

Polygraph

VOLUME 35

2006

NUMBER 1

Contents

Interrogation Using Polygraph and its Scientific Basis Yuri Kholodny	1
A Case Study of a Polygraph Examination on an Examinee with an Implanted Defibrillator Nathan Gordon and William Fleisher	22
Effects of Audiovisual Presentations of Test Questions During Relevant-Irrelevant Polygraph Examinations and New Measures John Kircher, Dan Woltz, Brian Bell, and Paul Bernhardt	25
Comparison of Evidentiary and Investigative Decision Rules: A Replication Donald Krapohl and Barry Cushman	55
Instructions to Authors	64

Interrogation Using Polygraph and its Scientific Basis¹

Yuri Kholodny²

After many decades of negative attitude to the option of polygraph application for prevention, detection and investigation of crimes, this psychophysiological “lie detection” method used in the device has been “legalized” in the country in 1993. One year later the interrogation using polygraph (IUP) have been incorporated in the national criminalistics (science of crime detection).

A huge work on introduction of the method into the Russian law enforcement practice have been carried out throughout the ten years elapsed. The polygraph interrogation assessment from the positions of criminalistic science methodology have enabled us to prove that it is one of the particular methods of criminalistics diagnostics. Entering of polygraph interrogation into contemporary criminology has given rise to a principally new direction in criminalistic science which has obtained the name of “criminalistic polygraphology”.

The thesis that the methods recommended to be applied for prevention, detection and investigation of crimes shall meet the number of requirements where scientific justification is one of them is a steady stable one for the national criminalistics.

The specialists throughout over centennial history of practical application of “lie detection” psychophysiological method using polygraph have kept trying to offer

scientific explanation and theoretical justification for those complicated processes going on with human mind and body and permitting to reveal from them the concealed information at polygraph interrogation.

It should be acknowledged that natural scientific explanation of the mechanisms forming the ground for polygraph interrogation is one of the most crucial scientific and applied research issues faced by global community of scientists and specialists engaged in the field.

Understanding of the nature of the processes going on in human psyche at polygraph interrogation apparently raises not only scientific, theoretical interest but it is as well of principal importance from practical point of view as it allows giving “transparency” to the method, increasing effectiveness and setting out scientifically grounded limits for its application.

The issues of polygraph interrogation technique theory are not properly treated in Russian scientific and special literature. The authors of small number of native publications on polygraph interrogation issue being aware of complexity of natural scientific justification of “lie detection” psychophysiological method using polygraph either avoid referring to this issue³ or offer vulgarized presentation of some theoretical concepts borrowed mainly from foreign scientific literature⁴.

¹This article was reprinted with permission of the publishing house “Spark” and the editors of “Vestnik kriminalistiki” (“Bulletin of Science of Crime Detection”).

²Head of Department of the Institute of Criminalistics of Federal Security Service of Russian Federation, D.Sc. (Law), Ph.D. (Psychology)

³Petrov A.M., Myagkikh S.G. From the polygraphologist’s notebook. Perm: 2003. 202 p.; *Belyushina O.V., Ladchenko A.G.* Polygraph in Business. M.: NOU ShO “Bayard”, 2004. 118 p; *Zhurin S.I.* Practice and Theory of Lie Detectors Application. M.: Goryachaya Liniya - Telekom (Hot Line - Telecom), 2004. 143 p.

⁴*Varlamov V.A., Varlamov G.V.* Psychophysiology of Polygraph Tests. Krasnodar: 2000. 239 p., *Zubrilova I.S., Skrypnikov A.I.* Application of Polygraph for Detecting Crimes by the Internal Affairs Authorities. Course of Lections. M.: VNII MVD (All-Russian Research Institute of the Ministry of the Interior), 1999. 117 p

The present article shall continue the work⁵ initiated by us to present theoretical aspects of polygraph interrogation: it is devoted to analysis of various approaches to resolution of this problem suggested by foreign and native scientists.

I.

By early 90s of the XX century there have been formed several theoretical approaches abroad (mainly in the US) trying to explain with different levels of success how lie of a human proves to be detected during IUP.

These theoretical approaches “can be classified into two major classes: (a) theories that emphasize motivational and emotional factors as the important determinants of psychophysiological differentiation (e.g., feelings associated with deception, fear of the consequences of the polygraph test’s results, motivation to deceive) and (b) theories that are based on cognitive factors (e.g., knowledge and awareness of certain information, attentional mechanisms that operate while processing the questions)”⁶.

According to the opinion of the US Congress experts carried out special study of the set of issues related to use of the polygraph tests «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 deceptively⁷. This theory was named **threat-of-punishment theory** and it refers to the first one of the classes mentioned above.

With a view of opening the essence of

this theoretical concept to a greater extend, Lynn P. Marcy wrote: “The basic theory of polygraphy is that under certain circumstances, questions the truth of which may have grave consequences for the subject will stimulate the sympathetic division of autonomic nervous system and cause physiological changes which can be measured, recorded, and analyzed. For this reason, the verbal answer which is articulated by the subject may not necessarily affect the physiological responses which is demonstrated by the instrument. That is to say, if the subject is asked the question, “Did you kill X?” and he is at that time aware that he did kill X, a physiological response would likely result even if he admitted his guilt and answered in the affirmative...

If, in response to this question the subject were to untruthfully deny his complicity, the fear of discovery of the truth as he knows it will cause changes in the function of each of the systems measured and recorded by the polygraph and permit the examiner to view a visible physiological response which both in theory and as demonstrated empirically by hundreds of thousands of polygraph tests can be correlated with deception.

If the subject were truthfully denying involvement in the crime, no crisis would be present and the question would not stimulate the sympathetic nervous system into action... The absence of the responses must mean that the subject is telling the truth, whereas the presence of responses means and means only that he is withholding information which he believes to be relevant to the question put to him”⁸.

⁵Mitrichev V.S., Kholodny Yu.I. Polygraph as the Tool to Obtain Orienting Criminalistic Information // Zapiski Kriminalistov (Notes of Criminologists). 1993. Issue 1. P. 173-180.; Kholodny Yu.I., Savelyev Yu.I. The Problem of Application of Polygraph Tests: Invitation to Discussion // Psikhologichesky Zhurnal (Psy

⁶Ben-Shakhar G., Furedy J. Theories and applications in the detection of deception. A psychophysiological and international perspective. Springer-Verlag New-York Inc., 1990. P. 101.

⁷Scientific validity of polygraph testing: a research review and evaluation - a technical memorandum. Washington, DC: U.S. Congress, Office of Technology Assessment, 1983. P.6.

⁸Nagle D.E. The polygraph in employment: applications and legal considerations // Polygraph. 1985. V. 14. № 1. P. 4-5.

Somewhat different interpretation of threat-of-punishment theory was offered by Davis R., according to whom “lying is an avoidance reaction with considerably less than 100 % chance of success, but the only one with any chance of success at all. The physiologic reaction would be the consequence of an avoidance reaction which has a low probability of reinforcement, but not too low. If the theory has any validity at all it must be supported that the physiologic reaction is associated with a state of uncertainty. It does seem that a lie told with a complete certainty of its acceptance would be unlikely produce much reaction; and on the other hand we have the experimental evidence ... that a lie told with no prospect of success whatever is also poorly detected”⁹.

It is easy to note that threat-of-punishment theory and both its interpretations above are rather vulnerable. First, the threat-of-punishment theory generates skeptical attitude among the critics who believe, that “in this theory, the polygraph instrument is measuring the fear of detection rather than deception per se”¹⁰. Second, one can hardly unconditionally agree with the opinion on the single determinative role of sympathetic nervous system in development of physiologic reactions in course of the polygraph tests. It is known that by no means all changes in a body happening on psychophysiological level are caused by effect of this very element of vegetative nervous system: for example, the decrease of heart rate, happening in response to presenting significant questions to the interrogated person which is frequently observed in course of the polygraph tests, is determined not by sympathetic but by parasympathetic nervous

system¹¹. Third, the threat-of-punishment theory creates serious complexities in explaining high effectiveness of experimental researches – for example those carried out in the circumstances of guessing selected figure or card tests¹² – where the threat of “ruinous consequences” for lie to experimentalist is excluded completely. Fourth, it follows from the theory under consideration that expressiveness of physiological reaction to that or other question in the course of “lie detection” process is the function of “threat-of-punishment avoidance reaction” And “if the subject were unaware that his autonomic responses were being monitored, detection rate would be minimal”¹³. However the experimental¹⁴ research undertaken demonstrated invalidity of this assumption: in the cases when it proved to be possible to persuade the examinees that polygraph was switched off (the reactions were registered telemetrically by the removed unit) it was established that there was observed no significant aggravation of physiological reactions expressiveness.

At the same time it is noteworthy to mention that threat-of-punishment theory discovers some experimental and solid practical confirmation: as evidenced by the US Congress experts likelihood of discovering the concealed information using polygraph interrogation method is typically higher in real-life conditions rather than in laboratory¹⁵ environment.

Apart from the threat-of-punishment theory the concept whose essence was formed by Luria A.R. ideas sounded in early 20s by the same is also referred to this class of “polygraph theories”.

⁹Davis R.C. Physiological responses as a means of evaluating information // The manipulation of human behavior / A.D. Biderman & H. Zimmer (Eds.). N-Y.: Wiley, 1961. P. 163.

¹⁰Scientific validity of polygraph testing: ..., P. 6.

¹¹Janig W. Vegetative Nervous System // Human Physiology. Nervous System. M.: “Mir” (World), 1985. V. 1. P. 167-219.

¹²Hassett J. Introduction to Psychophysiology. M.: “Mir” (World), 1981. 248 p.; Horneman C., O’Gorman J. Detectability in the card test as a function of the subject's verbal response // Psychophysiology. 1985. V. 22. №. 3. P. 330-333; и др.

¹³Barland G.H., Raskin D.C. Detection of deception // Electrodermal activity in psychological research. N-Y.: Academic Press, 1973. P. 446.

¹⁴Thackray R., Orme M. A comparison of physiological indices in detection of deception // Psychophysiology. 1968. V. 4. P. 329-339.

¹⁵Scientific validity of polygraph testing: ..., P. 97.

Let us recall, that Luria A.R. while studying affect state of criminals and having summarized huge experimental material established the following: “the acute state of the trauma, complicated by the necessity of the concealing it, bound in by the fear of expressing itself, creates in the criminal a state of exceedingly acute affective tension; this tension is very probably exaggerated because the subject is under the fear of disclosing his crime; the more serious the crime, the more marked the affect, and the greater the danger of disclosing it, the more this complex is suppressed... Such a tension is undoubtedly one of the most serious factors for the criminal in the recognition of the guilt. By confession the criminal has the means to avoid the affective traces, to find an exit for the tension, to discharge that affective tonus which created within him an unbearable conflict. Conflict can eliminate this conflict and restore the personality in a certain degree to a normal state, and this is its psychophysiological significance”¹⁶.

The Luria’s ideas have been transformed into **conflict theory**, which established “that a large physiological disturbance would occur when two incompatible reaction tendencies are aroused simultaneously, such as a tendency to tell the truth and the tendency to lie about the specific incident”¹⁷. The conflict theory in general stands together with the experimental data and assertion of Davis R., that detection would be easier to carry out, the more the examinee tries to conceal their lie, have been confirmed in the studies of the number of scientists. In particular improvement of concealed stimuli discrimination when the examinees seek to “deceive the device”, i.e. at intensification of

“conflict”¹⁸ has been demonstrated in laboratory environment.

Certain researchers stating their support of this theory point out that excitement at lie caused by conflict may be characterized as an inhibitive one related to parasympathetic nervous system¹⁹ activation. The experimental data on change of heart rate and T-wave amplitude during lie action²⁰ are quoted to confirm the above.

However most of the specialists recognize that conflict theory is rather weak and caution from far-reaching conclusions. According to Davis R. if conflict stands as a ground or reason for intensive reactions which mean lie then there exists a certain danger of being misled due to more intensive reactions connected with personal emotive problems. Furthermore, from the conflict theory point of view the well known fact of arising of intensive reactions at presentation of psychically significant stimuli, when no answers are required from an examinee (so called **silent test**) and likelihood of “contenting tendencies conflict”²¹ arising is practically excluded, defies explanation.

The **conditioned response theory** grounded by the principles discovered by Pavlov I.P. at studying highest nervous activity is the last one within the “motivational emotional theories” class. “This theory rests on the assumption that the relevant question ... produce differential physiological responsibility because they were conditioned to the subject’s past experience (e.g., crime). According to this account, the more serious the crime, the stronger the reactions that would be evoked by the relevant questions”²².

¹⁶Luria A.R. The nature of human conflicts. N-Y.: Liverwright, 1932. P. 114.

¹⁷Barland G. H., Raskin D. C. Detection of deception..., P. 446.

¹⁸Gustafson L., Orme M. Effects of heightened motivation on the detection of deception // J. of Applied Psychology. 1963. V. 47. P.408-411.

¹⁹Heslegrave R. An examination of the psychological mechanisms underlying deception // Psychophysiology. 1982. V. 19. P. 323.

²⁰Heslegrave R., Furedy J., Matyas T. Studies on the validity and utility of T-wave amplitude // Psychophysiology. 1982. V. 19. P.323.

²¹Horvath F., Reid J. The polygraph silent answer test // J. of Criminal Law, Criminology and Police Science. 1972. V. 63. P. 285-293.

²²Ben-Shakhar G., Furedy J. Theories and applications ..., P. 102.

With all external simplicity and seeming obviousness this theoretical concept is presumably even more vulnerable than conflict theory. If one agrees to this theory it would not be possible to offer acceptable explanation for psychophysiological reactions to lie in course of laboratory experiments where the detection percentage is rather high (for example in course of the experiments of identification of a card chosen by an examinee).

The major problem of the theories that emphasize motivational and emotional factors "is the difficulty in accounting for significant detection rates under mild conditions, when subjects are not specifically motivated to avoid detection, when subjects are not attempting to conceal the relevant information, and even when subjects are unaware of the fact that their responses are monitored by polygraph"²³.

The theories whose grounds are formed by "cognitive factors" related to perception and processing of stimuli presented to an examinee at polygraph test try to a certain extent to eliminate the imperfection mentioned above.

So, the fourth polygraph theory is the so-called **arousal theory**: "this theory avoids use of emotions such as fear or guilt. It states that detection occurs because of the differential arousal value of the various stimuli"²⁴.

The definition of "guilty knowledge" is utilized for experimental substantiation of this theory. The essence of this definition lies in the fact that the sign of crime "for the guilty subject only, the "correct" alternative will have a special significance, an added "signal value" which will tend to produce a stronger orienting reflex than that subject will show to other alternatives"²⁵.

When commenting the concept of guilty knowledge, G. Ben-Shakhar and J. Furedy pointed out that "clearly, for subjects who do not possess the guilty knowledge, all items are equivalent and evoke regular orienting reactions that will habituate with repetitions"²⁶. Exactly this determines a "cognitive" element of arousal theory: "the emphasis here is on the fact that an individual know something, rather than on the individual's emotions, fears, conditioned responses, or deception"²⁷.

In general, this theory closely matches the results of many laboratory researches carried out in the field. In particular, application of arousal theory facilitates understanding of reasons for existence of significant differences in effectiveness of psychically significant stimuli discrimination in the circumstances of different levels of motivation²⁸. (It should be mentioned herewith, that the results of experimental researches designated to confirm arousal theory have been based usually on registration of galvanic skin reflex /GSR/ - the only physiological indicator towards which the objective quantitative evaluation of the observed reactions could have been applied by foreign researchers).

The arousal theory has not gained wide recognition among the polygraphologists. Reid G. and Inbau F., the leading American polygraphologists of the 40-70s XX century, "suggest that the arousal theory may be predominant in laboratory experimentation, but that in the field situation the fear of punishment overrides the effect of alertness and attention found in the laboratory. This distinction is used by these workers and other field examiners to explain the effectiveness of electrodermal activity in the laboratory but not in the field"²⁹.

²³Ben-Shakhar G., Furedy J. Theories and applications ..., P. 106-107.

²⁴Barland G.H., Raskin D.C. Detection of deception..., P. 447.

²⁵Lykken D. Psychology and the lie detection industry // American Psychologist. 1974. V. 29. P. 728.

²⁶Ben-Shakhar G., Furedy J. Theories and applications ..., P. 107.

²⁷Ben-Shakhar G., Furedy J. Theories and applications ..., P. 108.

²⁸Orne M.T., Theckrey R.I., Paskevitz D.A. On detection of deception. A model for the study of physiological effects of psychological stimuli // Handbook of psychophysiology.- N-Y.: Holt Inc., 1972.- P.743-785; Reid J., Inbau F. Truth and deception...; Ben-Shakhar G., Furedy J. Theories and applications ...

²⁹Barland G.H., Raskin D.C. Detection of deception..., P. 447.

The Israeli psychophysiologicalists when carrying out neutral and significant stimuli reactions research experiments have empirically³⁰ established, that “psychophysiological detection depends on the relative frequency of the relevant stimuli in the stimulus set presented to the subject”³¹ in course of the polygraph test.

The researchers offered dichotomization theory: “according to the theory, persons who choose certain (relevant) stimuli will manifest independent habituation processes to the two types of stimuli (relevant and neutral)”³² to explain the empirical rule discovered.

The founders of this theory (Lieblich I., Ben-Shakhar G. et al.) hoped that methodical principles developed on its basis would allow separating in future the complex sets of stimuli in the groups and establish their subjective significance³³ determining consistent pattern of a subject addiction to each of the groups. However the researchers faced certain contradictions throughout their experiments while relying on dichotomization theory postulates.

First, “a strict dichotomization theory predicts that in a situation where the relevant and neutral categories are equiprobable (e.g., .50), it would not be possible to differentiate between them using a psychophysiological measure. However, in most studies using such base rate conditions, the skin conductance response (SCR) evoked by a relevant stimulus was greater than the SCR to a neutral stimulus”³⁴. Second, it was revealed

that the presented stimuli rarely caused more expressive SCR reaction than the neutral stimuli presented in the same circumstances. The dichotomization theory is in general rather far from real-life polygraph tests and is applicable only for a limited set of laboratory tasks. As its founders assert “further research is needed to understand the mechanism of habituation of differential autonomic responsivity”³⁵.

At the turn of the 80s-90s the attempts of foreign scientists to form solid theoretical ground for “lie detection” psychophysiological method using polygraph have not been exhausted with the five “polygraph theories” examined above. Heslegrave R. pointed out for example: “four theories were postulated to explain the increased arousal during deception. The Amount of Information theory states that the greater arousal during deception is because more information (honest and deceptive) receives attention and pressing during deception. The Retrieval Difficulty theory states that deceptive information is more difficult to retrieve than honest information and this enhances arousal. The Novelty theory states that the enhanced arousal is because of the novel association of the unfamiliar deceptive response with the question”³⁶. The Canadian scientist finally came to the conclusion that, according to his point of view the most fruitful is the conflict theory, as namely “conflict plays the primary role during the act of deception”³⁷. However certain *pro*- and *contra*- within this “polygraph theory” have been already mentioned above.

³⁰Ben-Shakhar G. A further study of the dichotomization theory in detection of information // *Psychophysiology*. 1977. V. 14. P. 408-413; Ben-Shakhar G., Lieblich I. The dichotomization theory for differential autonomic responsivity reconsidered // *Psychophysiology*. 1982. V. 20. P. 277-281.

³¹Ben-Shakhar G., Furedy J. Theories and applications ..., P. 111.

³²Ben-Shakhar G., Lieblich I. The dichotomization theory ..., P. 277.

³³Ben-Shakhar G., Lieblich I. A psychophysiological method for detecting cognitive categories // *Psychophysiology*. 1981. V. 19. P. 306.

³⁴Ben-Shakhar G., Lieblich I. The dichotomization theory ..., P. 277.

³⁵Ben-Shakhar G., Lieblich I. The dichotomization theory ..., P. 281.

³⁶Heslegrave R. An examination of the psychological mechanisms..., P. 323.

³⁷Heslegrave R. An examination of the psychological mechanisms..., P. 323.

Thus by the early 90s of XX century “some thirteen theories have been proposed to explain why people react when they are deceptive, although none can yet account for all of the facts”³⁸.

The leading Israeli and Canadian specialists came to the similar conclusion: Ben-Shakhar G., Furedy J. stated, that “no single theory or single theoretical approach is capable of providing a full account for the data”³⁹, observed in the course of polygraph tests carried out in real-life and experimental conditions.

II.

In the early XXI century the problem of natural scientific basis for lie detection psychophysiological method using polygraph has become especially vital in the USA. The attention to this problem was conditioned by the fact that Ministry of Energy which is responsible for nuclear power stations operation has taken the decision of applying polygraph screen tests so as to ensure security when working with human resources.

As per the ministry request and under the auspices of the National Academy of Science the Committee to Review the Scientific Evidence on the Polygraph (hereinafter – the Committee) has been formed; its name reflected the task assigned to the same.

After 19 months of work the Committee consisted of several dozens of scientists, who were not engaged in polygraph researches before, elaborated an extensive review⁴⁰ of applicative and theoretical aspects of the

modern technique of polygraph application in law enforcement practice. Inter alia this review paid great attention to “polygraph theories” and examination of contemporary approaches to the nature of processes in human psyche and body which enable detecting their lie.

Within the review the Committee specialists examined and analyzed the conflict theory, conditioned response theory, threat-of-punishment theory, arousal theory, dichotomization theory at the same time having combined the three last theories in an integrated group of the **mental attitude theories**.

In addition to those mentioned the review examined the **orienting theory** suggested by the Israeli researcher Kleiner M. as the general theoretical justification for polygraph testing. The Sokolov E.N. studies of orienting response, published in foreign scientific literature in the 60s last century⁴¹ have been put as the ground for orienting theory. Basing upon the definitions of “stimulus novelty” and “orienting response” Kleiner M. tried to explain difference between human response to control and test questions in the course of their testing using polygraph.

The significant achievement of the concept suggested by the Israeli scientist proved to be the fact that it introduced the definition of “significance of the stimulus” into theoretical constructions. This originality resulted in shift at test results evaluation from the definitions of “deception indicated” and “no deception indicated” to the definitions of “significant responding” and “no significant responding”.

³⁸Barland G. The polygraph test in the USA and elsewhere // The polygraph test. Lies, Truth and Science. London.: SAGE Publications, 1988. P. 83.

