific problem areas can be ascertained so that treatment can be directed toward those areas. There are other values in breaking through the patient's denial. These people have led a life of secrecy and lying and it has become a large part of their life style. Breaking through this allows the patients to work on their problems more effectively by becoming more able to discuss them in treatment.

Mr. Wygant appears to lack some information on the psychodynamics of pedophiles, but this same lack also exists in so far as pre-employment testing is concerned. He stated that " ... there is little difference between the pre-employment tests that were so offensive to law makers and a new procedure [disclosure tests*] that could become equally controversial". I would submit that these tests were not offensive to legislators, but rather, to the unions. Since unions have such large voting blocs, they strongly influenced the legislators in enacting anti-polygraph legislation. It is highly unlikely that legislators would support a bill favoring pedophiles.

In this vein, Wygant went on to list five faults that he found in pre-employment testing that he feels exist in abuser testing. Polygraphists allow non-examiners to determine the questions to be asked. This, of course, is seen in specific testing as well and it does not relate to the technique but to the competence of the examiner. Only the polygraphist can determine which questions can be appropriately used. A second issue was that examiners asked

---

* author's addition.

personal questions unrelated to the purpose of the examination. The personal questions that he referred to in discovery testing are sexual background questions. Therapists believe that it is difficult treating someone in denial. How could patients be treated for sexual abuse if they denied that they had ever committed the acts? Without knowing about these "personal issues", how could risk be predicted or what likelihood of success in treatment would there be. The individual could have been charged with exposure and treated for that, or the therapist might never be aware that the patient was also a rapist.

Wygant's third concern related to the subject being forced to take the test involuntarily. This is an interesting issue that involves the rights of privacy and freedom from testifying against oneself. A precedent has been set by the Oregon Court of Appeals in State v. Wilson (1974) which found that the test was voluntary since probationers had the choice of prison for their crime or probation with polygraph supervision. The appellant claimed that this was no choice at all, but the appellate court disagreed with her indicating that she made that decision and had to live with it. I am, however, in agreement with the fourth point raised by Wygant, that there was not adequate research of the pre-employment testing as well. Pre-employment testing was essentially eliminated by the OTA Report (1983), not because the studies on this procedure resulted in negative findings, but because there was little evidence of its effectiveness one way or the other. Research on the disclosure test can not be initiated, however, unless the tests are administered and then evaluated. In the study by Abrams et.al. (1991), it was learned that the tests were useful in that many admissions were made. Of more importance, they were found to be valid. However, much more research needs to be conducted.

Wygant's final point was also well taken. It related to examiners being more interested in their immediate financial gains than to the long-term implications of the testing. Just as in employee testing, in Oregon and Washington sex abuser testing has become competitive and fees have been reduced to the extent that for many polygraphists, it is not financially feasible to test these people. As fees have been cut some examiners tended to short-cut their procedure and inevitably errors have been made. Others have been examining large numbers of subjects, testing six or more people a day. The age of the chart rollers has returned. It is here that we risk losing this approach, not in the legislature. We must recognize that every time a polygraphist produces an inaccurate finding, it hurts all of us. Therefore, the local and national societies and the licensing boards must set up specific requirements for testing. Movements to accomplish this are now being made in the Northwest. We must police our own.

It was also reported by Wugant that the motivation to make admissions in a pre-employment test in contrast to a discovery test is that the employees' incentive is to avoid admissions because they might not be hired, whereas, the sex abuser is at risk if he/she does not make admissions. In actuality, among the reasons for testing both of these groups, one is to obtain admissions, and admissions that are made by both groups are made to avoid failing the test.