³⁹Ben-Shakhar G., Furedy J. Theories and applications ..., P. 113.

⁴⁰The Polygraph and Lie Detection. Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2003. 398 p.

⁴¹Sokolov E. N. Perception and the Conditioned Reflex. N-Y.: McMillan, 1963; Sokolov E. N. Orienting reflex as information regulator // Cognition and Categorization / Rosch E. and Lloyds B. (eds). Hillsdale, NJ: Lawrence Erlbaum Associates, 1966. P. 79-98.

Meanwhile the Committee specialists did not agree to Kleiner M.⁴² opinion that orienting theory may serve as comprehensive natural scientific justification for polygraph tests technique.

First of all they paid attention to the fact that “the practice of previewing questions with examinees is problematic under orienting theory”⁴³.

Second, proceeding from orienting theory the comparison questions should have been constructed in absolutely different manner: “instead of designing them to induce reactions in nondeceptive subjects, they would probably be designed to be nonevocative, as they are in the relevant-irrelevant technique”⁴⁴. The Committee specialists in generally “do not take very seriously the argument that ... polygraph examination procedures based on comparison question technique can be justified in terms of orienting theory”⁴⁵.

It is noteworthy to mention that carrying out review of “polygraph theories” the Committee specialists analyzed not all theoretical concepts available nowadays within global practice. In particular a very interesting theoretical concept of Polish polygraphologists which was named by its authors as **memory traces identification** concept proved to be out of the Committee focus.

The Polish researchers basing upon the accumulated experience of polygraph tests came to the conclusion over twenty years ago, “that the American theory of detection of deception did not provide sufficient explanation for the physiological phenomena registered during the tests”⁴⁶.

Throughout the next years the Polish researchers developed their own theoretical concept of human testing using polygraph. From their point of view such test “consists of four basic elements: 1. The character of this examination is to reproduce memory traces; 2. Examination procedure takes account of the principles used in psychological experiments; 3. The examination is a method of criminological identification; 4. The examination is aimed at retrieving information needed by law enforcement agencies. The above concept is based on the assumption of revealing memory traces of criminal offences”⁴⁷.

Although this theoretical concept proceeding from the article published can not claim to be exhaustive one in nature, it is nevertheless noteworthy to point out the most important achievement by Polish researchers, as per our opinion, which was made by them by the opening of XXI century: it was for the first time directly stated in the foreign scientific literature on the issue of lie detection psychophysiological method that polygraph test carries out human memory examination aimed at revealing existence (or absence) of traces of events having criminal relevant meaning.

However let us come back to the review elaborated by the Committee specialists.

The analysis done by the Committee specialists lead them to the same conclusion Davis R.⁴⁸ came to over 40 years ago: “it is possible that different theories are applicable in different situation. The dichotomization

⁴²Kleiner M. Physiological detection of deception in psychological perspectives: a theoretical proposal // Handbook of Polygraph Testing. San-Diego: Academic Press, 2002. P. 127-182.

⁴³The Polygraph and Lie Detection..., P. 77.

⁴⁴The Polygraph and Lie Detection..., P. 77.

⁴⁵The Polygraph and Lie Detection..., P. 77.

⁴⁶Krzyżewcin A. The Debate Over Polygraph in Poland // Polygraph. 2000. V. 29. № 3. P. 227.

⁴⁷Krzyżewcin A. The Debate Over Polygraph in Poland..., P. 227.

⁴⁸Davis R.C. Physiological responses ...

and orienting theories, for instance, may be more applicable to tests in which the signal value of the stimulation is more pertinent than the threat of severe consequences of detection: for example, when an investigation is aimed at identifying witnesses with knowledge about an incident even if they are innocent. The conflict, set, punishment, and arousal theories, in contrast, may be more applicable for identifying individuals guilty of serious crimes or those hiding dangerous plans or associations⁴⁹.

In 1973 Barland G. and Raskin D. stated "it is unfortunately that so little theoretically oriented research has been conducted during the past half century that the polygraph has been used as a "lie detector". Fortunately, this situation seems to be changing..."⁵⁰. In 1983, the experts, when performing under the US Congress assignment the analysis of situation in the field of polygraph application in various spheres of life of American society, came to the conclusion that in order to build the comprehensive "polygraph theory" it is necessary that "basic polygraph research should consider the latest research from the field of psychology, physiology, psychiatry, neuroscience and medicine"⁵¹.

20 years later – in 2003 – the Committee specialists had been also forced to state that "a solid theoretical base is necessary to have confidence in tests for the psychophysiological detection of deception... For the most part, polygraph research has focused on a few physiological responses for which measures have been available since at least the 1920s and tried to make the best of them by testing variations of them in practice,

without doing nothing much to develop the underlying science... There has been no systematic effort to identify the best potential physiological indicators on theoretical grounds or to update theory on the basis of emerging knowledge in psychology or physiology"⁵².

The Committee to Review the Scientific Evidence on the Polygraph finally came to the conclusion that "the theoretical rationale for the polygraph is quite weak"⁵³.

III.

The theoretical aspects of psychophysiological method of "lie detection" with use of polygraph draw attention of the scientists in the USSR for the first time in late 60s⁵⁴. Analysis of scientific and other information from overseas lead to the conclusion, that the "polygraph theories" existent at that period did not offer sufficient explanation of the nature of the phenomena observed in real and laboratory environment. In the light of the above it was proposed to examine the polygraph test process from the positions of informational theory of emotions proposed by the Academician Simonov P.V.⁵⁵ in 1965.

According to informational theory of emotions there exist stable dependence of the psychic tension grade upon demand ratio and difference between required and available information. At the same time the emotion itself acts as "reflection by the human brain ... of any actual demand (its quality and ratio) and probability (possibility) for its satisfaction evaluated by the brain basing upon... previously acquired individual experience"⁵⁶.

⁴⁹The Polygraph and Lie Detection..., P. 77.

⁵⁰Barland G.H., Raskin D.C. Detection of deception..., P. 471.

⁵¹Scientific validity of polygraph testing..., P. 106.

⁵²The Polygraph and Lie Detection..., P. 92.

⁵³The Polygraph and Lie Detection..., P. 213.

⁵⁴Kholodny Yu.I. Use of polygraph for prevention, detection and investigation of crimes. M: Mir bezopasnosti (World of Security), 2000. 160 p.

⁵⁵Simonov P.V. On the Role of Emotions in Adaptive Behavior of Living Systems /// Voprosy psikhologii (Problems of Psychology) 1965, No. 4, P. 75-84.

⁵⁶Simonov P.V. Emotional Brain (Physiology, Neuroanatomy, Psychology of Emotions) M.: Nauka (Science), 1981. 221 p.

Proceeding from the aforementioned point Simonov P.V. deduced the rule of emotions emerging which was expressed by the following structural formula:

$$\mathbf{E} = \mathbf{f} [- \mathbf{P}, (\mathbf{I}_n - \mathbf{I}_s), \dots]$$

where: **E**- emotions, their grade, quality and sign;

P - power and quality of actual demand;

(I_n - I_s) - evaluation of probability (possibility) for satisfaction of the demand basing upon acquired experience;

I_n - information on the means which are predictively necessary to satisfy the demand;

I_s - information on the means available with the subject at the moment⁵⁷.

Simonov P.V. and Zanicheva A.A. formulated the first native concept of interrogation using polygraph by 1970⁵⁸.

According to the concept, the need to conceal the information known to the person under examination and not to manifest their selective attitude to a particular fact or event (be it a criminal offense or the card chosen in the course of game/stimulating test) is the direct goal of the person being examined.

It was presumed thereat that the person under examination was not confident in their opportunities and did not know what happened to their physiological measures evaluated by the polygrapher when registering reactions. This causes deficit of pragmatic information for the person examined and leads to emerging of spontaneous emotional reactions accompanied by changes in physiological functions dynamics.

Basing on demand category the theoretical concept gave rather sufficient psychological explanation of human opportunities with use of polygraph, first of all in the field and laboratory environment and secondly in the cases when the verbal replies were not required from the examined person.

Progressiveness of the aforementioned concept lied in the fact that informational

theory of emotions being its fundament was first to point out principally new – **neurophysiological** – approach to studying “lie detection” method psychophysiological mechanisms and named basic brain structures (neocortex, hippocampus, amygdale) that participate in genesis of the reactions registered in the course of interrogation using polygraph.

The theoretical concept proposed by Simonov P.V. and Zanicheva A.A. along with the advantages mentioned above was not recognized as comprehensive as it was not in a position to give explanation to the number of the facts observed during IUP. It is particularly difficult to explain existence with the same person of virtually same physiological reactions at presenting test questions in the course of real or game like examination (when a card is to be guessed), although the needs of the examined person to conceal their awareness of these facts – signs of criminal act or the card chosen – would be definitely different.

After the specialized laboratory for IUP had been established in summer 1975 in the USSR the natural science grounding of permissibility to use the device for the purposes of detecting from the human the information concealed by them has gained its relevance as well.

The scientists and specialists pointed their attention to the fact that “polygraph theories” being formed by the early 80-s of the last century had the notional, descriptive nature and thus were not able to serve as the solid basis for fundamental scientific studies of the phenomenology of obtaining information from a human with use of polygraph.

The analysis of foreign “polygraph theories” and Simoniv-Zanicheva concept demonstrated that their formation would follow basically the same scheme: first a thesis, formulated on the basis of empiric observations (threatening punishment theory,

⁵⁷Simonov P.V. Higher Nervous Activity. Motivational and Emotional Aspects. M.: Nauka (Science), 1975. 173 p.

⁵⁸Theoretical researches in the field by Simonov P.V. and Zanicheva A.A. were not published in scientific literature due to the extremely negative attitude to polygraph, existent in that period in the USSR.

as an example) or scientifically established statements (conditioned reflex theory, as an example), was declared and then to confirm the thesis the facts from real or experimental (laboratory) practice of “lie detection” psychophysiological method application were selected.

Finally there was made a conclusion that such an approach was deadlock: it was not possible to offer theoretical grounding for psychophysiological method which is a polygraph interrogation method, by selecting one or another hypothesis basing upon the physiology data available without summarizing huge number of experimental data accumulated by psychophysiology and neurophysiology. It is possible to form the integrated theory capable of explaining set of facts stably observed at IUP of a human only in consequence to research of psychological phenomena in their correlation to the brain neurophysiologic activity mechanisms.

Emotions – the special class of mental processes and statuses, related to demands and motives, reflecting in the form of direct experiences the significance of particular situations for a human providing influence on the same in the course of their vital activity – are the psychological “center” of the threatening punishment theory⁵⁹ and starting point for Simonov-Zanicheva theoretical concept.

Psychological science has established long ago that emotions reflect human’s evaluative attitude to the situations, being formed or being possible, to their activity and/or to their expressions⁶⁰ in these situations.

Carrying out IUP of a person with the purpose of revealing concealed information

from them takes place always at certain psychological tension. As per the opinion of the persons both committed criminal act and groundlessly suspected, the IUP procedure proved for them to be always subjectively significant and emotionally saturated. Exactly this forced researchers to adhere the determining role to emotions at forming “polygraph theories”.

Without any sign of denying existence of emotional component in the current state of a human exposed to polygraph interrogation Russian specialists approached to analysis of what is going on from somewhat different positions back a quarter of a century ago.

At explanation of the essence of polygraph interrogation the foreign specialists would provide more or less detailed description of the procedure and there could be given a number of examples to this effect. However it could have not been possible to get formalized definition of the phenomenon forming the ground of the method in any foreign study of polygraph issue. At the same time the clear definition of the phenomenon forming the fundament being the basis of polygraph interrogation technique and which, according to the American scientists “is probably a fundamental paradigm for psychophysiology”⁶¹, is obviously necessary from the theoretical, methodical and particularly practical point of view.

The analysis of concealed information technique test (that is guilty knowledge test and peak of tension test) approaches undertaken by the Russian specialists in the late 70s – early 80s allowed suggesting hypothesis on existence of a certain unified phenomenon being the basis of this and other polygraph testing techniques. The conditional

⁵⁹The threatening punishment theory remains most popular among the US poly-graphers (e.g., see: *Polygraphs and Security*. A study by a Subpanel of Scandia’s Se-nior Scientists and Engineers. October 21, 1999. 25 p.; *Gordon N. J., Fleisher W. L.* Effective interviewing and interrogation techniques. San Diego.: Academic press, 2002. 173 p.).

⁶⁰Great attention is paid to emotions studies in contemporary psychological science. Detailed examination of various emotions psychology aspects stays beyond the subject of this study. Only thoroughly studied and experimentally grounded statements and only to the extend necessary for research of polygraph examination analysis technique will be used for further presentation of the material and discussing the particular issues of emotions psychology, psychological processes, statuses and functions.

⁶¹*Orne M.T., Theckrey R.I., Paskevitz D.A.* On detection of deception. A model for the study of physiological effects of psychological stimuli // *Handbook of psychophysiology*. N-Y.: Holt Inc., 1972. P.744.

working name “psychophysiological phenomenon”⁶² was given to this phenomenon for its convenient application in future.

Psychophysiological Phenomenon – if described from the point of view of IUP technology – consists in the fact that external stimuli (a word, a thing, a photo, and so forth) bearing for a human significant information, in particular situation, on the event infixed in their memory, stably generates the physiologic reaction exceeding reaction towards associated (homogeneous) stimuli presented under the same conditions but not related to the aforementioned event and bearing no significant information⁶³ for the human.

The suggested definition of Psychophysiological Phenomenon proved to be practically useful and productive from the methodological point of view. However appearance of this definition gave rise to at least two questions.

First – could this definition be considered as universal one?

Second – does this definition grasp in full the phenomenon it is meant to describe?

In order to reply the aforementioned questions the analysis of studies in the field of experimental and applied psychophysiology was undertaken which allowed establishing that Psychophysiological Phenomenon in the way it was formulated was not a “private property” of IUP and was observed not only in

course of testing a human with polygraph, but under some other methodical conditions as well.

First, Psychophysiological Phenomenon is stably observed in course of operator profession persons activity (flight dispatchers, radar station operators and so forth) at fulfilling the tasks on detection, identification and classification of targets, objects, etc. Psychophysiological Phenomenon, realized in such conditions, is rather thoroughly studied in engineering psychology and the researches outcomes have been provided in the studies of the scientists of the Institute of Psychology of Russian Academy of Science⁶⁴. It may be noted in general that Psychophysiological Phenomenon functioning in the course of human testing using polygraph does not differ much methodically from its (Psychophysiological Phenomenon) manifestations in the conditions of operator work.

Second, Psychophysiological Phenomenon may be observed in the conditions of subliminal perception⁶⁵ when human’s psyche is being tested by subjectively significant for them but unconscious stimuli.

The researches have shown that “subliminal effect of emotional word consists in the fact that change of vegetative functions ... has been registered before the person tested could read the same”⁶⁶ The essence of this phenomenon consisted in the fact that “weak sensor stimuli may generate activation of

⁶²Mitrichev V.S. and Kholodny Yu.I. Polygraph as a tool... .

⁶³The concept of «psychophysiological phenomenon» was suggested by Aza-rov Yu.K. in the 80s last century and at first the Psychophysiological Phenomenon definition mentioned “the events, infixed in human *consciousness*”. However theoretical analysis of Psychophysiological Phenomenon functioning in different methodical conditions, brief presentation of which will be provided below, has given the ground for the author of this article to introduce clarification into the definition of Psychophysiological Phenomenon: it shall be referred to “events infixed in human *memory*”.

⁶⁴Methods of Psychic Statuses Diagnostics and of Human Activities Analysis / Under general editorship of Dikaya L.G. M: IP RAN (Institute of Psychology, RAS), 1994. 206 p.; etc

⁶⁵Coconscious perception is the “form of direct psychic reflection of reality, conditioned by the stimulants on whose influence a subject may not report to themselves” (Psychology. Dictionary. M.: Publishing House for Political Literature, 1990 P. 388-389), or whose influence they do not realize.

⁶⁶Kostandov E.A. Perception and Emotions. M. Meditsina (Medicine). 1977 P. 117.

cortical neurons (which render control signal for occurrence of changes in vegetative functions – *Yu.Kh.*), but space-temporal parameters of this excitement are not enough for a stimuli to be created⁶⁷. The examples of multi-channel registration of Psychological Phenomenon in such conditions are widely presented in the studies performed in the 70s-80s of the XX century under the guidance of Kostandov E.A.⁶⁸.

In late 70s the phenomenon of coconscious perception was studied by Russian specialists from the “lie detection” psychophysiological method position. Particularly guilty knowledge test was realized at subliminal level using tachistoscope⁶⁹ in the experimental study (in which the author of this article participated). The experiments demonstrated that the tested person selected one of 5-6 double-digit figures in the number sequence suggested and concealing this figure from the examiner as per the instruction received did not manage to view any of them on the screen. This was confirmed with interrogation of the tested person after the experiment was over. At the same time the physiologic reactions registered using polygraph allowed to observe rather stably in the course of the experiment the figure the tested person was concealing within the sequence examined.

However of greatest interest with respect to research of the mechanisms being the ground for detection of the concealed information from a human in the course of its testing using polygraph is the form of Psychophysiological Phenomenon realization in conditions of sleeping⁷⁰. Over forty years ago American researchers stated that “it

clearly demonstrated that sleeping subjects can make complex discrimination between repetitive auditory stimuli. They can for instance discriminate between meaningful words” and “if the stimuli are personally significant, corticofugal signals to the brainstem may evoke, in turn, arousal signs which may be electro-encephalographic, autonomic, or behavioral⁷¹”.

In the late 70s the phenomenon of coconscious perception under conditions of sleeping was modeled by the specialists engaged in studies of polygraph interrogation psychophysiological mechanisms. In course of the experimental study (in which the author of this article participated) the tested person being in the second stage of slow wave sleep was presented with the stimuli being of higher significance for this subject in wakeful state along with the neutral stimuli (as per polygraph interrogation technique). Thereat all demonstrations of “excitement” described by the American researches were observed, including Psychophysiological Phenomenon which was being registered rather stably. It was quite impressive to observe the way the tested person was sleeping and was not realizing the words or word combinations played from a tape-recorder (and it was confirmed with interrogation of the tested person upon their waking up), when the changes in dynamics of breathing and in cardiovascular system signaled subjective significance for the sleeping subject of the certain semantic stimulus perceived by their brain. In particular any sleeping human at this stage of sleep manifests reactions to their own name or to the name of any event, fact that are quite significant in the current timespan of their life.

⁶⁷Kostandov E.A. Perceived and Non-perceived Forms of Human Higher Nervous Activity // Human Brain Activity Mechanisms. Part 1. Human Neurophysiology / Under editorship of Bekhterev N.P. L.: Nauka (Science). 1988. P. 494.

⁶⁸Kostandov E.A. On Physiological Mechanisms of “Psychological Defense” and Unconscious Emotions // Unconscious. Tbilisi: Metsniereba. 1978. V. 1. P. 633-651

⁶⁹Intensity and length of double-digit figures demonstration by tachistoscope on the screen were gradually diminished and selected in the way the tested person was unable to identify the value of the figures although they have seen some flare on the screen.

⁷⁰Dixon N. Subliminal perception. London.: McGraw-Hill. 1971. 363 p.

⁷¹Oswald I., Taylor F., Treisman M. Discriminative responses to stimulation during human sleep // Brain. 1960. V. 83. P. 450.

The above listed four variants of Psychophysiological Phenomenon realization in different methodological conditions (IUP, operator activity, subliminal perception and subliminal perception in the state of sleep) demonstrated that the definition of this phenomenon suggested above was true and allowed stating at least three important conclusions.

First, indeed the universal neurophysiologic mechanism forms the ground of Psychophysiological Phenomenon and stably functions regardless substantially different methodical conditions and modality of the stimuli perceived. In the light of this the position of P. Davis⁷² was recognized as prospectless (which was supported in particular by the specialists of the Committee for studying scientific relevance of polygraph in the early XXI) according to which various theories were applicable in various situations. Psychophysiological Phenomenon is an objective reality, “fundamental mechanism of psychophysiology” of humans. And Psychophysiological Phenomenon therefore shall have the single theoretical justification and explanation of its mechanisms may not be dependent upon one or another technique applied during IUP.

Second, Psychophysiological Phenomenon realization at unconscious perception (both in the state of sleep and wakeful state) proves that its neurophysiologic mechanism does not depend upon human consciousness, it operates autonomously, aside from will and desire of a human being. The second conclusion lead to the important inference: if Psychophysiological Phenomenon gives rise to a stronger physiologic reaction towards one of the stimuli under the conditions of unconscious perception then some specific feature is inherent to this stimulus. **Significance of the stimuli** is such feature which in the 70s-80s of the last century did not attract due attention: it determines relevance of the information contained in the stimuli to the sense of the task decided by a human in particular situation.

External stimuli are ranged by the level of their subjective significance for each human. It happens supraliminally in wakeful state. A human reacts unconsciously towards the stimuli significant for them at subliminal perception or during sleep. Such reaction proves that when there happens no realization of stimuli, psyche keeps classifying the stimuli perceived externally according to their subjective significance (for the current moment in the human life).

Finally, **third**, Psychophysiological Phenomenon realization under the conditions of unconscious perception in the wakeful or sleep state (that is when the perceived stimuli are not recognized) lead to a thought that the decisive role in the mechanism of its realization may be played not by human emotions (which can not arise under the conditions mentioned above in principle) but by their **memory**.

Indeed, in the circumstances mentioned some stimuli bear information on some event which is of subjective significance for a human: such stimuli would certainly be perceived and evaluated by human psyche aside from their desire and will. This process will be accompanied by expressive physiological reactions of the body which will be observable during IUP. If the stimulus perceived in the same conditions is subjectively insignificant for a human the reaction to such stimuli will not have stable and expressive character. Concealed information technique is based, in particular, on the principle specified.

The important advantage of concealed information technique is the fact that it serves as “incorporate an extremely effective safeguard against false positive errors – the innocent person cannot determine which question is the critical question (that is a relevant question - *Yu.Kh.*), and therefore cannot consistently react to it regardless of how nervous or fearful he is”⁷³.

⁷²Davis R.C. Physiological responses...

⁷³The accuracy and utility of polygraph testing (Department of Defense, Washington, D.C.) // Polygraph. 1984. № 1. P. 59.

Basing upon the own experimental researches described above, the data on identified by that period brain structures, involved in genesis of emotional statuses and memory⁷⁴ and results of the studies in the field of neurophysiology⁷⁵, obtained by the mid 80s, the author of the present article attempted for the first time in the Russian science to explore neurophysiologic mechanisms of Psychophysiological Phenomenon realization.

The researches conducted those years demonstrated that “if in the previous experience of the system the stimuli concerned (or same one) coincided with a certain biologically important activity, there happened activation of memory traces with transfer of excitement to the subcortical centers of emotions and motivations correspondent to the activity concerned. All these stages of stimuli information processing seem to be obvious... The association cortex, secondary and tertiary areas of this analyzer and hippocampus structures take part in it”⁷⁶.

When studying possible neurophysiologic mechanisms of Psychophysiological Phenomenon the attention was paid to the fact that amygdale is closely tied with hippocampus and these two structures participate together in organization of various forms of emotional behavior.

The researches allowed to get to the hypothesis that amygdale plays specific role in Psychophysiological Phenomenon realization under the conditions of testing human during IUP. Scientific data demonstrated in particular that “amygdale plays an important if not a major role in evaluating such feature of signals

as emotiveness. Marking the coming signals on the basis of the past experience the amygdale forms hierarchic relations in the signals flow... Participation of amygdaloid complex in the memory processes may consist in regulation of the flow of the signals forwarded for fixation and preservation... The comprehensive signal formed with its participation proves to be significantly more resistant to interfering influence of the similar ones because it possesses additional parameter – significance realized in neuron signals”⁷⁷. Thus, proceeding from the outcomes of the theoretical and experimental researches conducted, the author of this article suggested in 1987 the **theoretical concept of purposeful memory examination**, which gave explanation to the mechanisms forming the ground of revealing information from human in the course of IUP.

IV.

Memory plays fundamental role in the mechanisms of “lie detection” psychophysiological method, and many facts point to this. Starting from the first described experiment of “lie detection” with use of laboratory devices, performed by C. Lombroso in the late XIX⁷⁸ century, polygraph test practice pushed the researchers to the analysis of the role the memory plays in this psychophysiological method technique. However instead of this, the researchers would focus on emotional tension state of a human when he was in the situation of polygraph testing. Such an approach would bring up emotions of the tested person as the most important psychophysiological component, pushing their memory aside.

⁷⁴Simonov P.V. Human Higher Nervous Activity... ..; Simonov P.V., Ershov P.M. Temperament. Character. Personality. M.: Nauka (Science), 1984. 160 p.

⁷⁵Vinogradova O.S. Hippocampus and Memory. M.: Nauka (Science), 1975. 333 p.; Natural Scientific Fundamentals of Psychology. M.: Pedagogika (Pedagogy), 1978. 368 p.; Chepurinov S.A., Chepurinova N.E. Brain Amygdaloid Complex. M.: Moscow State University (MSU) Publishing House, 1981. 255 p.; etc.

⁷⁶Ivanitskiy A.M., Strelets V.B., Korsakov I.A. Informational Processes of Brain and Psychic Activity. M.: Nauka (Science) 1984. P. 101.

⁷⁷Amygdaloid Complex (Correlations, Behavior, Memory). Novosibirsk: Nauka (Science), 1981, P. 170-172 p.

⁷⁸When testing some Tosetti suspected in a murder of a six-year old girl C. Lombroso “adopted the plethysmograph and found a slight diminution of the pulse when Tosetti was set to do a sum; when, however, skulls and portraits of children covered with wounds were placed before him, the line registered showed no sudden variation, not even at the sight of the little victim’s photograph. The results of the foregoing examination proved conclusively that Tosetti was innocent of a crime” (quoted from *Trovillo P. V. A history of lie detection // J. of Criminal Law and Criminology. 1939. V. 29. No. 6. P. 863*).

The results of the researches by Voronin L.G. and Konovalov V.F. serve as strong argument to the fact that memory is a leading psychological function at revealing from a human using psychophysiological method the information which may be concealed; they applied elements of polygraph testing technique in the experiments when studying the mechanisms of memory functioning in the early 70s of last century.

When explaining selection of actually prohibited in that time “lie detection” method as a research tool the scientists pointed that “any technique is perspective if using the same one can manage to discover changes of vegetative and other reactions arising at emerging, preservation and interaction of affective afterimage”⁷⁹ in human memory. Voronin L.G. and Konovalov V.F. came to the conclusion that “traces of stimulation discoverable through electrosensitive reactions are neurophysiologic basis of memory... For long-term memory ... (these reactions – *Yu.Kh.*) accompany the process of retrieval of information from memory”⁸⁰. When applying ECG, EEG, galvanic skin reaction and other electrophysiological research methods for memory mechanisms examination, the scientists discovered that “if strong excitement of signal systems emerges, irradiating emotional sphere of brain activity, the same will find its reflection in electrographic components (for example, in GSR). This is especially expressively manifested if emotions along with signal systems create the specific state which is usually called concern”⁸¹.

Criminalistics science has been leading to understanding that memory plays an important role in polygraph interrogation mechanisms. It is known that criminalistics faces two classes of traces during investigations of crime - materially fixed traces

and afterimages (that is “ideal traces”, “prints” of the crime event), infixed in human memory.

As mentioned above, Polish criminalists came to the same point of view after the Russian specialists and independently. The opinion that during IUP examination of events traces preserved in human memory is carried out became apparently common in Polish criminalistics in the mid 90s. The Polish researchers⁸², having truthfully determined memory as fundamental basis of polygraph interrogation unfortunately adjoined “ideal traces” kept in memory to the emotions having thus obtained modified threat-of-punishment theory as a result.

Memory is the form of mental reflection of reality consisting in infixing, preservation and further reproduction of human experience by the same. Connecting past with the present and future, memory is the most important cognitive function of human, forming the ground for their development and education. Contemporary psychophysiological science understands totality of memory systems by the function of “memory”: long-term and short-term memory, procedural and declarative memory (where the latter is divided into episodic and semantic one), etc.

Now, nineteen years after, it is possible to admit that the suggested hypothesis on the role of amygdale in neurophysiologic mechanisms of Psychophysiological Phenomenon proved to be correct. As the further researches demonstrated, participation of namely this structure “ensures formation of stably and durably preserved traces of emotional memory” rapid and firm infixing of emotional events in memory”⁸³, due to which “emotional memory trace is not erasable and subject to amnesia”⁸⁴.

⁷⁹Voronin L.G. and Konovalov V.F. *Electrographic Trace Processes and Memory*. M.: Nauka (Science), 1976. P. 88-89.

⁸⁰Voronin L.G. and Konovalov V.F. *Electrographic Trace Processes ...*, P. 145.

⁸¹Same as above, P. 102.

⁸²Wariographia kryminalistyczna. Wydawnictwo Wyższej Szkoły Policji, Szczytno, 1998.; *Krzyżecin A.* The Debate Over Polygraph in Poland // *Polygraph*. 2000. V. 29. № 3. P. 226-236.

⁸³*Danilova N.N.* *Psychophysiology* M.: Aspect Press. 1999. P.130.

⁸⁴Amygdaloid Complex ... P. 175

Polygraph application practice confirms accuracy of the scientific data obtained: in the course of real polygraph interrogations it proved to be possible to reveal from human the afterimages of higher significance (preserved in emotional memory) events backed in past by 15-20 years.

M. Kleiner, when developing the oriented “polygraph theory” mentioned above went farther than Polish polygraphologists: he was studying correlation between emotions and memory in course of polygraph interrogation and correlation between test and control questions in particular. When examining the place of memory in polygraph interrogation technique the Israeli polygraphologist applied analysis of neuron mechanisms of the function and also discovered important role of amygdale in formation of affected memory traces⁸⁵. Proceeding from the different scientific positions Kleiner M. finally came to assertions close to those stated by the native scientists in late 80s.

Thus many direct and indirect data obtained by the researchers during the last 20-25 years unambiguously testify – memory⁸⁶ is the fundamental psychic function subject to examination by “lie detection” psychophysiological method using polygraph.

The purposeful memory examination theory, developed in the Russian polygraphologist school, consists in the fact that during testing with polygraph the afterimages of events kept in memory of a human, may be intentionally actualized using the aim set and then detected through registered physiologic reactions as they arise in response to stimuli presented to them (human being) specially selected and grouped.

From the positions of task-oriented memory test theory many phenomena empirically observed in course of polygraph

interrogation may be successfully explained. For example with the help of the theory the following becomes clear:

- high effectiveness of polygraph testing in laboratory environment (which was difficult to explain from the positions of threat-of-punishment theory) and intensity of registered reactions in the cases when the tested person did not know about registration carried out (which is also not explainable from the point of view of the function of “threat-of-punishment avoidance”);
- reason for intensity of reactions registered in the course of silent test (which was not understandable from the point of view of conflict theory) or in laboratory environment (where the conditioned reflex theory “stumbled”);
- nature of “activating power” of stimuli presented during polygraph testing which under the influence of the aim set examine activated afterimages of human memory (the activation theory was unable to point what and where is activated in human mind and brain);
- necessity of prior discussion of the questions with the tested person before presenting them during polygraph testing which contradicted postulates of the orientation theory and the number of other facts.

Thus the theoretical concept of task-oriented memory test, according to our opinion forms solid basis for sufficiently clear explanation and acceptable natural scientific grounding of mechanisms ensuring accuracy of information obtained from a human in consequence to the polygraph test of the same.

⁸⁵Kleiner M. Physiological detection of deception ..., P. 147-149.

⁸⁶This assertion shall not be understood as denying the role of psychologically important components of current state of a human at IUP (emotions, attention, set, etc.). The issues of participation and influence of certain psychic processes will be examined in next articles.

References

- Abrams S. A* The Complete Polygraph Handbook. Lexington Books, 1989. 267 p.
- Amygdaloid Complex (Correlations, Behavior, Memory). Novosibirsk: Nauka (Science), 1981, P. 170-172 p.
- Belyushina O.V., Ladchenko A.G.* Polygraph in Business. M.: NOU ShO "Bayard", 2004. 118 p.
- Ben-Shakhar G.* A further study of the dichotomization theory in detection of information // Psychophysiology. 1977. V. 14. P. 408-413.
- Ben-Shakhar G., Lieblich I.* The dichotomization theory for differential autonomic responsivity reconsidered // Psychophysiology. 1982. V. 20. P. 277-281.
- Ben-Shakhar G., Furedy J.* Theories and applications in the detection of deception. A psychophysiological and international perspective. Springer-Verlag New-York Inc., 1990. 192 p.
- Barland G.* The polygraph test in the USA and elsewhere // The polygraph test. Lies, Truth and Science. London.: SAGE Publications, 1988. P. 73-95.
- Barland G.H., Raskin D.C.* Detection of deception // Electrodermal activity in psychological research. N-Y.: Academic Press, 1973. P. 419-477.
- Chepurnov S.A., Chepurnova N.E.* Brain Amygdaloid Complex. M.: Moscow State University (MSU) Publishing House, 1981. 256 p.
- Danilova N.N.* Psychophysiology. M.: Aspect Press. 1999. 373 p.
- Davis R.C.* Physiological responses as a means of evaluating information // The manipulation of human behavior / A.D. Biderman & H. Zimmer (Eds.). N-Y.: Wiley, 1961. P. 142-168.
- Dixon N.* Subliminal perception. London.: McGraw-Hill. 1971. 363 p.
- Feldman H.* Lie detection manual. (A workbook - textbook). Franklin Lakes: Allison Press, 1982. 201 p.
- Gordon N. J., Fleisher W. L.* Effective interviewing and interrogation techniques. San Diego.: Academic press, 2002. 173 p.
- Gustafson L., Orne M.* Effects of heightened motivation on the detection of deception // J. of Applied Psychology. 1963. V. 47. P.408-411.
- Gustafson L., Orne M.* Effects of perceived role and role success on the detection of deception. // J. of Applied Psychology. 1965. V. 49. P. 412-417.
- Hassett J.* Introduction to Psychophysiology. M.: Mir (World), 1981. 248 p.
- Heslegrave R.* An examination of the psychological mechanisms underlying deception // Psychophysiology. 1982. V. 19. P. 323.
- Heslegrave R., Furedy J., Matyas T.* Studies on the validity and utility of T-wave amplitude // Psychophysiology. 1982. V. 19. P.323.

- Horneman C., O'Gorman J.* Detectability in the card test as a function of the subject's verbal response // *Psychophysiology*. 1985. V. 22. № 3. P. 330-333.
- Horvath F., Reid J.* The polygraph silent answer test // *J. of Criminal Law, Criminology and Police Science*. 1972. V. 63. P. 285-293.
- Ivanitskiy A.M., Strelets V.B., Korsakov I.A.* Informational Processes of Brain and Psychic Activity. M.: Nauka (Science), 1984. 200 p.
- Janig W.* Vegetative Nervous System // *Human Physiology. Nervous System*. M.: Mir (World), 1985. V. 1. P. 167-219.
- Kleiner M.* Physiological detection of deception in psychological perspectives: a theoretical proposal // *Handbook of Polygraph Testing*. San-Diego: Academic Press, 2002. P. 127-182.
- Kholodniy Yu.I., Savelyev Yu.I.* The Problem of Application of Polygraph Tests: Invitation to Discussion // *Psikhologichesky Zhurnal (Psychological Magazine)*. 1996. V. 17. No. 3. P. 53-69.
- Kholodniy Yu.I.* Application of Polygraph for Prevention, detection and Investigation of Crimes. M.: Mir Bezopasnosty (World of Security), 2000. 160 p.
- Kholodniy Yu.I.* Interrogation Using Polygraph and Psychic Reflection // *Bulletin of Volzhskiy University. Series "Jurisprudence"*. Toliatti, V. 18. P. 205-209.
- Kholodniy Yu.I.* Criminalistic Polygraphology and its Application in Law Enforcement Practice // *Informational Bulletin № 21*. M.: Academy of the MIA, 2003. P. 14-19.
- Kostandov E.A.* Perception and Emotions. M.: Meditsina (Medicine) 1977. 247 p.
- Kostandov E.A.* Realized and Non-realized Forms of Human Higher Nervous Activity // *Human Brain Activity Mechanisms. Part 1. Human Neurophysiology / Under editorship of Bekhtereva N.P.* L.: Nauka (Science). 1988. P. 491-526.
- Kostandov E.A.* On Physiologic Mechanisms of "Psychological Defense" and Non-reported emotions // *Unconscious*. Tbilisi: Metsniereba. 1978. V. 1. P. 633-651.
- Krzywocin A.* The Debate Over Polygraph in Poland // *Polygraph*. 2000. V. 29. № 3. P. 226-236.
- Luria A.R.* The nature of human conflicts. N-Y.: Liverwright, 1932. 164 p.
- Lykken D.* Psychology and the lie detection industry // *American Psychologist*. 1974. V. 29. P. 728.
- Matte J. A.* Forensic psychophysiology using the polygraph // N-Y.: J. A. M. Publications, 1996. 793 p.
- Methods of Psychic Statuses Diagnostics and of Human Activities Analysis / Under general editorship of Dikaya L.G.* M: IP RAN (Institute of Psychology, RAS), 1994. 206 p.;
- Mitrichev V.S., Kholodniy Yu.I.* Polygraph as the Tool to Obtain Orienting Criminalistic Information // *Zapiski Kriminalistov (Notes of Criminalists)*. 1993. V. 1. P. 173-180.
- Nagle D.E.* The polygraph in employment: applications and legal considerations // *Polygraph*. 1985. V. 14. № 1. P. 1-33.

- Natural Scientific Basics of Psychology. M.: Pedagogika (Pedagogics), 1978. 368 p.
- Orne M.T., Theckrey R.I., Paskevitz D.A.* On detection of deception. A model for the study of physiological effects of psychological stimuli // Handbook of psychophysiology.- N-Y.: Holt Inc., 1972.- P.743-785.
- Oswald I., Taylor F., Treisman M.* Discriminative responses to stimulation during human sleep // Brain. 1960. V. 83. P. 440-453.
- Reid J., Inbau F.* Truth and deception. The polygraph ("lie-detector") technique. Baltimore: The Williams & Wilkins Co., 1977 (2-nd edit.). 430 p.
- Petrov A.M., Myagkhih S.G.* From the Polygraphologist Notebook. Perm: 2003. 202 p.
- Podshibyakin A.S., Kholodniy Yu.I.* On Clarification and Addition of the Criminalistic Diagnostics Objects // Russian Juristic Doctrine in XXI century: Problems and Approaches for Decision: Scientific Practical Conference. (October 3-4, 2001). Saratov 2001. P. 225-227.
- Polygraphs and Security. A study by a Subpanel of Scandia's Senior Scientists and Engineers. October 21, 1999. 25 p.
- Psychology. Vocabulary. M.: Publishing House for Political Literature, 1990. 494 p.
- Scientific validity of polygraph testing: a research review and evaluation - a technical memorandum. Washington, DC: U.S. Congress, Office of Technology Assessment, 1983. 132 p.
- Simonov P.V.* On the Role of Emotions in Adaptive Behavior of Living Systems // Voprosy psikhologii (Problems of Psychology) 1965, No. 4, P. 75-84.
- Simonov P.V.* Higher Nervous Activity. Motivational and Emotional Aspects. M.: Nauka (Science), 1975. 173 p.
- Simonov P.V.* Emotional Brain (Physiology, Neuroanatomy, Psychology of Emotions). M.: Nauka (Science), 1981. 221 p.
- Simonov P.V., Ershov P.M.* Temperament. Character. Personality. M.: Nauka (Science), 1984. 160 p.
- Sokolov E. N.* Perception and the Conditioned Reflex. N-Y.: McMillan, 1963.
- Sokolov E. N.* Orienting reflex as information regulator // Cognition and Categorization / Rosch E. and Lloyds B. (eds). Hillsdale, NJ: Lawrence Erlbaum Associates, 1966. P. 79-98.
- Thackray R., Orne M.* A comparison of physiological indices in detection of deception // Psychophysiology. 1968. V. 4. P. 329-339.
- The accuracy and utility of polygraph testing (Department of Defense, Washington, D.C.) // Polygraph. 1984. № 1. P. 1-143.
- The Polygraph and Lie Detection. Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2003. 398 p.
- Trovillo P. V.* A history of lie detection // J. of Criminal Law and Criminology. 1939. V. 29. №. 6. P. 848-881.

-
- Varlamov V.A., Varlamov G.V.* Psychophysiology of Polygraph Tests. Krasnodar: 2000. 239 p.
- Vinogradova O.S.* Hippocampus and Memory. M.: Nauka (Science), 1975. 333 p.
- Voronin L.G., Konovalov V.F.* Electrosensitive Trace Processes and Memory. M.: Nauka (Science), 1976. 166 p.
- Wariographia kryminalistyczna. Wydawnictwo Wyższej Szkoły Policji, Szczytno, 1998. 141 p. (Poland).
- Zhurin S.I.* Practice and Theory of Lie Detectors Application. M.: Goryachaya Liniya - Telekom (Hot Line - Telecom), 2004. 143 p.
- Zubrilova I.S., Skrypnikov A.I.* Application of Polygraph for Detecting Crimes by the Internal Affairs Authorities. Course of Lectures. M.: VNII MVD (All-Russian Research Institute of the Ministry of the Interior), 1999. 117 p.

A Case Study of a Polygraph Examination on an Examinee with an Implanted Defibrillator

Nathan J. Gordon and William L. Fleisher¹

Abstract

There is little, if any, research published concerning the examination of subjects with implanted defibrillators and the subsequent affect this device would have on the outcome of a Psychophysiological Detection of Deception (PDD) examination. The evidence of an error or inconclusive charts produced by a polygraph subject with an implanted pacemaker defibrillator is anecdotal at best. An implanted pacemaker defibrillator is a battery operated electronic device usually implanted subdermally to correct arrhythmias of the heart and prevents myocardial infraction. Unless disclosed in the pre-test interview, it is highly unlikely that an examiner would be aware that an examinee has such a device. This article provides an account of a polygraph examination of an examinee with an implanted pacemaker defibrillator.

Background

A PCSOT examinee scheduled for a monitor examination disclosed during the pre-test interview that he had a pacemaker defibrillator implanted in his chest to control for a rapid heartbeat. He produced a medical card that identified the device as a model 7230, and included the company name and phone number.

The defibrillator is a small metal device that containing electronic circuitry and a battery, similar to a pacemaker. Pacemakers are implanted to increase a slow heartbeat. Defibrillators, on the other hand, detect and correct both fast and slow heart rates. The first defibrillation system was implanted in 1980. Today, over 35,000 defibrillators are implanted each year (Medtronic, Inc, 2005). The model implanted in the examinee was the Marquis DR Model 7274, from Medatronic. A technical support person from the company advised that the polygraph examination would not have any adverse affect on the defibrillator. However, representative advised that the defibrillator may have an adverse affect on the cardio component of the polygraph. In addition to the implant the examinee reported using Verapamil, a

medication for his heart rate (Arky, 1995), as well as insulin for his diabetes (Arky, 1995).

Instrumentation

A Lafayette LX-4000 computerized polygraph system, software version 9.5.1, was utilized for the examination.

Procedure

A demonstration test was administered, followed by four charts of an Integrated Zone Comparison Technique multi-issue examination (Gordon & Fleisher, 2000). The comparison and relevant questions were:

C5 While in treatment have you deliberately misled your group about anything?

R6 Since April, 2004, have you had sexual contact with anyone under the age of 18?

C8 While on parole, other than what we discussed, have you lied to your Parole Officer about anything?

R9 Since April, 2004, have you done anything sexually you could be arrested for?

¹Academy for Scientific Investigative Training

C11 While in treatment, other than what we discussed, have you lied to anyone about anything?

R12 Since April, 2004, have you sex with anyone you deliberately did not tell me about?

Results

While there did not appear to be any noticeable changes in the examinee's pulse rate, very observable blood volume changes did occur, as can be seen in R6 and R9 in figure one.