Wygant suggested that polygraphy will suffer by association with sex abuser therapy and the penile plethysmograph. Both of these areas are rather new and sex abuse has become a big business. Inevitably, those who treat and those who administer the plethysmograph vary in ability. Washington has just begun certifying its therapists and other states will follow. Wygant stated that showing pornographic pictures to adolescents who are abusers could be controversial. He should be aware that there are programs available which do not include "pornographic pictures". Of the children I have seen, I doubt whether they would be shocked by the sexual acts that they have committed. The plethysmograph has been found to be an effective tool in helping to evaluate pedophiles. Through this approach they can determine which areas of aberrant behavior exist and also how the individual is progressing in therapy. Wygant stated that he is aware of adolescents who responded to many different stimuli; same-sex children, different-sex children, animals, and rape. So, too, do some adults because many pedophiles have many sexual aberrations and because of that, they are more difficult to treat. This, however, is important information for the therapist to have. It also is important to know that only a small portion of adolescents respond in this manner. Wygant's statement that adolescents are being tested extensively in the Northwest, is a bit misleading. While there is a movement in the direction of testing adolescents more often, it is still the adult who is mainly being examined. I am, however, completely in favor of testing adolescents because this has to provide a better opportunity of successfully treating them, as in contrast to when they become adults. Wygant sees not only that testing but the therapy as "punitive and confrontational" suggesting that they more traditional treatment methods should be used. Unfortunately, these patients are difficult to treat and even at this age they have learned to lie well. In my opinion, traditional approaches would not be very effective, and the individuals whom I have tested have often continued to offend even while they are in therapy. Without polygraphy, the therapist does not know if his patient is reoffending and if the community is safe. Therapy and testing are not punitive, but both are intrusive. Wygant should be aware that all therapy is intrusive, as a therapist, I not only

need to know what a patient does but what that patient thinks and feels regardless of their problems.

Mr. Wygant discussed the validity of the admissions made during the disclosure test. He stated that the thousands of new admissions from an individual usually turn out to be contacts in crowded places of which the victims may not even be aware. It is true that admissions do result in exposing some relatively minor areas of sexual abuse such as voyeurism and exposure, but in addition, abusers also list large numbers of children whom they have penetrated. One only has to read the newspapers to become aware of the numbers of nursery school operators, religious figures, coaches, and scout leaders, who, over the years have molested hundreds of children. This is not simply rubbing up against someone in a bus, as Wygant would have us believe. I would agree with Wygant that admissions of sex abusers are not always accurate. Some exaggerate to shock or impress, but most have to estimate how many children they have victimized because they cannot recall all of them.

In regard to the use of control questions, Wygant appears to be misinformed. Controls related to fantasy can be used, just as questions related to purposely lying to one's therapist or probation officer.

Wygant stated that sex abuse is adjudicated as a criminal activity, but as a mental illness for sentencing purposes. This is true only in so far as making treatment a condition of probation. He further stated that this dichotomy does not exist in other crimes. In contrast to his statement, it is also true with drug and alcohol use and certainly in instances in which an individual is found to not be responsible for his/her acts because of a mental disease. This is particularly important in child abuse, because if an abuser spends two years in prison and is released, he is still an abuser and will offend again. If there is any chance that therapy will be effective, then it has to be attempted, but during that period polygraphy at least will offer some degree of protection to the community through periodic testing.

One should recognize that the conditions that cause the sexual abuse of children persist and that approximately 25% of those abused become abusers themselves (Abrams 1989). Added to this, we have no clear idea as yet as to the effectiveness of treatment, which could mean that many pedophiles will continue with their abuse pattern despite treatment. Therefore, I would expect that this problem will persist, possibly even becoming greater. Thoughts of eliminating any technique that could assist in reducing this problem is nonsensical. It is my opinion, that polygraphy can

serve a very important role in coping with this problem, and we should do all that we can to support it.

References Cited

Abrams, S. (1991). Polygraph, A New Beginning. Polygraph, 20, 204-213.

Abrams, S. (1989). The Complete Polygraph Handbook. Lexington, MA: Lexington Books.

Abrams, S., Hoyt, D., and Jewell, C. (1991). The Effectiveness of the Disclosure Test with Sex Abusers of Children. Polygraph, 20, 192-203.

Finkelhor, D. (1986). A Source Book on Child Sexual Abuse. Beverly Hills, CA: Sage.

State v. Wilson, 17 Ore.App. 375, 521, P.2d 1317 (1974), cert. den. 420 U.S. 910, 95 S.Ct. 829, 42, L.Ed. 2d 89 (1975).

Scientific Validity of Polygraph Testing: A Research Review and Evaluation - A Technical Memorandum (Washington, D.C.: U.S. Congress, Office of Technology Assessment, OTA-TM-H-15, 1983).

Wygant, J.R. (1992). Full Disclosure: An Ethical Question. Polygraph, 21, 51-57.

* * * * * *