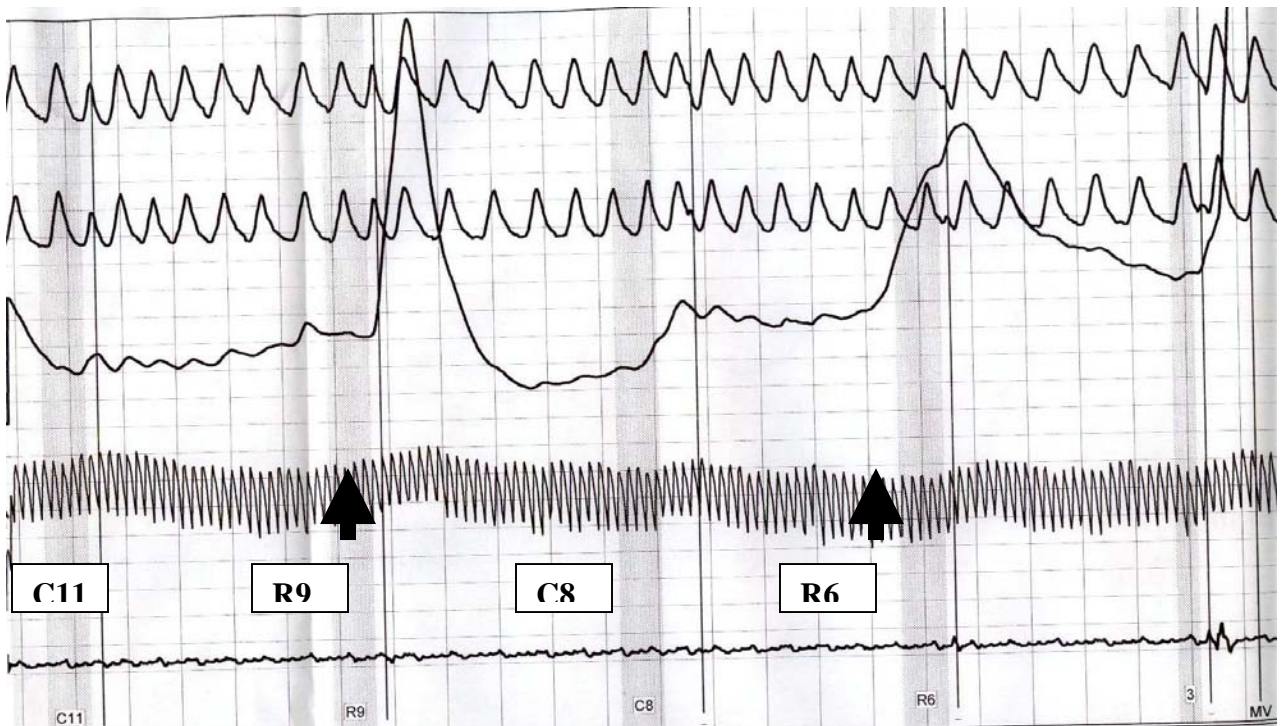


FIGURE 1

The charts were scored using three algorithms: Polyscore® 5.5, Objective Scoring System (OSS) (Krapohl & McManus, 1999) and ASIT PolySuite (Gordon & Cochetti, 1987). It should be noted that none of these algorithms are validated for multi issue examinations (spot question decisions). Polyscore® 5.5 indicated: INCONCLUSIVE - PROBABILITY OF DECEPTION IS LESS THAN 0.24. Polyscore® did showed signal weights of its decision based on BLOOD VOLUME as a -0.37 and PULSE +0.01 (RESPIRATION +0.33, ELECTRODERMAL +0.29). The strongest index indicative of deception was blood volume.

Objective Scoring System indicated DECEPTION INDICATED. Based on the published field research, the likelihood that these polygraph data were produced by a

truthful subject is 3% or lower (Krapohl & McManus, 1999).

ASIT PolySuite, which combines the "Horizontal Scoring System," with the Academy for Scientific Investigative Training's Algorithm for Chart Interpretation (Gordon, 1999) indicated: DECEPTION.

ASIT PolySuite showed component contributions resulting in its deceptive scores as BLOOD VOLUME 40% (PNEUMO 36% and ELECTRODERMAL 24%). The strongest index indicative of deception was blood volume.

When the examinee was confronted with the results of the examination he admitted that he had sex with his daughter's mother, on numerous occasions. These acts took place at her house, which was a direct

violation of the special conditions of his parole, specifically, 1) the parolee is not allowed outside the City of Philadelphia without permission and 2) being in a house where there was a minor child (his daughter).

Conclusion

This case study examined the feasibility of whether a successful Psychophysiological Detection of Deception (PDD) could be conducted on a subject with an implanted defibrillator, and whether the device would have an adverse affect on the data collected in the cardio parameter.

Based on the information provided by the medical technician the examination can be conducted without any concern of the polygraph instrumentation interfering or affecting the device. Based on the data collected and analysis of component contributions toward the examination scores, it does not appear the device had any adverse affect on blood volume changes which is the primary indicator of reaction in the cardiograph component. To provide a solid empirical foundation for these conclusions, further research needs to be conducted in this area.

References

- Arky, R. (1995). *Physicians' Desk Reference*. Medical Economics Data Production Company, Montvale, NJ
- Gordon, N.J. (1999). The Academy for Scientific Investigative Training's Horizontal Scoring System and Examiner's Algorithm for Chart Interpretation. *Polygraph*, 28(1), 56-64.
- Gordon, N. J., & Cochetti, P. M. (1987). The Horizontal Scoring System. *Polygraph*, 16(2), 116-125.
- Gordon, N.J., Fleisher, W. L., et al. (2000). A Field Validity Study of the Integrated Zone Comparison Technique. *Polygraph*, 29(3), 220-225.
- Krapohl D., & McManus, B. (1999). An Objective Method for Manually Scoring Polygraph Data. *Polygraph*, 28(1), 209-222.
- Matte, J. A. (1996). Using the Polygraph. In *Forensic Psychophysiology*. J.A.M. Publications: Williamsville, NY.
- Medtronic Inc. (2005). *History of Defibrillators: Technology Development through the Decades*. Web site: www.medtronic.com.

Effects of Audiovisual Presentations of Test Questions During Relevant-Irrelevant Polygraph Examinations and New Measures¹

John C. Kircher², Dan J. Woltz³, Brian G. Bell⁴, and Paul C. Bernhardt⁵

Abstract

The present study tested the prediction that audiovisual presentations of test questions would increase the accuracy of relevant-irrelevant (RI) tests. Ninety-six male and female university students and staff were tested about information they had provided about themselves on a simulated employment application. All subjects falsified answers to two items on the employment form. Subjects were then given two RI tests. On each test, the subject answered deceptively to one of four questions about the employment form. Subjects were offered \$50 if they could convince the polygraph examiner that all of their answers on both tests were truthful. In the control condition, the polygraph examiner presented test questions only aurally. In the experimental condition, questions were presented visually, in large characters on a computer monitor, as well as aurally by the polygraph examiner. Polygraph outcomes were determined by computer analysis of the physiological recordings.

Audiovisual presentations of test questions had no discernable effects on the accuracy of test outcomes or individual physiological measures. A computer model of electrodermal, cardiovascular, and respiration measures based on prior polygraph research correctly identified 65% of the false statements by subjects. In 19% of the cases, the computer indicated that the subject was deceptive but identified the wrong test question (false positive). In the remaining 16% of cases, the computer indicated that the subject was truthful (false negative). A new computer model was developed that yielded 80% correct decisions, 8% false positive, and 12% false negative outcomes.

Electrodermal, respiration, absolute blood pressure, peripheral vasomotor activity, heart rate, and vagal tone discriminated between truthful and deceptive responses during the polygraph examinations. Continuous measures of absolute blood pressure were more diagnostic than measures derived from the traditional cardiograph. However, subjects' level of involvement in the task may have been too weak to observe large effects on the cardiograph measures.

Subjects reported using various strategies to defeat the tests. Although different strategies were used in audiovisual and audio only conditions, none of the countermeasures correlated with false negative decision outcomes. Additional research is needed to determine if subjects' knowledge of the rationale that underlies RI tests, or if instructions and training in the use of specific countermeasures reduce the accuracy of RI tests.

¹ This article originally took the form of a final report to the United States Government on October 1, 1998. The views expressed in this article do not reflect those of the United States Government.

²Ph.D. (Principal Investigator), Department of Educational Psychology, University of Utah, Salt Lake City, UT 84112, voice: (801) 581-7130, fax: (801) 581-5566, email: kircher@gse.utah.edu

³Ph.D. (Co-principal Investigator), Department of Educational Psychology, University of Utah, Salt Lake City, UT 84112, voice: (801) 581-2047, fax: (801) 581-5566, email: woltz@gse.utah.edu

⁴Ph.D. (Research Associate), Department of Educational Psychology, University of Utah, Salt Lake City, UT 84112

⁵M.S. (Graduate Assistant), Department of Educational Psychology, University of Utah, Salt Lake City, UT 84112

Introduction

Relevant-irrelevant (RI) polygraph tests are commonly administered to prospective government employees and are periodically given to employees who have access to classified information (OTA, 1983). In security screening examinations, the costs associated with false negative polygraph outcomes often outweigh those of false positive errors, and some research suggests that false negative errors are more common. Barland, Honts, and Barger (1989) reported that the risk of false negative errors in government screening examinations may be as high as 66%. This figure is many times greater than the 3% to 8% false negative error rates typically obtained for specific-incident comparison question tests (e.g., Raskin, Kircher, Honts, & Horowitz, 1988).

One possible reason for false negative errors in screening examinations is that government employees with access to classified documents are typically brighter and better informed about polygraph techniques than the average citizen. They may be better able to use countermeasures or may have access to specialized training that would increase the effectiveness of their attempts to defeat the test (Honts, Raskin, & Kircher, 1994). Although subjects' awareness of the rationale underlying polygraph tests and subjects' spontaneous use of dissociation as a countermeasure do not appear to increase the risk of false negative errors in specific-incident comparison question tests (Honts, Raskin, & Kircher, & Hodes, 1988; Rover, Raskin, & Kircher, 1979), these factors may account for some of the false negative errors in screening tests. Whether or not subjects report it, we suspect that most individuals will attempt some form of mental dissociation to minimize the magnitude of their physiological reactions to test questions.

The present study will attempt to identify the types of dissociative techniques used spontaneously by subjects during RI polygraph tests by interviewing subjects immediately after they have taken two such tests. Data from these interviews will be used to explore the possibility that some dissociative techniques may be more effective

at masking deceptive responses than others.

Whether or not dissociation is a serious problem in RI tests, one possible approach to countering dissociation and improving the accuracy of RI tests is to increase the demands on subject's attentional resources. Test questions presented simultaneously to two sensory modalities rather than one may make it more difficult for subjects to divert their attention from the issues under investigation.

The primary objective of the present study was to evaluate the effects of combining audio and visual presentations of test questions on subjects' physiological responses during polygraph tests. It was expected that the combination of visual and audio presentations of test questions would increase the differences between the physiological responses associated with truthful and deceptive answers and improve decision accuracy. Underlying this prediction was the possibility that audiovisual presentations would command more attention and interfere more with subjects' spontaneous attempts to dissociate during the test than would audio presentations alone. To test this prediction, we tested only deceptive subjects who had falsified information on a simulated employment application.

Two prior investigations revealed no significant differences in the decision accuracies produced by visual and aural presentations of test stimuli (Biejk, 1980; Carlson & Smith, 1991). However, a detection of information paradigm was used in both of those studies; Biejk used numbers tests, and Carlson and Smith used concealed information tests. These test protocols have limited practical utility, and they produce results that may not generalize to the types of tests commonly used in the field. Neither study examined the combined effects of auditory and visual presentations of test stimuli on physiological responses. Neither study offered subjects strong incentives to appear truthful on the test; and neither study evaluated physiological measures other than skin conductance or resistance.

The present study examined all of the physiological measures traditionally used by

field examiners to decide if the subject was truthful or deceptive on the test. These included thoracic and abdominal respiration, skin conductance, cardiograph, and finger pulse amplitude. The proposed research also examined several new measures as well. Recent research by John A. Podlesny (personal communication, 1996) suggests the Finapres blood pressure monitor might be an effective replacement for the traditional cardiograph. The cardiograph currently used by field polygraph examiners is obtained from a blood pressure cuff that is wrapped around the upper arm and inflated to about 60 mm Hg. Because the cuff occludes blood flow in the arm, it is uncomfortable, and the level of discomfort increases the longer it remains in place. Consequently, the traditional cardiograph limits the time available to collect data from subjects and may ultimately become a distraction that competes for the subject's attention to test questions.

In contrast to the cardiograph, the Finapres device outputs continuous measures of blood pressure in absolute units (mm Hg). The cuff for the Finapres fits on a finger, is relatively unobtrusive, and can be inflated for long periods with little or no discomfort to the subject. Podlesny's research suggests that the Finapres may be useful for comparison question tests, but its utility in RI tests has never been evaluated.

In the present study, concurrent measures of cardiovascular activity were obtained with the cardiograph and with the Finapres. Multiple features were extracted from each of these waveforms to determine if either channel provided data that are diagnostic of truth and deception in RI tests.

Vagal tone is another measure that shows promise (Porges, Bohrer, Cheung, Drasgow, McCabe, & Keren, 1980). Vagal tone is a measure of the extent to which heart period covaries with frequencies associated with normal respiration (sinus arrhythmia). Although many physiological factors affect heart rate, respiration probably accounts for the greatest proportion of variance in heart rate over the short span of a polygraph examination.

Prior research with the comparison

question test revealed that vagal tone 15 to 20 seconds after question onset discriminated between truthful and deceptive subjects (Raskin & Kircher, 1990). As predicted, vagal tone was lower when guilty subjects answered relevant question than when they answered control questions, and innocent subjects showed the reverse pattern.

Our preliminary findings with vagal tone were encouraging. Vagal tone is mediated by the vagus nerve, which is part of the parasympathetic nervous system. Conversely, most of the physiological changes currently scored by field polygraph examiners and by our computerized scoring system are mediated by the sympathetic nervous system. This raises the possibility that vagal tone may be relatively independent of the measures we currently use to discriminate between truthful and deceptive subjects. To the extent that physiological measures are both diagnostic and independent of one another, they would provide complimentary sources of information about the subject's truthfulness and would improve the accuracy of polygraph outcomes.

Methods

Subjects

After pilot testing, 48 male and 48 female subjects were recruited from the general university community by means of advertisements in the university newspaper and postings on bulletin boards around campus. The advertisements offered \$10 per hour and a possible \$50 bonus to qualified students and non-faculty staff for participation in a study about lie detection.

Prospective subjects called a secretary and were interviewed over the telephone. Callers qualified for the study if (1) they were at least 18 years of age, (2) had at least a high school education, (3) took no prescription medications except birth control pills, (4) never had a polygraph test, and (5) were fluent in English.

Demographic data were obtained from 90 of the 96 participants. The mean age of the sample was 24.5 years (SD=6.8). Years of education ranged from 12 to 22 (M=15.5, SD=1.7). Most participants were Caucasian

(90%) or Asian (7%) and unmarried (72%). The participants were currently or had been university students who majored in social or behavioral sciences (24%), biological sciences (21%), arts and humanities (13%), physical sciences and math (10%), or other (31%). The obtained distribution of study participants was similar in age, sex, education, and ethnicity to the target population of job applicants and employees of the government sponsor.

Apparatus

The CPS-LAB system (Scientific Assessment Technologies, SLC, UT) was used to configure the data collection hardware, specify storage rates for the physiological signals, and build automated data collection protocols. CPS-LAB was also used to collect, edit, and score the physiological data.

The physiological data acquisition subsystem (PDAS) of CPS-LAB generated analog signals for thoracic and abdominal respiration, skin conductance, cardiograph, finger pulse amplitude, and cardiometer. Calibrated analog output from a Ohmeda 2300 Blood Pressure Monitor was routed to a general-purpose coupler on the PDAS. Each of the seven analog signals was digitized at 1000 Hz with a Metrabyte DAS 16F analog-to-digital converter installed in a 50 MHz PC compatible 486 computer with 16 MB of RAM.

Respiration was recorded from two Life-Tech respiration transducers secured with Velcro straps around the upper chest and the abdomen just below the rib cage. The Life-Tech transducer contained an indium-gallium strain gauge that changed in resistance as the subject breathed. Resistance changes were recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 13\text{Hz}$.

Palmer skin conductance was obtained by applying a constant voltage of .5V to two 10mm Beckman Ag-AgCl electrodes filled with .075M NaCl in a Unibase medium. The electrodes were taped to the distal or middle phalanx of the first and second fingers of the left hand with adhesive collars. The signal was recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 6\text{ Hz}$.

Changes in cardiovascular activity

(cardio) were recorded from a blood pressure cuff wrapped around the right upper arm and inflated to 45 mm Hg at the beginning of each chart. The cuff was connected by rubber tubing to a Motorola MPX10DP pressure transducer in the PDAS. The output from the pressure transducer was amplified and recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 8.8\text{ Hz}$. The DC output to the analog-to-digital converter was split and sent to another channel of the PDAS where it was AC-coupled with a .2-second time constant and a 2-pole, low-pass filter, $f_c = 10\text{ Hz}$.

Finger pulse amplitude was obtained from a UFI photoplethysmograph attached to the thumb of the left hand with a Velcro strap. The signal from the photocell was AC-coupled with a .2-second time constant and a 2-pole, low-pass filter, $f_c = 10\text{ Hz}$.

The electrocardiogram was obtained from Lead I or II using disposable, pre-gelled Red Dot™ Ag-AgCl snap electrodes. The PDAS generated a 20 ms square wave pulse that coincided with the R-wave in the electrocardiogram. The square wave from the PDAS was routed to the analog-to-digital converter, and the CPS-LAB software measured and stored the time between successive pulses to the nearest ms (interbeat interval).

The finger cuff of the Finapres Blood Pressure Monitor was attached to the middle phalanx of the ring finger with Velcro. Continuous calibrated voltage changes from the Finapres Monitor were routed to a general purpose coupler on the PDAS where it was recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 10\text{ Hz}$. Voltage changes were converted to absolute blood pressure in mm Hg.

Although all channels were sampled at 1000 Hz, the data were reduced before they were stored in files on the computer harddisk by averaging the samples for successive epochs. Respiration and skin conductance channels were stored in data files at 10 Hz. Cardiograph, finger pulse, and blood pressure signals were stored at 100 Hz. The cardiometer produced an interbeat interval for each heart beat. Successive interbeat intervals were stored in the data file.

The text of test questions and associated background characters were presented to the subject on a 21" high-resolution monitor. A program running on a dedicated 166 MHz 586 PC compatible computer controlled the visual display. The computer and monitor were placed on a movable cart in the subject room. Communication between the subject computer and the CPS-LAB data collection computer was accomplished with interconnected digital I/O ports. The program running on the subject computer presented the appropriate visual display when it received digital input from the CPS-LAB computer. The CPS-LAB computer signaled the subject computer to display the question and background character when the polygraph examiner pressed the space bar and began to read the question. The examiner pressed the space bar on the CPS-LAB computer again when the subject answered. At that time, the CPS-LAB computer signaled the subject computer to remove the text of the question from the display. The background character remained on the display until the examiner presented the next question. Background characters were 120 mm in height and were bright red. The characters that comprised the question text were 25 mm in height and were white. The text was presented against a black screen background.

Procedure

If the secretary determined that the caller qualified for the study, the secretary provided a brief description of the study and the essential features of informed consent. Informed consent included identification of the government sponsor. Callers who agreed to participate were given an appointment.

The secretary instructed participants bring as many as possible of the following items with them to their appointment:

- Driver's license
- Birth certificate
- Proof of current address
- Social Security card
- Proof of college attended (diploma, transcript, or yearbook picture)
- Proof of high school attended (diploma, transcript, or yearbook picture)
- Automobile registration
- Proof of auto insurance
- Check or deposit slip from a checking account
- Credit cards

Subjects were called the evening before their scheduled appointment and reminded of the documents needed to verify their answers on the simulated employment application.

Preliminary data gathering. When the subject arrived, a research assistant described the study and obtained the subject's informed consent. The assistant then asked the subject for the requested documents. The assistant used the documents to complete a simulated employment application with 14 items. The research assistant left items blank if the subject failed to bring the necessary information. Eight completed items were selected at random, four for each of two RI tests.

The subject was then asked to complete a copy of the employment form. The subject was told to choose one of the four items that had been verified by the research assistant for each test and to falsify their answer to that item. They were then instructed to deny that they had falsified any of the items during the polygraph test. Consequently, subjects' answers concerning one item were deceptive and three items were truthful. The research assistant then prepared lists of the two sets of four relevant questions to be used on the RI tests.

The subject was informed that in order to receive the \$50 bonus, they must appear truthful to all of the questions on both polygraph tests. Whether the subject passed both polygraph tests and received the \$50 bonus was determined by a computer analysis of the polygraph charts (described below).

The research assistant directed the subject to the restroom to wash their hands with soap and water. When the subject returned from the restroom, the research assistant gave the application form and the

two lists of questions to the subject. The subject was told to report to the laboratory, introduce him or herself to the polygraph examiner, and give the materials to the examiner. Only the subject and research assistant knew the questions to which the subject had agreed to lie. The polygraph examiner was not informed until after the tests were administered and scored by the computer.

Pretest procedures. The subject was seated in a small room within the laboratory that was isolated from the recording equipment. The subject was given a brief description of how the test would be conducted, and a second standard polygraph consent form was administered. After the subject signed the form, the polygraph examiner attached the transducers and gave a brief explanation of the autonomic nervous system and its role in the detection of deception.

A numbers test was then administered to adjust the polygraph, habituate the subject to the attachments and the questioning procedure, and demonstrate the effectiveness of the technique. The subject was told to choose a number between 3 and 6, disclose it to the polygraph examiner, and to deny having selected each of the numbers 1 through 7 when asked on the test. The examiner then left the subject room, began physiological data collection, and questioned the subject about having chosen the numbers 1 through 7.

At the conclusion of the numbers test, the examiner re-entered the subject room and told the subject that the strongest reaction occurred to the number they had chosen. The subject was also told that their reactions when lying and telling the truth were easily distinguishable and they were suitable for further testing.

The subject was then shown their employment application and told that the test will focus on the information provided by the subject on the employment form. The exact wording of the first set of four relevant questions was read to the subject, and the subject was instructed to answer each question "Yes" or "No" during the polygraph test.

In addition to the four relevant questions, two neutral (irrelevant), and one sacrifice relevant question was included in the question sequence. After reviewing the relevant questions, the examiner told the subject that it would be necessary to ask some questions that the subject and examiner know will be answered truthfully. The subject was told that these questions would be used to determine if they continue to react appropriately throughout the test and continue to be a suitable subject for a polygraph test. The neutral questions were then reviewed with the subject.

Standard Audio-Only and Audiovisual Question Presentation Formats. Each subject was given two RI tests. One test was performed in the standard manner; the polygraph examiner presented the test questions only verbally. In the other test, the subject not only heard each question but also saw the text of the question on a 21" computer monitor positioned approximately 1m in front of the subject. The order in which subjects received the standard auditory and experimental audiovisual formats was counterbalanced. Half of the male and half of the female subjects received the standard audio-only format first and the audiovisual format second, and the remaining subjects received the formats in reverse order.

Audiovisual Format and Verbal Conditioning. In the audiovisual format, a large alphanumeric character that uniquely identified the specific content of each question appeared in the background, behind the text of the question. For example, a large 'B' appeared in the background for a question about the bank where the subject had a checking account. The entire text of the question and background character was presented the moment the examiner began to ask the question and it remained on the computer display until the subject answered "Yes" or "No." At that time, the text was removed from the display but the background character remained on the screen until the next test question replaced it.

Prior to conducting the polygraph test that used audiovisual presentations of test questions, the subject was trained to recognize the character uniquely associated with each

test question. The character to be associated with the three neutral and four relevant questions was presented on the computer

monitor adjacent to the text of each question as illustrated in Table 1.

Table 1. Example Computer Display of Background Characters and Associated Test Questions

7:	Are we on the <u>7</u> th floor?
L:	Do you intend to <u>lie</u> to any of the questions on this test?
B:	Did you falsify your place of <u>birth</u> ?
N:	Did you falsify your last <u>name</u> ?
A:	Did you falsify your <u>address</u> ?
Y:	Is this the <u>year</u> of 1997?
M:	Did you falsify your <u>month</u> of birth?

The subject was given one to two minutes to study the characters and associated questions. At the end of the study period, the display was turned off and the subject's memory for these associations was tested. The polygraph examiner named the characters in random order, and the subject was asked to state the question verbatim or to identify the issue addressed by the question. The examiner gave immediate verbal feedback about the correctness of the subject's response. If the subject's answer was incorrect, the examiner stated the correct answer before he proceeded to another character on the list. If the subject made any mistakes, the subject was given another opportunity to study the list. This procedure was repeated until the subject completed the entire list with no mistakes. No subject required more than two opportunities to study the list.

To explore the possibility that an analysis of the voice may be used to detect deception, audio recordings were obtained for a subsample of 45 subjects. After the verbal learning task, key terms for the seven test questions were listed on the monitor. The examiner played an audio recording that instructed the subject to make a statement about each of the topics listed. The polygraph examiner told the subject that their statements should be completely consistent with the information they reported on their

employment application; e.g., "I was born in Philadelphia, Pennsylvania." Therefore, subjects' statements concerning six of the key terms were truthful, and one statement was false. The results of our analysis of the audio recordings will be provided later.

Polygraph Testing. After completing the pretest procedures, the examiner left the subject room and presented the test questions. The minimum inter-question interval was 30 s. A set of physiological recordings was stored on the computer's hard disk for each presentation of a question sequence (chart). Two charts were collected for each test. After the first chart of each test, the subject was told to relax and was given a 1-3 minute break.

The test questions on the two charts for a given test were presented in one of 16 preprogrammed orders. The 16 orders conformed to rules of question presentation currently used by the sponsor. The particular order was determined randomly by the computer subject to the constraint that no order was used twice for a particular subject. In all cases, each of the four relevant questions was presented three times. Each relevant question appeared once on one chart and twice on the other chart. Example orders for the two tests administered to a subject are shown in Table 2.

Table 2. Question Sequences for the First and Second Polygraph Tests for a Subject

Test	Chart	Question Sequence
1	1	N1 SR R3 R1 R2 N2 N1 R4 R2 R1
1	2	N2 N1 R4 R3 N1 R1 R3 R2 R4 N2
2	1	N2 SR R3 R4 N1 R2 R1 N2 R3 R2
2	2	N1 N2 R4 R1 N1 R3 R2 R4 N2 R1

Note: (N) refers to a neutral question, and (R) refers to a relevant question.

At the completion of the first polygraph test, the examiner positioned the computer monitor in front of the subject or removed it as appropriate. He then reviewed the test questions for the second test and collected the second set of two charts. At the conclusion of the second test, the sensors were removed and the subject was escorted back to the room where they had completed the employment form. There the subject was interviewed by the research assistant and completed posttest questionnaires.

General Arousal Index. After the subject left the lab to be interviewed by the research assistant, the polygraph examiner edited artifacts from the charts and executed a computer program that evaluated the polygraph charts for a decision. For each test separately, the computer measured the amplitude of skin conductance (SC) responses, the amplitude of baseline increases in the cardiograph, and the cumulative vertical excursion of thoracic and abdominal respiration tracings following the onset of each relevant question. The 12 measurements (4 relevant questions X 3 repetitions) for each type of measurement were converted to 12 z-scores. A composite score was then computed for each presentation of a relevant question. The composite was the sum of weighted z-scores for the skin conductance (50%), cardiograph (25%), and the two respiration measurements (12.5% each). The weights assigned to the various measurements are similar to those that optimize the discrimination between truthful and deceptive subjects in comparison (control) question tests (Kircher, Raskin, Honts, & Horowitz, 1994).

The composite scores served as the dependent variable in a series of four planned comparisons. Each relevant question was

presented three times. For each relevant question separately, the mean of its three composite scores was compared to the mean of the nine scores for the remaining three relevant questions. A t-ratio was then computed for each relevant question. The numerator of the obtained t-ratio for a relevant question was the difference between the mean for that relevant question and the mean of all remaining relevant questions. The denominator was the pooled within-question standard deviation of composite scores. A large positive t-ratio for a question indicated that the subject showed a relatively strong physiological reaction to the question. This t-ratio served as the *general index of differential arousal* for a relevant question.

Computer Decisions. A one-tailed t-test was performed to determine if the mean reaction to the target relevant question was significantly greater than the mean for the remaining three questions ($t > 1.397$, $p < .10$). If the general arousal index for a relevant question exceeded this cutoff, the subject was considered deceptive to the question and was ineligible for the \$50 bonus. When the general arousal indices for all eight relevant questions on the two tests were non-significant, the subject was paid the \$50 bonus in addition to their regular pay of \$20. The criterion t-ratio was established during pilot testing to achieve a 20% chance (approximately) that the subject would pass both tests and receive the bonus.

Debriefing and Post-test Interview. While the physiological data were being analyzed, a research assistant interviewed the subject in an attempt to discover the strategies they used to appear truthful on their tests. For each type of test, subjects were asked to recall as many as possible of the thoughts they had and techniques they used in an effort to

avoid detection. For each technique, they were asked how often they used it and were asked to rate how effective they thought the technique was in helping them to escape detection. The subject then completed a questionnaire that listed several types of thoughts and techniques reported by pilot subjects and by subjects in prior research on spontaneous countermeasures (Honts, Raskin, Kircher, & Hodes, 1988). The questionnaire was completed once for the audio-only test and again for the audiovisual test. For each test and question category, subjects rated the percent of time they used the countermeasure and its perceived effectiveness.

Results

Effects of audiovisual presentations of test questions were assessed at three levels of analysis: computer decisions, the general arousal index, and individual physiological response waveforms.

Effects of Audiovisual Presentations of Test Stimuli on Computer Decisions

As described above, computer decisions were based on general indices of differential reactivity to the four relevant questions on a polygraph test. The indices of differential arousal were t-ratios from the four statistical comparisons of each relevant question to the mean of the three remaining relevant questions. A decision was classified as correct if the relevant question to which the subject was deceptive produced the greatest differential arousal score, and that score exceeded a predetermined critical value ($t=1.397$, one-tailed $p < .10$). A decision was classified as a false positive if a relevant question to which the subject was truthful produced the greatest general arousal score,

and that score exceeded the critical value. A decision was classified as a false negative if none of the four arousal scores exceeded the critical value. Percent outcomes are presented in Table 3.

There was little difference between Audio Only and Audiovisual formats in decision outcomes. Both formats yielded slightly less than 50% correct decisions and about 7% false positive errors. McNemar's test of the difference between Audio Only and Audiovisual conditions revealed no significant difference for correct decisions or for false alarms. At the level of computer decisions, there was no evidence that audiovisual presentations of test questions improved the accuracy of diagnoses.

Although the percent correct decisions was less than 50% for both types of tests, it should be recalled that the critical value of t was arbitrarily set at 1.397 to achieve a cumulative false negative rate of 20% on both tests. That is, we planned to pay the bonus to 20% of our subjects, and subjects were paid only if they passed both tests. The multiplication rule of probability for independent events states that the joint probability of a false negative result (pass) on both tests would be $\text{Pr}(\text{false negative on first test}) \times \text{Pr}(\text{false negative on the second test}) = .448 \times .469 = .21$.

Although false negative events on the two tests may not have been conditionally independent, the product of the two individual probabilities (.21) approached our target value (.20). The point here is that it was necessary to have high false negative rates on the two

Table 3. Percent Outcomes Based on Computer Decisions for Audio Only and Audiovisual Presentations of Test Questions

	Correct	False Positive	False Negative
Audio Only	49.0	6.3	44.8
Audiovisual	46.9	7.3	45.8

tests to ensure that we would award the \$50 bonus in approximately 20% of the cases.

Effects of Audiovisual Presentations of Test Stimuli on the General Arousal Index

It is well known that non-parametric tests of outcomes in cross-classification tables are often less powerful than parametric tests of the continuous measures that underlie those classifications. To explore the possibility that a parametric test might be more sensitive to effects of audiovisual presentations of test questions, the general arousal index for the question answered deceptively served as the dependent variable in a Sex X Test Format split-plot ANOVA. Sex was a between-group factor with two levels (male and female); and Test Format was a within-subjects factor with two levels (audio-only and audiovisual). Table 4 shows the means and standard deviations of the arousal index for males and females in the audio-only and audiovisual conditions.

ANOVA revealed no significant difference between audio-only (M=1.67) and audiovisual (M=1.78) arousal scores. For females, it appeared that greater discrimination between truthful and deceptive responses was found in the audio-only condition (M=1.52) than the audiovisual condition (M=1.30). Conversely, for males, it appeared that the reverse was true. However, the Sex X Test Format interaction was not significant, $F(1,94) = 2.99, p < .09$. Interestingly, there was a main effect of Sex. Overall, males reacted more strongly when they lied (M=2.04) than did females (M=1.41), $F(1,94) = 5.53, p < .03$. In general, Test Format did not significantly affect general arousal scores.

Effects of Audiovisual Presentations of Test Questions on Electrodermal and Cardiovascular Response Waveforms

The general arousal index and computer decisions were based on an a priori weighted combination of SC, cardiograph, and respiration features. The choice of features was based on prior research with comparison question tests and mock crime scenarios. In the present study, RI tests were conducted concerning answers on a simulated employment application form. Since the situations differed in several respects, the selected response parameters may have been insensitive to the effects of audiovisual presentations of test questions. Therefore, additional analyses of individual physiological waveforms were performed to test for differences between the audio-only and audiovisual formats.

Second-by-second response waveforms were generated for SC, cardiograph systolic levels, cardiograph diastolic levels, decreases in finger pulse amplitude, Finapres systolic levels, Finapres diastolic levels, and heart rate. In general, the computer extracted its measurements from the response to each relevant question for a period of 20 seconds following question onset. Each physiological response was defined as a series of 20 points, one for each poststimulus second. Each point in the response waveform was a weighted mean level of the measurements made during that one-second interval (Kircher & Raskin, 1988). In addition to the second-by second measures, successive measurements of vagal tone were derived for successive 5-second epochs that began four seconds prior to question onset.

For each subject, the mean response waveform was computed for the three presentations of the relevant question

Table 4. Means (and Standard Deviations) of the General Index of Differential Arousal

	Females	Males	Marginal
Audio-Only	1.52 (1.36)	1.83 (1.50)	1.67
Audiovisual	1.30 (1.66)	2.26 (1.92)	1.78
Marginal	1.41	2.04	1.73

answered deceptively. Another mean response waveform was computed for the nine presentations of relevant questions answered truthfully. The two mean response waveforms per subject were analyzed with a Sex X Test Format X Deception X Seconds split-plot ANOVA. Sex was a between-group factor with two levels (male and female); Test Format was a within-subjects factor with two levels (audio-only and audiovisual); and Deception was a within-subject factor with two levels (RQs answered truthfully and RQs answered deceptively). With the exception of vagal tone, Seconds was a within-subjects factor with 20 levels. In the case of vagal tone, the Seconds factor was replaced by Epochs with 6 levels. To correct for bias due to violations of the sphericity assumption, Geisser-Greenhouse corrected (GGC) p-values (conservative) were used for all statistical tests that involved repeated measures with more than two levels (Keppel, 1992).

Only one of eight ANOVAs revealed any main or interaction effects that included Test Format as a factor. The ANOVA of SC responses yielded a significant Test Format X Deception X Seconds interaction, $F(19, 1786) = 4.45$, GGC $p < .01$. Visual inspection of plots of means suggested that the SC responses associated with *truthful* answers to relevant questions recovered less quickly in the audiovisual condition than in the audio-only condition. Otherwise, the audiovisual presentations of test questions had little or no effect on SC and cardiovascular response waveforms.

Effects of Deception on Electrodermal and Traditional Cardiovascular Response Waveforms

Figure 1 through Figure 4 show mean second-by-second electrodermal and traditional cardiovascular responses to relevant questions answered deceptively and relevant questions answered truthfully. To produce a common starting point for the response curves in a figure, the level during the first second was subtracted from each poststimulus level that defined the response curve. The figures also show response waveforms for neutral questions, although the data for neutral questions were not included in the ANOVAs. Responses to neutral questions

were provided as a baseline against which responses to relevant questions might be compared visually.

The Sex X Test Format X Deception X Seconds split-plot ANOVA of skin conductance (described above) revealed a significant main effect of Deception, $F(1,94) = 34.0$, $p < .01$, as well as a significant Deception X Seconds interaction, $F(19,1786) = 24.59$, GGC $p < .01$. As shown in Figure 1, skin conductance responses to relevant questions were greater when subjects lied than when they told the truth.

Diastolic cardiograph responses are presented in Figure 2. Contrary to expectations, analyses of diastolic cardiograph responses revealed no main effect of Deception, nor was there evidence of any interaction effects that included Deception as a factor. Responses to neutral questions appeared stronger than responses to relevant questions. This was an artifact of the position of neutral questions in the question sequence. The first and often second position of every question sequence was occupied by a neutral question. During the initial 45 s to 60 s seconds of a chart, the cuff pressure rose as the volume of blood in the arm distal to the cuff increased and then stabilized at a relatively constant baseline level.

Systolic cardiograph responses are shown in Figure 3. ANOVA revealed a main effect of Deception, $F(1,94) = 4.25$, $p < .05$. As predicted, relevant questions answered deceptively evoked stronger systolic cardiograph responses than relevant questions answered truthfully. Deception did not interact with Sex, Test Format, or Seconds.

Second-by-second changes in finger pulse amplitude are shown in Figure 4. The data were expressed as proportions of the amplitude of the finger pulses at stimulus onset. ANOVA revealed a main effect of Deception on the overall amplitude of finger pulses across the 20-second interval, $F(1,94) = 7.51$, $p < .01$. Relevant questions answered deceptively produced greater vasoconstriction in the finger than did relevant questions answered truthfully. In addition, the Deception X Seconds interaction was significant, $F(19,1748) = 4.95$, GGC $p < .01$.

Examination of Figure 4 suggests that the interaction was due primarily to the rate at which the amplitude of finger pulses reached the minimum. The reduction in the amplitude of finger pulses occurred more quickly when the subject was deceptive than when the subject was truthful.

Effects of Deception on New Cardiovascular Response Waveforms

Second-by-second plots of diastolic and systolic blood pressure as measured in the finger by the Finapres device are shown in Figure 5 and Figure 6. ANOVA revealed significant main effects of Deception on both diastolic, $F(1, 94) = 6.24, p < .02$, and systolic blood pressure waveforms, $F(1, 94) = 26.11, p < .01$. The Deception X Seconds interactions were also significant for diastolic, $F(19, 1786) = 6.24, GGC p < .01$, and systolic blood pressure waveforms, $F(19, 1786) = 8.90, GGC p < .01$. Examination of Figure 5 indicates that diastolic blood pressure increased about 1 mm Hg more when the subject was deceptive than when the subject was truthful. For systolic blood pressure, the difference between truthful and deceptive responses was about 4 mm Hg (Figure 6).

Plots of heart rate for relevant and neutral questions are presented in Figure 7. ANOVA yielded a significant main effect for Deception, $F(1,94) = 120.93, p < .01$, as well as a significant Deception X Seconds interaction, $F(19, 1786) = 25.29, GGC p < .01$. The pattern of differences between relevant questions answered truthfully and relevant questions answered deceptively is consistent with prior research with comparison question tests (Raskin, 1979). Deception was associated with a small initial increase in heart rate followed by a pronounced deceleration and slow recovery to baseline. In contrast, truthfulness was associated with an initial increase in heart rate that recovered about 10 seconds following question onset.

Vagal tone was measured every five seconds for a period of 30 seconds that began four seconds prior to question onset and ended 24 seconds after question onset. Differences from prestimulus vagal tone for relevant questions and neutral questions are shown in Figure 8. Vagal tone was measured

in natural logarithms of the variance in heart period associated with normal adult respiration frequencies (Porges et al., 1980). ANOVA indicated that the main effect of Deception was significant, $F(1, 94) = 22.66, p < .01$. On average, subjects showed more vagal tone when they answered relevant questions deceptively than when they answered relevant questions truthfully.

The Deception X Epochs interaction was also significant, $F(5, 470) = 2.94, GGC p < .02$.

For relevant questions answered deceptively, vagal tone increased then stabilized at the higher level. For relevant questions answered truthfully, vagal tone dropped slightly late in the scoring window.

Exploratory Analyses of Response Waveform Features

The CPSLAB software permits exploration of multiple features of response waveforms. The procedures for measuring features that may be extracted from response waveforms are included in Appendix A. Exploratory analyses of multiple features were conducted in an attempt to identify a diagnostic subset of measures that might improve upon the accuracy achieved by the general arousal index. These exploratory analyses were justified on the grounds that the general arousal index was an *a priori* composite of measures known to be useful for comparison question tests. To our knowledge, little exploratory research has been conducted to identify computer measurements that might be of value for RI tests.

To assess the diagnostic validity of a given feature on a given polygraph test, the 12 measurements of that feature (4 relevant questions X 3 repetitions) were correlated with a dichotomous variable (point-biserial correlation, r_{pb}). The dichotomous variable distinguished between the three repetitions of the relevant question answered deceptively (coded 1) and the nine repetitions of relevant questions answered truthfully (coded 0). A high positive point-biserial correlation indicated that the three measurements obtained for the relevant question answered deceptively were greater than the nine measurements of relevant questions answered

truthfully. That is, high scores on the dichotomous variable X ($X=1$) were associated with high scores on Y (measurements of deceptive responses), and low scores on X ($X=0$) were associated with low scores on Y. Conversely, a negative r_{pb} indicated the extent to which deceptive responses were associated with relatively small measured responses.

By design, the r_{pb} was attenuated because there were three times as many relevant questions answered truthfully as there were relevant questions answered deceptively. Therefore, the r_{pb} was adjusted upward or downward as necessary to correct for attenuation (Cohen & Cohen, 1983).

Since two RI tests were conducted on each of the 96 subjects, a total of 192 within-subject r_{pb} were computed for each feature extracted by CPSLAB. The mean within-subject r_{pb} for 23 features of skin conductance, cardiograph, finger pulse amplitude, and blood pressure (Finapres) responses are presented in Table 5.

In addition to the features listed in Table 5, the sum of absolute differences between adjacent samples of the thoracic respiration recording (TR excursion) was measured from response onset for 10 seconds, as was abdominal respiration (AR) excursion. The r_{pb} for TR excursion and AR excursion were -0.353 and -0.255 , respectively. Since deception is usually associated with suppressed respiratory activity (less excursion), negative correlations were expected for the respiration measures.

To measure heart rate deceleration, the mean heart period for two seconds prior to question onset was subtracted from the heart period for each of the 20 poststimulus seconds. A difference greater than zero indicated that the interval between heartbeats had increased (heart rate decreased). Differences less than zero were set to zero. Peak amplitude of this derived cardiac deceleration curve was the point of maximum deceleration. Peak amplitude and area to full recovery for the cardiac deceleration curve yielded r_{pb} of 0.270 and 0.240 , respectively.

For vagal tone, point-biserial correlations were computed for each poststimulus epoch. The r_{pb} ranged from a low of 0.053 for the first epoch (0-4s) to a high of 0.131 for the third epoch (5-9 s; see Figure 8).

Variable Selection Procedures.

Based on examinations of the within-subject point-biserial validity coefficients and intercorrelations among the various measures, 18 variables were selected for possible inclusion in a new general arousal index. The 18 variables are listed in Table 6.

In the initial assessments of features described above, the measurement window was the same for all physiological channels. It was arbitrarily set to range from the moment of question onset to 20 seconds after question onset. To improve discrimination between truthful and deceptive responses, the end points of scoring windows were adjusted in 1-second increments to maximize the r_{pb} for 5 of the 18 measures. The five measures were thoracic and abdominal respiration excursion, skin conductance amplitude, systolic blood pressure amplitude (Finapres), and finger pulse area to half recovery.

An index of differential arousal was computed for each polygraph test and each variable in Table 6. The procedure for calculating an index of differential arousal for a particular feature was similar to the method used for the general arousal index. For a given test and waveform feature, the mean of nine measurements for relevant questions answered truthfully was subtracted from the mean of the three measurements for the relevant question answered deceptively. The difference between the means was divided by the pooled within-question standard deviation. Again, this differential arousal index was a t-ratio that measured the extent to which reactions to relevant questions answered deceptively could be distinguished from relevant questions answered truthfully. For each variable, Table 6 shows the mean and standard deviation of arousal indices for the 192 polygraph examinations.

Table 5. Mean Within-Subject Point-Biserial Correlations (Validity Coefficients) for Electrodermal and Cardiovascular Waveforms

Feature	SC ^a	CDD ^b	CDS ^c	BPD ^d	BPS ^e	FPA ^f
Peak amplitude	0.398	-0.001	0.016	0.064	0.101	0.155
Area to half recovery	0.375	-0.014	0.014	0.038	0.092	0.178
Area to full recovery	0.382	0.002	0.018	0.049	0.117	0.181
Area to peak amplitude	0.336	-0.013	0.019	0.000	0.046	0.074
Area from peak to full recovery	0.383	0.022	-0.003	0.085	0.128	0.204
Risetime from response onset	0.174	-0.029	0.010	-0.025	-0.003	-0.027
Risetime from first low point	0.076	-0.031	-0.011	-0.090	-0.063	-0.084
Half recovery time	0.219	0.021	0.004	0.062	0.104	0.195
Full recovery time	0.226	0.048	0.006	0.082	0.102	0.121
Duration to half recovery	0.217	-0.023	0.018	0.009	0.056	0.105
Duration to full recovery	0.245	0.017	0.028	0.041	0.091	0.087
Latency to first low point	-0.083	0.004	-0.007	-0.025	-0.049	0.009
Latency to response onset	-0.181	-0.008	-0.033	-0.061	-0.083	-0.050
Riserate from onset of epoch	0.363	0.037	0.023	0.123	0.130	0.159
Riserate from response onset	0.309	0.030	0.008	0.070	0.038	0.068
Half recovery rate	0.240	-0.001	0.008	0.051	-0.017	-0.049
Full recovery rate	0.253	-0.008	0.009	0.000	0.000	0.000
Excursion	0.340	-0.076	-0.058	0.026	-0.005	-0.009
Number of responses	-0.011	-0.150	-0.139	-0.060	-0.140	-0.045
Burst frequency	-0.025	-0.112	-0.143	-0.091	-0.150	-0.002
Mean of successive amplitudes	0.381	0.043	0.053	0.047	0.100	0.026
Std Dev of successive amplitudes	0.092	0.020	0.032	0.074	0.100	0.149
Level	0.188	0.002	0.004	0.100	0.179	0.074
Standard deviation	0.383	-0.006	0.015	0.066	0.085	0.122

^aSkin Conductance, ^bCardiograph Diastolic, ^cCardiograph Systolic, ^dFinapres Diastolic Blood Pressure, ^eFinapres Systolic Blood Pressure, ^fFinger Pulse Amplitude (reflected)

The third column of values in Table 6 reports the results of a single-sample t-test that compared the mean index of differential arousal shown in the first column to 0.0. Zero was the expected value of the difference under null conditions. The results of those tests indicated that all 18 variables reliably discriminated between truthful and deceptive responses.

The fourth column of Table 6 reports the mean within-subject point-biserial correlation (r_{pb}) after adjusting the scoring windows to maximize the r_{pb} for selected key variables

As expected, the most diagnostic measure was skin conductance amplitude. Respiration excursion measures were also highly diagnostic, as were measures of the magnitude of cardiac deceleration. It may be recalled that the general arousal index, which was used to make decisions, was based on SC amplitude, respiration excursion, and amplitude of baseline increases in the cardiograph recordings. Except for the use of the cardiograph measure, it appears that the variables for making decisions were reasonably well chosen.

Table 6. Diagnosticity Statistics and Tests for Significance for 18 Waveform Parameters

Channel	Feature	Mean	SD	t ¹	r _{pb}
TR	Excursion	-1.42	1.23	-11.28	-0.38
AR	Excursion	-0.99	1.23	-7.89	-0.27
SC	Amplitude	1.52	1.09	13.60	0.44
SC	Half recovery time	0.77	0.92	8.22	0.23
SC	Latency	-0.56	1.08	-5.08	-0.19
BPD	Amplitude	0.39	1.09	3.51	0.10
BPD	Rise rate	0.32	0.97	3.28	0.09
BPD	Level	0.47	0.91	5.05	0.13
BPS	Amplitude	0.51	1.13	4.48	0.13
BPS	Rise rate	0.43	0.94	4.47	0.12
BPS	Level	0.68	1.03	6.53	0.19
FPA	Amplitude	0.68	1.22	5.50	0.17
FPA	Area to half recovery	0.82	1.26	6.36	0.20
FPA	Area peak to full recovery	0.61	1.16	5.16	0.17
FPA	Rise rate	0.44	0.99	4.37	0.13
HR	Max deceleration	0.79	1.09	7.09	0.21
HR	Area to full recovery	1.00	1.36	7.20	0.24
HR	Vagal tone	0.44	0.96	4.48	0.05

¹All t-ratios were significant at $p < 0.01$.

Selection of an Alternative Subset of Physiological Measures for Discriminating Between Truthful and Deceptive Responses

All-possible-subsets regression analysis was used to select from the available pool of 18 arousal indices a small subset that optimized discrimination between truthful and deceptive responses to relevant questions. Ordinarily, the criterion to be predicted by a set of physiological measures is a dichotomous variable that distinguishes between truthful and deceptive subjects. Since all subjects in the present study were deceptive, it was necessary to create a sample of 'truthful' cases.

Deceptive and truthful cases were obtained from the available sample in the following manner. Each subject had been given two polygraph tests. One test from each subject was considered a deceptive case, and the other test was considered a truthful case. If the test was treated as a deceptive case, one of the three relevant questions answered truthfully was selected at random and discarded. Reactions to the relevant question answered deceptively were then compared to the data for the two remaining questions, both

of which had been answered truthfully.

If the test was treated as a truthful case, the relevant question answered deceptively was discarded. The three remaining relevant questions were answered truthfully. One of the three remaining questions was selected at random, and reactions to that question were then compared to the reactions to the other two questions. Thus, for all cases, indices of differential reactivity were t-ratios from planned comparisons of one question to the mean of two others. For deceptive cases, a question answered deceptively was compared to the mean of two questions answered truthfully. For truthful cases, a question answered truthfully was compared to the mean of two other questions answered truthfully. The cases assigned to truthful and deceptive categories were balanced in terms of Sex, Order of Presentation, and Test Format (audio only vs. audiovisual).

Since each of the 96 subjects contributed two polygraph tests, 192 cases were available for the regression analysis. The sample was split in half, and each half contained 48 truthful and 48 deceptive cases.

Again, split halves were balanced in terms of Sex, Order of Presentation, and Test Format.

The two split halves were independently analyzed with the all-possible-subsets program available in BMDP. This algorithm begins by ranking, in order of proportions of variance explained, the top 10 subsets composed of one predictor variable. It then ranks the top 10 subsets composed of two predictor variables, then the top 10 subsets composed of three variables, and so on. The algorithm uses several criteria to choose an optimal model for predicting the criterion (BMDP Manual Version 7.0, 1992).

In one split half, the best subset consisted of seven variables. In the other sample, the best subset consisted of six variables. A model was chosen with five predictors variables, four of which were common to the top-ranked subsets for the two split halves. The 5-variable model was ranked second among subsets of five variables in one split half, and it ranked fourth among subsets composed of five variables in the other split half.

The selected variables were then used to create a regression equation to predict the criterion in one split half (standardization sample). The remaining subjects were set aside to validate the model. The variables are listed in Table 7 along with their simple correlations with the criterion, standardized regression coefficients, and t-ratios that reflect their relative contributions to the model.

The 5-variable model was dominated by

skin conductance measures. When subjects were deceptive, their skin conductance responses had greater amplitude, longer recoveries, and shorter latencies to response onset than when they were truthful. Deception was also associated with faster reductions in finger pulse amplitude and greater respiratory suppression. In the standardization sample of cases, the 5-variable model accounted for 46% of the variance in the criterion.

Decision Accuracies for the Original 3-Variable Model and the New 5-Variable Model

A new general arousal index was computed using the standardized regression coefficients as weights for the five variables selected for the new model using the procedures described earlier for the original 3-variable model. The new general arousal index was used to classify cases in the validation sample. As before, the decision was considered correct if the general arousal index for the relevant question answered deceptively was greater than the arousal index for any other question and it exceeded some critical value. The decision was considered a false positive error (false alarm) if the general arousal index for a relevant question answered truthfully yielded the greatest score and that score exceeded the critical value. The decision was considered a false negative error (miss) if no relevant question produced a general arousal score that exceeded the critical value.

Correct decisions and false positive errors for the original 3-variable model and the

Table 7. Univariate and Multivariate Correlation/Regression Analyses of Selected Variables for Predicting Truth and Deception in the Standardization Sample (96 polygraph tests from 48 subjects)

Variable	Correlation with Criterion	Std. Coefficient	Reg. t-ratio
SC Amplitude	0.49	0.41	5.07
SC Half Recovery Time	0.24	0.13	1.67
SC Latency	-0.23	-0.25	-3.07
FPA Rise Rate	0.24	0.16	2.01
TR Excursion	-0.39	-0.33	-4.12

new 5-variable model are plotted in Figure 9 as a function of different critical values. Two curves are shown for the new model, one for the standardization sample ($n=48$) and one for the validation sample ($n=48$). The results for the original 3-variable model were obtained from the entire sample of 96 subjects. The vertical line at the critical value of 1.397 marks the cutoff used with the original 3-variable model to decide if a subject failed a test.

As expected, both detection and false positive rates decreased as greater evidence of differential reactivity to a relevant question (the critical value) was required to make a decision. At a critical value of 0.0, 84% of the lies by subjects in the validation sample were detected. Sixteen percent of the subjects in the validation sample were considered deceptive to the wrong relevant question (false positive). When the critical value was 0.0, there were no misses; that is, every subject was classified as deceptive to one question or another. With a critical value of 0.0, it was a foregone conclusion that every subject would be classified as deceptive because the mean for at least one relevant question had to exceed the mean of the remaining three questions. This situation resulted in at least one positive, general arousal index (t -ratio) that exceeded the critical value of 0.0.

To make a decision with a critical value of 2.0, the mean for a relevant question had to be 2.0 or more standard deviations above the mean of the remaining three. At that extremely conservative cutoff, neither the 3-variable model nor the 5-variable model produced any false positives errors. However, the original 3-variable model produced only 38% correct decisions, and missed 62% of the lies. The 5-variable model produced 48% correct decisions and missed 52% of the lies.

It may be seen that the 5-variable model was uniformly more sensitive (produced more correct decisions) and more specific (produced fewer false positive errors) than the original 3-variable model. Overall, the 5-variable model increased the hit rate by about 10% and reduced the rate of false alarms by about 10%. Where the 5-variable model produced just under 10% false positive errors ($t = .80$), 80.2% of its decisions were correct,

8.3% were false positives, and 11.5% were false negatives on cross-validation. At the same cutoff, the original 3-variable model produced 65.1% correct, 18.7% false positives, and 16.2% false negatives.

It may also be seen that there was no loss in decision accuracy when the 5-variable model was cross-validated. The weights that were optimal for the standardization sample worked about as well as for the validation sample.

Finally, it should be noted that substantial improvement in the performance of the 3-variable model would have been achieved by simply dropping the cardiograph measure from the model and basing decisions on only skin conductance amplitude and respiration excursion. It may be recalled that cardiograph responses did not distinguish between truthful and deceptive responses in the present study. Consequently, inclusion of the cardiograph measure added error variance to the general arousal index and reduced its predictive validity. The hit and false alarm curves for the model with only skin conductance amplitude and respiration excursion were approximately mid-way between the correct and false alarm curves for the 3-variable and 5-variable models shown in Figure 9.

Spontaneous Countermeasures

The spontaneous use of countermeasures was assessed first by means of an interview then by computerized questionnaires. The primary purpose of the interview was to determine if subjects would report using countermeasures that were not included on the computer questionnaire. The interview was conducted prior to the computer questionnaire to ensure that these reports would not be influenced by exposure to the computer questionnaire.

The interviewer asked subjects what they had done to defeat the polygraph when the questions were presented on the monitor and what they had done when the monitor was absent. The interviewer was non-directive, non-suggestive, and solicited information from subjects with prompts such as, "Did you do anything else to try to appear truthful?" and

“Tell me more about that.” When the subject described a countermeasure, they were asked how often they used it and to rate its perceived effectiveness. Effectiveness was assessed on a 5-point scale that ranged from *not effective* to *highly effective*. The subject was also asked to indicate if they had used the countermeasure on falsified items, truthful items, or both.

Eighty-three subjects reported using at least two countermeasures during each of their polygraph tests. The remaining 13 subjects reported using only one countermeasure. The one or two countermeasures each subject reported using most often were included in the analysis. The countermeasures described by subjects could be broadly classified as mental countermeasures, physical countermeasures that involved attempts to inhibit or control

some physiological system such as respiration, and physical countermeasures that involved attempts to activate some physiological system. The three global categories were further subdivided into the 15 subcategories listed in Table 8.

A research assistant who had conducted most of the interviews coded all of the interviews. Another assistant coded 24 of the interviews with female subjects. Each assistant classified each statement by a subject into one of the 15 subcategories. The inter-rater agreement was 74%.

Mean utilization and effectiveness ratings are presented in Table 8 for each category. Percentage use statistics do not sum to 100% because subjects often reported using more than one type of countermeasure.

Table 8. Countermeasure Utilization Derived from Interviews

	Audio Only		Audiovisual	
	% Use	Effective	% Use	Effective
Mental Countermeasures	7.3	3.6	9.0	3.6
Faulty Thinking: e.g., convincing oneself that no lie had been told	10.3	3.6	7.2	3.5
Resignation: e.g., believing that one could not defeat the polygraph	2.5	3.3	2.6	3.0
Attention Focussed Inward: the tendency to think about something very specific but external to the testing situation	4.8	4.2	2.3	4.3
Mental Work: a mental activity, such as mental arithmetic, that is not tied to a specific external referent	18.3	3.6	24.2	3.5
Attention Focussed Outward: attention focussed on an object in the room; e.g., looked at the carpet to avoid the monitor	8.0	3.5	16.8	3.6
Other mental activity	0.0	0.0	1.0	4.0
Inhibitory Countermeasures	9.0	3.5	6.5	3.8
Controlled breathing	18.2	3.4	13.3	3.6
Controlled heart rate	4.5	3.4	3.0	3.7
Controlled speech: e.g., spoke in monotone	4.3	3.8	3.9	4.3
General relaxation	16.0	3.4	10.6	3.7
Other physical control	2.0	3.5	1.9	3.4
Activating Countermeasures	1.0	3.6	1.0	2.8
Activation of breathing	1.1	3.5	1.4	3.5
Activation of heart rate	1.0	3.0	1.0	2.0
General arousal	1.8	3.3	1.4	3.0
Other physical activity	1.0	5.0	0.0	0.0

Paired t-tests revealed that *attention focussed inward*, $t(95) = -2.18$, $p < .05$, and *attention focussed outward*, $t(95) = -3.74$, $p < .01$ were used more often in the audiovisual condition than in the audio only condition. In contrast, *faulty thinking*, $t(95) = 2.10$, $p < .05$, *controlled breathing*, $t(95) = 3.05$, $p < .01$, and *general relaxation*, $t(95) = 2.72$, $p < .01$, were used more often in the audio only condition. No other comparisons of audiovisual and audio only conditions were significant.

To determine if a particular type of countermeasure was effective in defeating the test, a dichotomous variable was created to indicate if the subject reported using a particular countermeasure ($X=1$) or not ($X=0$). This dichotomous variable was correlated with another dichotomous variable that indicated if the outcome of the polygraph test was correct ($Y=1$) or not ($Y=0$). A third dichotomy was created to indicate if the outcome of the polygraph test was a false positive. Correlations were computed separately for audiovisual and audio only conditions. None of these correlations was significant. There were also no significant differences between male and female subjects in the use of any of these countermeasures.

Countermeasures applied selectively while answering truthfully or deceptively to test questions may be more effective than countermeasures applied generally throughout the polygraph test. To evaluate this possibility, a dichotomous variable was created to indicate if the subject stated that they had used a particular countermeasure selectively during the polygraph test. This variable was then correlated with one measure that indicated if the decision was correct and another measure that indicated if the decision was a false positive. Neither correlation was significant for either the audio only or audiovisual treatment conditions.

After the interview, subjects completed two computerized questionnaires. One computerized questionnaire asked subjects which countermeasures they had used to appear truthful in the audiovisual condition, and another questionnaire asked subjects which countermeasures they had used in the audio only condition. Each questionnaire assessed frequency of use and perceived effectiveness of 10 countermeasures. Frequency of use and perceived effectiveness were rated on 11-point scales that ranged from 0% to 100% in 10-point increments. The anchors for the perceived effectiveness scale were *completely ineffective* and *extremely effective*. The 10 countermeasures and mean ratings are presented in Table 9 for the audio only and audiovisual conditions.

To compare audiovisual and audio only conditions, a separate paired t-test was performed for each of the 10 items on the computer questionnaire. Consistent with the findings obtained from the interview data, subjects reported using *self-deception* (faulty thinking), $t(95) = 3.72$, $p < .01$, and *general relaxation*, $t(95) = 2.19$, $p < .05$, more often during the audio only test than the audiovisual test. No other differences were significant.

A dichotomous variable was created for each of the 10 items on the computer questionnaire to indicate if the subject had used the countermeasure. None of those variables correlated significantly with correct decisions or false positive outcomes. Nor was there any evidence that the responses of male and female subjects differed to any of the items on the computer questionnaires.

Table 9. Countermeasure Utilization Assessed by Computer Questionnaires

Countermeasure	Audio Only		Audiovisual	
	% Use	% Effective	% Use	% Effective
Rationalization: e.g., told oneself, "I'm not telling a real lie."	39.7	45.3	37.6	46.0
Self-deception: e.g., told oneself, "I falsified none of my answers on the form."	40.9	43.9	34.0	39.3
Dissociation: tried to blank the mind and think about nothing	39.0	47.1	41.9	49.6
Imagined a peaceful scene to produce relaxation	22.5	47.3	20.6	47.7
Imagined a fearful or anger-provoking situation	9.3	30.4	6.8	33.5
Relaxation	63.5	48.8	59.5	49.7
Controlled breathing	52.8	47.2	52.5	47.2
Controlled heart rate	29.7	39.1	28.1	43.3
Produced pain: e.g., bit tongue	2.6	26.0	1.4	12.5
Tensed muscles: e.g., pressed toes against the floor	5.5	18.6	5.2	22.8

Discussion

Audiovisual versus Audio-Only Presentations of Test Questions

The primary objective of the present study was to explore possible benefits of presenting questions visually as well as aurally during RI tests. Effects of audiovisual presentations of test questions were assessed at the level of decisions, a composite measure of the relative strength of response to relevant questions answered deceptively, and at the level of individual response waveforms. In general, none of the analyses revealed any advantage to audiovisual presentations of test questions.

At the level of decisions, there were no significant differences between audio-only and audiovisual presentations of test questions in correct decisions or false positive outcomes. The two test formats differed by less than 2.2% in correct decisions, false positives, and false negatives. Audiovisual tests yielded slightly fewer correct decisions than audio-only tests (46.9% vs. 49.0%) and slightly more false positive outcomes (7.3% vs. 6.3%).

Decisions were based on composite measures of differential skin conductance,

respiration, and cardiograph responses to relevant questions. Analysis of these composite scores also failed to reveal any effects of audiovisual presentations of test questions. Interestingly, males showed significantly greater discrimination between truthful and deceptive responses than did females. There was also a tendency for males to be more reactive when they lied in the audiovisual condition and for females to be more reactive when they lied in the audio-only condition, although the interaction was not statistically significant.

The features of skin conductance, respiration, and cardiograph measures used for the composite measure of differential reactivity were based on prior research with comparison-question tests. Since RI tests were conducted in the present study, it was possible that effects of audiovisual presentations might be revealed in other characteristics of the physiological response waveforms or in other channels of physiological activity altogether. Therefore, repeated measures ANOVAs were performed on individual response waveforms. No effects of audiovisual presentations were observed in any waveform analyses of cardiograph, blood pressure, finger pulse amplitude, heart rate, or vagal tone. One significant result was

obtained for skin conductance responses. Skin conductance responses to relevant questions answered truthfully recovered less quickly in the audiovisual condition than in the audio-only condition. This effect may indicate that subjects paid attention to the background character that remained on the screen after the subject answered a relevant question truthfully. However, the effect was small (estimated $\eta^2 = .05$), and in light of the number of main and interaction effects tested in these analyses, this one significant finding may have been due to chance.

The failure to observe effects of audiovisual presentations of test questions in the present study is consistent with prior research that compared aural and visual presentations of test items (Biejk, 1980; Carlson & Smith, 1991). Null results were obtained in the present study despite numerous improvements in the design and analysis of physiological measures. The present study was designed to enhance the effects of visual presentations of test questions and the utility of findings for field applications. In contrast to prior research, the present study examined the combined effects of auditory and visual presentations rather than presenting test items only aurally or only visually. It examined the effects of audiovisual presentations using a common field technique (RI test) rather than a guilty knowledge test. Each test question was presented against a large background character that had been associated with the question using a pretest question recognition task. The background character remained on the screen after the subject answered and the text of the question was removed. The background character was intended to stimulate retrieval of the question and interfere with attempts to dissociate during the test.

An attempt was made to achieve a high degree of statistical power (sensitivity) to detect effects of audiovisual presentations of test questions. The sample was large. One hundred and ninety-two polygraph tests were administered to 96 subjects. Statistical tests focused on within-subject differences between audio-only and audiovisual conditions, and within subjects designs are typically more powerful than between-group designs. Finally, there were no ceiling effects to limit our ability

to detect a meaningful change in the accuracy of test outcomes. Both test formats yielded detection rates that exceeded change levels of accuracy, but there was still considerable room for improvement.

In light of the present results and those of previous investigations, it does not appear that visual presentations of test questions enhance detectability or offer any clear advantage over the simpler and less costly aural presentation format.

Discrimination between Truthful and Deceptive Responses to Relevant Questions

The ANOVA of skin conductance and all cardiovascular response waveforms except one of two cardiograph measures revealed significant differences between relevant questions answered truthfully and relevant questions answered deceptively. As predicted, subjects showed greater increases in skin conductance and blood pressure, and showed greater reductions in finger pulse amplitude, heart rate, and respiratory activity when they lied than when they told the truth. With the exception of the cardiograph, the observed pattern of changes in the various response channels is consistent with the changes reported previously for comparison-question tests (e.g., Kircher & Raskin, 1988; Podlesny & Raskin, 1978; Raskin & Hare, 1978; Rovner et al., 1979).

The results from the cardiograph were disappointing. Although deception was associated with statistically significant increases in systolic cardiograph levels, the magnitude of the effect was small, and no significant effect of deception was observed for increases in diastolic cardiograph levels. Since all other response channels showed effects consistent with those observed in other investigations of polygraph techniques, it is unlikely that the cardiograph is generally not useful for RI tests.

There are at least two other possible reasons for the failure to observe large effects for the cardiograph. The pressure in the cuff is typically set to about 60 mm Hg. For subject comfort in the present study, the cuff pressure was set to about 50 mm Hg. Therefore, the coupling between the arm and

the cuff may have been insufficient to adequately track changes in blood pressure and volume. However, the quality of the recordings at this low pressure did not appear to differ from that observed at higher pressures.

Another possibility is that the psychological impact of the manipulation was weak. Large effects on physiological measures are common in mock crime experiments where subjects engage in a reasonable facsimile of a real crime (Kircher, Horowitz, & Raskin, 1988). Subjects in these experiments are recruited from the general community and are placed in an unfamiliar environment. They receive tape-recorded instructions and meet with no one prior to committing the mock crime. They are lead to believe that they might 'get caught' and should prepare an alibi. They are instructed to tell no one that they are subjects in an experiment, and they commit the crime in an uncontrolled naturalistic setting in broad daylight. Under these conditions, subjects experience considerable emotional involvement during the commission of the crime and subsequent polygraph examinations.

Smaller or nonsignificant effects on physiological measures are often reported in studies that include weak manipulations or unrealistic crime scenarios (Kircher et al., 1988). In the present study, subjects were recruited from the campus and were familiar with their surroundings. Because most subjects were college students who were likely to have taken basic psychology and related courses, many may have been generally acquainted with the goals and methods of psychological experimentation. Prior to taking their polygraph examinations, subjects met with a research assistant who explained the task, directed them to falsify items on the simulated employment application form, and helped them complete the form. Subjects were informed that the polygraph examiner was unaware of the items that had been falsified on the application form. However, because the subject and research assistant worked together to complete the form, it was clear to the subject that the research assistant knew which items had been falsified. Therefore, the subject's deception was only partial and the situation was clearly contrived. All of these factors may have reduced subjects' level of

personal involvement in the task and established a psychological context that was insufficient to evoke a strong cardiovascular response.

New Cardiovascular Measures

In contrast to the weak results from the cardiograph, measures of diastolic and systolic blood pressure from the Finapres blood pressure monitor discriminated between truthful and deceptive responses. Stronger effects were obtained on systolic than diastolic blood pressure. Together, these findings suggest that the Finapres is more diagnostic than the cardiograph in RI tests. Since the Finapres device is also less uncomfortable than the arm cuff of the cardiograph, consideration should be given to replacing the cardiograph with the Finapres or similar device.

Results from a recent study of the probable-lie comparison-question test also indicated that the data from the Finapres were more diagnostic than those from the cardiograph (Podlesny & Kircher, in preparation). Since that study was specifically designed to compare the cardiograph and Finapres, their data provide much more conclusive support for the use of the Finapres.

On the other hand, the cardiograph is measured from a standard blood pressure cuff and an inexpensive pressure transducer. Together, the cuff and transducer cost about \$25. In contrast, the Finapres is a costly piece of medical equipment. It is no longer even manufactured. Although the company has replaced the Finapres with a new device known as the Portapres, the current price of the Portapres is about \$27,000. It is likely that a comparable recording technology could be developed and manufactured at a lower cost, but the cost might still be significant. An important question at this point is if measures extracted from other physiological channels can be identified that will compensate for the loss in predictive validity from not using the Finapres or a comparable device. In the present study, neither the cardiograph nor the Finapres contributed measures to a prediction equation that was optimal for discriminating between truthful and deceptive responses.

Effects of deception were also obtained for heart rate (HR). The pronounced HR deceleration associated with deception in the present study is consistent with prior research with comparison-question tests (Podlesny & Raskin, 1978; Rovner et al., 1979; Raskin, 1979). Raskin (1979) attributed the deceleration to a baroreceptor reflex initiated by a rapid rise in blood pressure at the aortic arch. However, in the present study, HR deceleration began at about 5 seconds after question onset. Neither systolic nor diastolic blood pressure increases in the finger distinguished between truthfulness and deception at the fifth poststimulus second. This suggests that the HR deceleration was not a consequence of increases in peripheral blood pressure, although it might be tied to an increase in blood pressure in large arteries such as the aortic arch, which was not recorded.

Cardiac deceleration is a component of the orienting reflex and may be viewed as an index of the signal value of the initiating stimulus (Graham & Clifton, 1966). Theoretically, relevant questions answered deceptively should have greater signal value than relevant questions answered truthfully or neutral questions (Raskin, 1979). This hypothesis is entirely consistent with the heart rate responses observed in the present study, since only relevant questions answered deceptively were associated with cardiac deceleration.

Effects of deception on vagal tone were obtained but they were not in the expected direction. Prior research revealed a small but significant effect on vagal tone in the five-second interval that began 15 seconds after question onset (Raskin & Kircher, 1990). In that study, subjects were guilty or innocent of falsifying answers on an employment application and were given comparison-question tests. Vagal tone for guilty subjects dropped in response to relevant questions, whereas vagal tone for innocent subjects dropped in response to comparison questions. In the present study, vagal tone increased when subjects were deceptive and varied about baseline levels when subjects were truthful. The maximum increase was observed 5-9 seconds after question onset.

Grossman, Wuentjies, and Defares (1984) and Hirsch and Bishop (1981) reported that vagal tone is positively related to the amplitude of respirations. In comparison-question tests, respiration is suppressed for about 10 seconds following question onset. After 10 seconds, there is a compensatory increase in the depth of respirations. If vagal tone covaries with respiration, one would expect an initial reduction in vagal tone followed by a compensatory increase. Therefore, the results reported earlier by Raskin and Kircher (1990) are not consistent with this predicted pattern.

The present results are also not consistent with the prediction that vagal tone is positively related to respiratory activity during polygraph tests. In the present study, vagal tone was elevated during the 15-second interval that breathing was suppressed. The effects on vagal tone were, however, consistent with the observed reductions in heart rate. The onset of increases in vagal tone corresponded to the onset of reductions in heart rate, and the changes in both channels peaked at about the same time. Moreover, McCabe, Yongue, Porges, and Ackles (1984) demonstrated that stimulation of the baroreceptors in the aortic arch of rabbits and cats increases vagal inhibitory action on the heart. Thus, as Raskin (1979) originally suggested, an increase in blood pressure might stimulate the baroreceptors in the aorta. This results in a reflexive increase in vagal tone that causes the heart to slow.

The hypothesized chain of events suggests that the vagal effects observed in polygraph examinations are mediated by changes in blood pressure and not by respiratory activity. The pattern of correlations among features extracted from respiration, blood pressure, heart period, and vagal tone waveforms in the present study provide empirical support for this hypothesis. Vagal tone was positively correlated with increases in systolic blood pressure in the finger ($r = .24$) and increases in heart period (reductions in heart rate; $r = .35$) but not thoracic ($r = -0.06$) or abdominal respiratory activity ($r = -0.09$).

Exploratory Analyses

Exploratory analyses revealed numerous differences between physiological reactions to relevant questions answered truthfully and relevant questions answered deceptively. All channels except the cardiograph yielded features that correlated significantly with a dichotomous variable that distinguished between relevant questions answered truthfully and deceptively. Overall, features extracted from the skin conductance waveform tended to be more diagnostic than those extracted from other waveforms. Measures of respiration suppression were also highly diagnostic. With the exception of the null results from the cardiograph, the observed effects are generally consistent with those obtained in mock crime studies of comparison-question tests. Honts, Devitt, Winbush, and Kircher (1996) reached similar conclusions in a study of the concealed information (guilty-knowledge) test.

Within-subject point-biserial correlations were computed in the present study to assess the diagnostic validity of various measures of physiological reactivity. These correlations were generally smaller than point-biserial validity correlations reported elsewhere (e.g., Kircher & Raskin, 1988). However, there were substantial differences between this and earlier studies in the manner in which these correlations were calculated. In earlier studies, a measure of the mean difference between comparison and relevant questions was computed for each guilty and innocent subject in the sample. The mean difference was then correlated with a dichotomous variable that represented group membership. Thus, only one correlation was obtained to assess the discrimination between groups of guilty and innocent subjects.

In the present study, all subjects were guilty of falsifying answers on their simulated employment applications. Since there was no true comparison group of innocent subjects, a within-subject point-biserial correlation was obtained for each guilty subject. For each subject, the within-subject correlation assessed the discrimination between questions answered truthfully and deceptively. The mean of within-subject correlations was then computed. Although the correlations reported

in the present study are related to those reported previously, it is difficult to compare their absolute magnitudes directly.

There is reason to believe stronger evidence of predictive validity for individual features would have been obtained with traditional between-group point-biserial correlation coefficients. To develop a new model of physiological measures for RI tests, a pseudo sample of truthful cases was created by excluding the data for the relevant question answered deceptively from half of the polygraph tests administered to subjects. Since this procedure provided samples of truthful and deceptive cases, it was then possible to compute the traditional between-group point-biserial correlation for each feature. These between-group correlations were uniformly greater in absolute magnitude than their within-subject counterparts. For example, the mean of the 192 within-subject correlations for SC amplitude and TR excursion were 0.44 and -0.35, respectively. The comparable between-group correlations were 0.53 and -0.46. The latter correlations are within the range of between-group correlations reported for comparison-question tests (Horowitz, Kircher, Honts, & Raskin, 1997; Kircher & Raskin, 1988; Raskin et al., 1988).

New Model of Differential Reactivity for RI Tests

A new model of physiological measures was developed for RI tests. Overall, as compared to the original 3-variable model, the new model increased the correct decisions by about 10% and reduced false positive outcomes by about 10%. The differences between the original and new models depended on the criterion used to decide if the subject was deceptive to a particular relevant question. With the criterion for the new model set to obtain a false positive rate that was less than 10% ($t = 0.80$), the new 5-variable model produced 80.2% correct, 8.3% false positives, and 11.5% false negatives on cross-validation. At the same cutoff, the original 3-variable model yielded 65.1% correct, 18.7% false positives, and 16.2% false negatives.

Although the new model substantially improved decision accuracy, estimates of the

model's predictive validity were positively biased. In general, inflated estimates of predictive validity occur because *variables* are selected to optimize the accuracy of diagnoses in the sample, and because *weights for those variables* are computed to optimize the accuracy of diagnoses in the sample. Although the variables and weights are ideal for the sample at hand, the sample at hand is not perfectly representative of the population. Therefore, one would not expect the accuracy to be as high in another independent sample from the same population.

In the present study, any bias due to the choice of weights for the variables was controlled. The weights for variables were derived using only half of the cases (standardization sample). The accuracy of diagnoses in the remaining cases (validation sample) provided an independent test of the validity of model weights for RI tests. However, all of the cases were used to select a subset of variables for the model from a pool of 18 variables, and that group of 18 variables had been pre-selected from a much larger pool of potential predictor variables. Since we did not control for bias due to the selection of variables, the reported accuracy of the new model may be too high. An independent and representative sample of new sample of cases would be needed to obtain an unbiased estimate of the validity of the 5-variable model for field use.

The new model contained measures of skin conductance responses, respiration suppression, and vasoconstriction in the finger. The skin conductance measures included peak amplitude, half-recovery time, and latency to response onset. Skin conductance amplitude and half-recovery time are commonly used in psychophysiological research as indicators of electrodermal response magnitude. It is therefore not too surprising that they discriminated between truthful and deceptive responses.

It is interesting that skin conductance response latency entered the model. In the present study, all relevant questions began with the phrase, "Did you falsify your...." In the audio-only condition, subjects had to wait until the polygraph examiner concluded the relevant question before they knew whether to

lie or tell the truth. Random variability in the time it took the polygraph examiner to state the question would contribute error variance to the response latency measure and reduce its predictive validity. Despite this source of random error in half of the polygraph tests, response latency significantly discriminated between truthful and deceptive cases.

In the audiovisual condition, relevant questions appeared on the computer monitor the moment the polygraph examiner began to state the question. Since the critical information was always visually available to the subject at question onset, the rate at which the polygraph examiner stated the question was not likely to be a significant source of error variance. Thus, one might expect that response latency would be more diagnostic in the audiovisual condition than in the audio-only condition. A post hoc comparison of response latencies in the audiovisual and audio-only conditions provided marginal support for this prediction ($p < .04$).

Finally, although measures derived from heart period and absolute blood pressure were not included in the selected model, they frequently appeared in alternate models listed by the all-possible-subsets regression algorithm. These promising new measures should be examined with more compelling laboratory paradigms that achieve levels of personal involvement that more closely approximate those in the field.

Spontaneous Countermeasures

Interviews and questionnaires revealed several differences in the types of countermeasures used by subjects during audio only and audiovisual tests. Two such differences were consistent across interview and computer questionnaire formats. However, the effects were generally small, and the number of significant effects was small relative to the number of statistical tests. Moreover, no clear pattern of strategies emerged that would suggest that the visual presentation of test questions fundamentally changed the way subjects attempted to defeat the test.

It is interesting to note the divergence of reports by subjects during the interview and on the computer questionnaire. A wide range of countermeasure techniques was revealed by both formats. However, subjects were generally more willing to endorse items on the computer questionnaire than to volunteer the same information during the interview. Subjects may have had difficulty recalling their thoughts and actions during the polygraph examinations, or they may have been reluctant to discuss these matters with the research assistant.

A wide range of mental and physical techniques was used by subjects to appear truthful during their polygraph tests. Most subjects reported that they used the same strategy throughout the test. However, a large percentage of subjects reported that they used countermeasures only when they answered a question truthfully, or only when they answered a question deceptively. None of the countermeasures, not even those used selectively in response to specific questions, appeared to affect the risk of false negative or false positive outcomes.

The null results from the present study are consistent with those obtained in studies using comparison question polygraph techniques (Honts, Hodes, & Raskin, 1985; Rovner et al., 1979). It appears that the RI technique is no more sensitive than

comparison question techniques to spontaneous countermeasure maneuvers by well educated but naïve deceptive subjects. However, an investigation similar to the study by Rovner et al. (1979) should be conducted with RI tests. Specifically, it is unknown if the accuracy of RI tests would suffer if subjects understand the rationale for making decisions, are given suggestions about countermeasures that are likely to be effective and when they should be applied, and are given opportunities to practice their countermeasure maneuvers while attached to a polygraph.

In conclusion, the present study failed to show effects of audiovisual presentations of test questions on decision accuracy. On the other hand, it revealed numerous effects of deception on traditional and new physiological measures. A promising avenue for future research would be the development of computer algorithms for conducting RI tests. The computer may be programmed to extract diagnostic information from these physiological measures and make statistical comparisons of reactions to the various test questions. By analyzing subjects' physiological reactions online, the computer could also automatically adjust the sequencing of test questions by the polygraph examiner. These innovations would improve the reliability of RI tests by formalizing the rules for conducting RI tests and interpreting the physiological data.

References

- Barland, G. H., Honts, C. R. & Barger, S. D. (1989). Studies of the accuracy of security screening polygraph examinations. Research Division Department of Defense Polygraph Institute, Fort McClellan, AL.
- Biejk, J. (1980). Experimental and procedural influences on differential electrodermal activity. *Psychophysiology*, 17, 274-278.
- BMDP Statistical Software Manual (1992). W. J. Dixon (Ed.) Berkeley: University of California Press.
- Carlson, B. L. & Smith, B. J. (1991). The effects of aural versus visual presentation of questions during a detection of deception task. Department of Defense Polygraph Institute, Fort McClellan, AL.
- Cohen, J. & Cohen, P. (1983). *Applied multiple regression: Correlation analysis for the behavioral*

- sciences (2nd Ed.)*. New York: Erlbaum.
- Graham, F. K. & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin*, 65, 305-320.
- Grossman, P., Wientjes, C., & Defares, P. (1984). Individual differences in cardiac parasympathetic control predicted by ventilatory parameters. *Psychophysiology*, 21, 579. (Abstract).
- Hirsch, J. A. & Bishop, B. (1981). Respiratory sinus arrhythmia in humans: How breathing pattern modulates heart rate. *Am. J. Physiol.*, 241, H620-H629.
- Honts, C. R., Devitt, M., Winbush, M. & Kircher, J. C. (1996). Mental and physical countermeasures reduce the accuracy of the concealed knowledge test. *Psychophysiology*, 33, 84-92.
- Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985). Effects of physical countermeasures on the physiological detection of deception. *Journal of Applied Psychology*, 70, 177-187.
- Honts, C. R., Raskin, D. C., Kircher, J. C., & Hodes, R. L. (1988). Effects of spontaneous countermeasures on the physiological detection of deception. *Journal of Police Science and Administration*, 16, 91-94.
- Honts, C. R., Raskin, D. C. & Kircher, J. C. (1994). Effects of physical and mental countermeasures on the detection of deception. *Journal of Applied Psychology*, 79, 252-259.
- Horowitz, S. W., Kircher, J. C., Honts, C. R., & Raskin, D. C. (1997). The role of control questions in the physiological detection of deception. *Psychophysiology*, 34, 108-115.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook (3rd Ed.)*. Englewood Cliffs, N. J.: Prentice Hall.
- Kircher, J. C., Horowitz, S. W., & Raskin, D. C. (1988). Meta-analysis of mock crime laboratory studies of the control question polygraph technique. *Law and human behavior*, 12, 79-80.
- Kircher, J. C. & Raskin, D. C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology*, 73, 291-302.
- Kircher, J. C., Raskin, D. C., Honts, C. R., & Horowitz, S. W. (1994). Generalizability of statistical classifiers for the detection of deception. *Psychophysiology*, 31, S73. (Abstract)
- McCabe, P. M. Yongue, B. G. Porges, S. W. & Ackles, P. K. (1984). Changes in heart period, heart period variability and a spectral analysis estimate of respiratory sinus arrhythmia during aortic nerve stimulation in rabbits. *Psychophysiology*, 21, 149-158.
- McCabe, P. M. Yongue, B. G., Ackles, P. K., & Porges, S. W. (1985). Changes in heart period, heart period variability and a spectral analysis estimate of respiratory sinus arrhythmia in response to pharmacological manipulations of the baroreceptor reflex in cats. *Psychophysiology*.
- Numaguchi, G., Kircher, J.C., Craig, R., Raskin, D. C., Woltz, D. J., & Packard, R. E. (1994). *The Erlanger method for measuring cardiovascular activity: Correlations with blood volume and arterial pressure. Psychophysiology*, 31, S11. (Abstract)
- Office of Technology Assessment [OTA] (1983). *Scientific validity of polygraph testing: A research review and evaluation -- A technical memorandum (OTA-TM-H-15)*. Washington, D. C.: U.S.

Government Printing Office.

- Podlesny, J. A. & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology*, 15, 344-359.
- Porges, S. W., Bohrer, R. E., Cheung, M. N., Drasgow, F., McCabe, P. M., & Keren, G. (1980). New time series statistic for detecting rhythmic co-occurrence in the frequency domain: The weighted coherence and its application to psychophysiological research. *Psychological Bulletin*, 88, 580-587.
- Raskin, D. C. (1979). Orienting and defensive reflexes in the detection of deception. In H. D. Kimmel, E. H. van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans*. Hillsdale, N. J.: Erlbaum.
- Raskin, D. C. & Hare, R. D. (1978). Psychopathy and the detection of deception in a prison population. *Psychophysiology*, 15, 126-136.
- Raskin, D. C. & Kircher, J. C. (1990). Development of a computerized polygraph system and physiological measures for detection of deception and countermeasures: A pilot study. Report to the U.S. Federal Government (Contract 88-L655300-000). Scientific Assessment Technologies, 1865 Herbert Avenue, Salt Lake City, UT 84108
- Raskin, D. C., Kircher, J. C., Honts, C. R., & Horowitz, S. W. (1988). A study of the validity of polygraph examinations in criminal investigation. Final report to the National Institute of Justice (Grant No. 85-IJ-CX-0040). Salt Lake City: University of Utah, Department of Psychology.
- Rovner, L. I., Raskin, G. C., & Kircher, J. C. (1979). Effects of information and practice on detection of deception. *Psychophysiology*, 16, 197-198. (Abstract)

Appendix A

Features Extracted from Response Waveforms by CPSLAB

General notes: CPSLAB permits specification of scoring windows relative to question onset. Unless otherwise noted in the text, the scoring window for each feature began at question onset and ended 20 seconds later. Measured times were reported in ms but are accurate only to the precision of the storage rate specified in the project file for a channel.

1. *Peak amplitude.* CPSLAB identified the time and level each change from negative or zero slope in the response waveform (low point) and each change from positive slope to zero or negative slope (high point). The difference between each low point and every succeeding high point was computed. Peak amplitude was defined as the greatest such difference if it exceeded some preset minimum. For skin conductance, this minimum was 0.02 μ Siemens. For all other waveforms, the minimum was zero.
2. *Area to half recovery* was area under the response curve from response onset to the time at which the tracing recovered to half the peak amplitude or to the end of the scoring window, whichever occurred first (time of half recovery).
3. *Area to full recovery* was area under the response curve from response onset to the time at which the tracing recovered to the level at response onset or to the end of the scoring window, whichever occurred first (time of full recovery).
4. *Area to peak amplitude* was area under the response curve from response onset to the time of peak amplitude.
5. *Area from peak amplitude to full recovery* was area under the response curve from the time of peak amplitude to the time of full recovery.
6. *Risetime from response onset* was the time in ms from response onset to the time of peak amplitude.
7. *Risetime from first low point* was the time in ms from the first detected point of change from zero or negative slope to positive slope to the point of peak amplitude.
8. *Half recovery time* was the difference in ms between the time of peak amplitude and the time of half recovery.
9. *Full recovery time* was the difference in ms between the time of peak amplitude and the time of full recovery.
10. *Duration to half recovery* was the difference in ms between the time of response onset and the time of half recovery.
11. *Duration to full recovery* was the difference in ms between the time of response onset and the time of full recovery.
12. *Latency to first low point* was the time in ms from the onset of the stimulus event to the first detected change from zero or negative slope to positive slope.

13. *Latency to response onset* was the time in ms from the onset of the stimulus event to response onset.
14. *Riserate from stimulus onset* was the linear rate of increase from the onset of the stimulus event to peak amplitude.
15. *Riserate from response onset* was the linear rate of increase from response onset to peak amplitude.
16. *Half recovery rate* was the linear rate of decrease from peak amplitude to the point of half recovery.
17. *Full recovery rate* was the linear rate of decrease from peak amplitude to the point of full recovery.
18. *Excursion* was the sum of absolute deviations between adjacent samples in the scoring window.
19. *Number of responses* was the number of changes from zero or negative slope to positive slope in the scoring window.
20. *Burst frequency* was the reciprocal of the shortest interval between each low point in the waveform and the second low point that followed.
21. *Mean of successive amplitudes* was the mean of successive differences between each low point in the scoring window and the following high point.
22. *Standard deviation of successive amplitudes* is the standard deviation of successive differences between each low point in the scoring window and the following high point.
23. *Level* was the mean of the samples that define the waveform segment.
24. *Standard deviation* was the standard deviation of samples that define the waveform segment.

Comparison of Evidentiary and Investigative Decision Rules: A Replication¹

Donald J. Krapohl and Barry Cushman

Abstract

The present investigation sought to replicate Krapohl (2005) in a comparison of traditional polygraph decision rules to those he proposed for use in courtroom and Paired Testing (Marin) Protocol application (American Society for Testing and Materials, 2005). Using a new sample of field cases, 10 experienced polygraph examiners conducted blind scoring of 100 polygraph examinations for which ground truth had been established. On average, Evidentiary Decision Rules and Investigative Decision Rules showed no difference in overall error rates (inconclusives excluded). However, there was a significant reduction in inconclusive decisions and an increase in correct decisions for the Evidentiary Decision Rules compared to the Investigative Decision Rules. Most of this improvement came from the correct detection of truthful cases, and there was a modest loss in detecting deceptive cases. Evidentiary Decision Rules also showed a more balanced accuracy for truthful and deceptive cases than did Investigative Decision Rules. The replication further supports the use of Evidentiary Decision Rules in settings where balanced accuracy and a low number of inconclusive results are preferred over decision rules designed only to minimize false negative decisions.

Krapohl (2005) outlined a rationale for the development of polygraph decision rules, or cutoff scores, that were suitable for evidentiary and paired-testing as specified by the American Society for Testing and Materials (ASTM, 2005). It was based on observations from several research publications that the cutoffs used by many investigative organizations (e.g., Light, 1999) were significantly less sensitive to truthfulness than deceptiveness (Blackwell, 1998; Franz, 1989; Podlesny & Truslow, 1993; Raskin, Kircher, Honts & Horowitz, 1988; Senter, 2003). These cutoffs, referred to hereafter as Investigative Decision Rules, tend to minimize false negative outcomes and may be quite sensible when used in settings where the costs of false negatives are high and those of false positives are low. They become more difficult to justify

in settings where the costs of errors shift, however. Krapohl devised a new set of decision rules that considered the factors that influenced decision errors. The new rules, called Evidentiary Decision Rules, were tested on four independent data sets, and they produced a more balanced accuracy than when the Investigative Decision Rules were applied to the same scores. Evidentiary Decision Rules also permitted most scorers to meet the stringent Marin Protocol standard of limiting inconclusive decisions to 20% or lower, and achieving an accuracy of 86% or greater with the remaining cases. Most scorers did not meet these requirements with the Investigative Decision Rules. These findings were promising, and because they were based on four different data sets, it suggests that they were robust.

¹This study would not have been possible without the participation of the 10 blind scorers in this study: Gerard Brady, Barry Colvert, Barry Cushman, Donnie Dutton, Robert Gilford, Ronald Homer, Danny Morgan, Robert Mylott, Stanley Slowick, and Michael Woodcock. We are also grateful to Rose Swinford for managing the data, to Dr. Stuart Senter for originally proposing a two-stage decision process, and to Dr. Tim Weber for his helpful comments to an earlier draft of this article. This is one in a series of papers under the heading *Best Practices*. The opinions expressed in this article are those of the authors, and do not necessarily represent those of the US Government, the Department of Defense or the Portland (Maine) Police Department. Comments or reprint requests should be sent to: Donald Krapohl, PO Box 10411, Ft. Jackson, SC 29207, or by e-mail to dkrapohl@aol.com.

The development of the Evidentiary Decision Rules was initially prompted by the introduction of the Marin (or paired-testing) Protocol to the field, a method that relies on a statistical principle called “joint probabilities” to minimize polygraph decision errors. Joint probabilities relate to the likelihood of two events occurring together, and paired-testing entails the conduct of polygraph examinations of two or more examinees. Paired-testing is used when different individuals are offering contradictory versions of an event, and one of them must certainly be lying. Using the principle of joint probabilities, when two individuals are making a claim for which one of them must be deceptive, and both undergo polygraph testing regarding that claim, the likelihood that the polygraph results of *both* examinees are erroneous is very small. Specifically, the probability of both of the examinations being in error is equal to the product of the two independent error probabilities, represented mathematically as $P(A \text{ and } B) = P(A) * P(B)$, or $P[A \cap B]$. For example, if the likelihood of a polygraph decision error is 14% (median error estimated by the National Research Council, 2003), the likelihood of two independent decision errors is 14% of 14%, or about 2%. The Marin Protocol exploits this statistical principle, and when applied under the narrow conditions described here, it permits a decision accuracy that is not possible in a single polygraph examination.

The development of the Evidentiary Decision Rules considered four important, proven factors. The first is that, on average, truthful examinees produce reactions to probable-lie comparison questions that are relatively less intense than the reactions of deceptive examinees to relevant questions. This asymmetry in responding means the scores of deceptive examinees will tend to be further below zero than the corresponding scores of truthful examinees will be above zero. Said another way, the underlying phenomenon is asymmetrical. If symmetrical cutoffs are placed over an asymmetrical phenomenon, the detection of one group (liars) will be better than the detection of the other group (truthtellers). If a user wants to have equivalent accuracy for both truthtellers and liars, cutoffs must be adjusted to account for this asymmetry. The cutoffs for the

Evidentiary Decision Rules were calculated from the asymmetry uncovered in the development of the Objective Scoring System (Krapohl & McManus, 1999), and subsequently cross validated (Krapohl, 2005).

A second factor affecting polygraph decisions came from the findings of Cullen and Bradley (2004) that the positional relationship of relevant and probable-lie questions could influence scores. In their analog study they found that the sequence of irrelevant-relevant-comparison (I-R-C) questions resulted in polygraph scores that were significantly more negative than the when the sequence of questions was comparison-relevant-irrelevant (C-R-I). This phenomenon was also found in the scores of field cases (Krapohl & Dutton, 2005). Average scores for truthful examinees with the I-R-C sequence had scores that were near or below zero in both studies. To avoid this problem, the Evidentiary Decision Rules require that the testing technique be a form of Zone Comparison Technique (ZCT). In the ZCT family of formats the relevant questions are always preceded by comparison questions, and therefore mitigates the negative shift in scores associated with the I-R-C sequence.

It has also been discovered that two-stage decision rules help to minimize inconclusive outcomes, the third important factor in the design of the Evidentiary Decision Rules. In 2003 Dr. Senter discovered that a scorer could significantly increase the proportion of definitive polygraph decisions by using a two-stage decision process. One version of Dr. Senter’s approach is used in the Evidentiary Decision Rules, which is discussed in the Methods section. The use of the two-stage decision rules in the Krapohl (2005) introductory paper on Evidentiary Decision Rules cut inconclusive decisions by more than half without any significant effect on overall decision accuracy, an effect predicted from Senter’s earlier work.

The final factor that can affect decision accuracy is the scoping of the relevant questions. It is generally recognized that decision accuracy is degraded when multiple issues are presented in the same test, or as relevant questions become more ambiguous. Screening tests are less accurate than single-issue examinations for these reasons.

Screening tests have more utility in the investigative process, to assist in the pursuit of information, but they cannot compete for accuracy with single-issue testing. If accuracy is of paramount importance in evidentiary polygraphy, a position with strong appeal, it is necessary that evidentiary polygraph examinations be single-issue examinations. The Evidentiary Decision Rules were based on single-issue examinations.

The purpose of the present effort is to replicate Krapohl's (2005) study with an independent data set. Scorers who were processing for certification under the Marin Protocol conducted blind scoring of 100 new cases drawn from a database of confirmed cases. Unlike many blind scoring studies where there is little incentive for examiners to perform at their best, the present project took advantage of their extrinsic motivation to obtain a prestigious certification of competence.

Method

Cases

A total of 100 cases were drawn at random from the same large database of confirmed cases used by Krapohl and McManus (1999). Half of these field cases were from deceptive examinees and the other half were from truthful examinees. Ground truth was established by a confession that inculpated the examinee or exculpated another examinee, by forensic evidence, or by the subsequent discovery that the reported crime did not occur. Weaker confirmatory criteria, such as judicial decisions or eyewitness testimony, were not used. All examinations had been conducted according to the federal method of the Zone Comparison Technique (ZCT) (Light, 1999) on Axciton computerized polygraphs.

Scorers

One of the requirements in the Marin Protocol certification process is to demonstrate competence in the analysis of polygraph charts (Veritas Center, 2005). Scores were taken from the work of the first 10 field examiners undergoing the Marin Protocol certification process. Decisions were imposed on them first using the Investigative Decision Rules, then the Evidentiary Decision Rules. The scorers

also made separate judgments of truthfulness or deception with the cases based upon the decision rules they preferred, but in the interest of standardization, only the two sets of decision rules being compared are reported here.

Procedure

Examiners who volunteered to submit to the Marin Protocol certification process were mailed a data disk that contained instructions, a score sheet, and 100 cases. The instructions advised the examiners to print 100 copies of the score sheet and to use them during their scoring (See Appendix A). The instructions also included a mailing address to return the score sheets and an e-mail address to send questions or comments about the project. The scorers were asked to use the standard 7-position numerical scoring scale.

The 100 cases (300 charts) had been converted to PDF format. The choice in digital format was based on two considerations: First, not all of the participants were likely to have Axciton software. Therefore, it was necessary to choose a generally available format. Second, the use of the PDF version over the Axciton version provided a safeguard that precluded the use of any of the automated algorithms to guide scoring. Because of the limitations of the PDF version, the examiners were required to score the charts from their computer screens as the printed versions were too small to be useful.

Once the scorers mailed in their score sheets, the scores were entered into an Excel spreadsheet. Formulas built into the spreadsheet automatically tested the two sets of decision rules for the scores and calculated the decision accuracy and inconclusive rates.

Decision rules

The processing of the data was identical to that used by Krapohl (2005) and is briefly repeated here. Two types of decision rules were applied to the scores. The first were the Investigative Decision Rules used for polygraph examinations by the U.S. Government (Blackwell, 1998; Light, 1999). According to these rules, an evaluation of DI is made when the total of all scores is -6 or lower, or when the sum of any one of the relevant questions is -3 or lower ("Spot Score Rule," or "SSR"). Decisions of NDI require a

positive total (spot) score for each relevant question and a grand total of +6 or greater for the sum of all spot scores. All other cases result in Inconclusive decisions.

Evidentiary Decision Rules entail a two-stage process as originally suggested by Senter (2001; 2003). They begin with asymmetric cutting scores: if the grand sum of scores equals -6 or lower, the call is DI; if the grand sum of all scores is +4 or greater, the determination is NDI. In those cases where the grand sums range from -5 to +3, the SSR is applied. For those cases, if a single relevant question has a sum of -3 or lower, the decision is DI. All other cases are deemed Inconclusive.

Data reduction

Zeros were assigned where the examiner made notations that the tracing was uninterpretable due to artifacts. Only spot scores and total scores were analyzed. Decisions for each case were based solely on the scores and the two sets of decision rules. Because of multiple comparisons, alpha for all statistics was set at .01.

Results

Investigative Decision Rules

Accuracy for the Investigative Decision Rules is shown in Table 1. Collectively, the overall correct decision rate was 86.1% without inconclusives, which met the ASTM accuracy standard for the Marin Protocol of 86.0% or greater. The inconclusive rate was 19.8%, which was also within the Protocol limit of 20%. However, only three of the ten individual scorers satisfied these requirements: scorers 1, 2, and 9.

Tests of proportion were applied to the polygraph decisions (Bruning & Kintz, 1997). Of the 500 results possible on deceptive cases (10 scorers of 50 deceptive cases), 412, or 82.4% of them were correct, and this was greater than chance ($z = 10.82, p < .01$). A total of 278 correct decisions out of 500 opportunities were made on truthful cases for an accuracy of 55.6%, and this did not exceed chance expectancy ($z = 1.77, p > .01$). Accuracy was higher for deceptive cases than truthful cases when inconclusives were removed ($z = 7.89, p < .01$), and when inconclusives were included ($z = 9.12, p < .01$).

The proportion of inconclusive results was greater for truthful cases than for deceptive cases ($z = 5.24, p < .01$). Error rates

Table 1. Decision accuracy using Investigative Decision Rules, in percentages.

Scorer	<u>Deceptive Cases</u>			<u>Truthful Cases</u>			<u>Overall</u>	
	Correct	Miss	Inc	Correct	Miss	Inc	Correct w/o Inc	Inc Rate
1	90.0	4.0	6.0	70.0	18.0	12.0	87.9	9.0
2	86.0	2.0	12.0	70.0	4.0	26.0	96.3	19.0
3	82.0	6.0	12.0	50.0	24.0	26.0	81.5	19.0
4	74.0	6.0	20.0	62.0	16.0	22.0	86.8	21.0
5	82.0	2.0	16.0	34.0	20.0	46.0	84.6	31.0
6	80.0	4.0	16.0	48.0	22.0	30.0	83.1	23.0
7	80.0	8.0	12.0	44.0	28.0	28.0	77.5	20.0
8	82.0	0.0	18.0	58.0	14.0	28.0	90.9	23.0
9	86.0	8.0	6.0	72.0	16.0	12.0	86.8	9.0
10	82.0	4.0	14.0	48.0	18.0	34.0	85.5	24.0
Average	82.4	4.4	13.2	55.6	18.0	26.4	86.1	19.8

Inc = Inconclusive

were significantly lower for deceptive cases than truthful cases ($z = 6.82, p < .01$). As was found in the original 2005 study, Investigative Decision Rules appear to make the examination sensitive to detecting deception, but there is a shortfall in its performance with truthful examinees: truthful examinees appear to incur significantly more errors and inconclusives.

Evidentiary Decision Rules

Table 2 shows the accuracy of the scorers when the Evidentiary Decision Rules are applied to the same scores. Of the 500 results possible on deceptive cases, 396, or 79.2% of them were correct, and this was greater than chance ($z = 9.65, p < .01$). A total of 412 correct decisions out of 500 opportunities were made on truthful cases for an accuracy of 82.4%, and this was also greater than chance expectancy ($z = 10.83, p < .01$). Correct decisions, as a percentage of all decisions, was 87.2%. The inconclusive rate was 7.3%. Collectively, both the accuracy and inconclusive rates met the Marin Protocol standard, as did seven of the ten individual scorers: scorers 1, 2, 4, 5, 6, 8, and 9

There were no statistically significant differences between the rates of correct decisions for truthful and deceptive cases when the Evidentiary Decision Rules were

used ($z = 1.28, ns$). Inconclusive rates were not significantly different between deceptive and truthful cases ($z = 1.58, ns$), nor were error rates ($z = 0.29, ns$). These findings replicate Krapohl's (2005) findings that the Evidentiary Decision Rules did not share the bias of the Investigative Decision Rules in correct decisions or inconclusives.

Comparison of Decision Rules

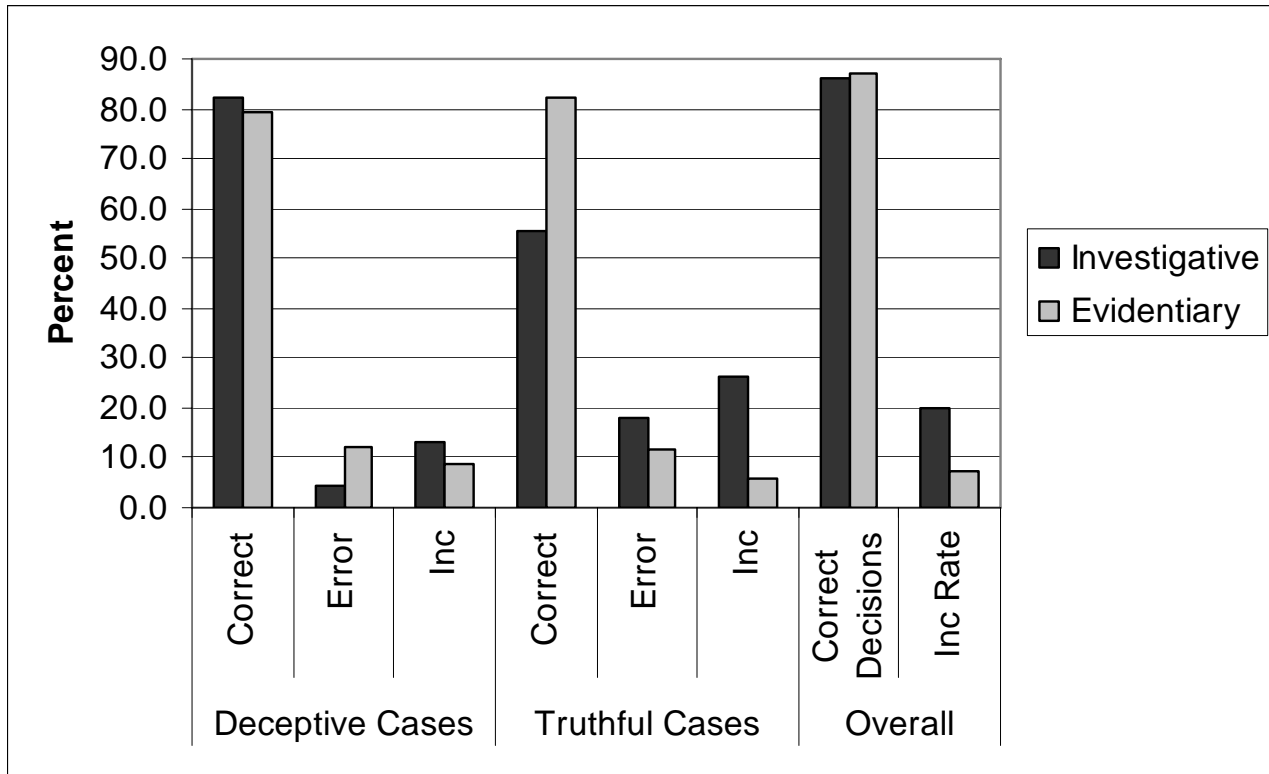
We next directly compared decision accuracy between the Investigative and Evidentiary Decision Rules. Scorers using Evidentiary Decision Rules made significantly more correct decisions overall than those using Investigative Decision Rules ($z = 6.09, p < .01$) with fewer inconclusives ($z = 8.17, p < .01$), and there were no differences in overall error rates ($z = 0.49, ns$). It appeared that most of the increase in correct decisions for Evidentiary Decision Rules over Investigative Decision Rules came from truthful cases ($z = 9.16, p < .01$) as had occurred in the Krapohl (2005) study. However, in contrast to previous results there were significant differences between the two sets of decision rules for deceptive cases ($z = 4.47, p < .01$): there were significantly more false negative errors with the Evidentiary Decision Rules than with the Investigative Decision Rules. Figure 1 compares accuracy between the decision rule

Table 2. Decision accuracy using Evidentiary Decision Rules, in percentages.

Scorer	Deceptive Cases			Truthful Cases			Overall	
	Correct	Miss	Inc	Correct	Miss	Inc	Correct w/o Inc	Inc Rate
1	86.0	12.0	2.0	92.0	8	0.0	89.9	1.0
2	80.0	10.0	10.0	96.0	0	4.0	94.6	7.0
3	76.0	14.0	10.0	76.0	16	8.0	83.5	9.0
4	72.0	12.0	16.0	84.0	6	10.0	89.7	13.0
5	80.0	8.0	12.0	74.0	16	10.0	86.5	11.0
6	76.0	10.0	14.0	80.0	14	6.0	86.7	10.0
7	78.0	18.0	4.0	82.0	14	4.0	83.3	4.0
8	82.0	10.0	8.0	74.0	14	12.0	86.7	10.0
9	84.0	10.0	6.0	82.0	14	4.0	87.4	5.0
10	78.0	18.0	4.0	84.0	14	2.0	83.5	3.0
Average	79.2	12.2	8.6	82.4	11.6	6.0	87.2	7.3

Inc = Inconclusive

Figure 1. Comparison of accuracy between Investigative and Evidentiary Decision Rules. Overall Correct Decisions excludes Inconclusives.



Discussion

Our results were consistent with those of Krapohl (2005). The Evidentiary Decision Rules brought about a significant reduction of inconclusives over the Investigative Decision Rules, with most of the improvement occurring with truthful cases. The use of Evidentiary Decision Rules also resulted in balanced accuracy, that is, the proportions of correct decisions for truthful and deceptive cases were equivalent. This contrasts markedly from the Investigative Decision Rules, which favor detection of deception over the detection of truthfulness.

Use of the Evidentiary Decision Rules resulted in a slightly higher but statistically significant increase in false negatives, a result that the earlier study failed to detect only by a very slim statistical margin. When considering both studies and looking at the pattern of correct and incorrect decisions in toto, it would be reasonable to conclude that the use of Evidentiary Decision Rules entails some risk

with deceptive cases. This observation serves as a caution against its use in settings where false negative results have a substantially higher cost than false positive results.

It is interesting to note that in this study the group results from Investigative Decision Rules would have met the Marin standards for accuracy and inconclusives, albeit by a small margin. The group averages for Investigative Decision Rules had an inconclusive rate of 19.8% (limit of 20% for Marin Protocol), and a decision accuracy of 86.1% (minimum of 86% under Marin). However, the imbalance in accuracy was still apparent: with the Investigative Decision Rules, 82.4% of the deceptive cases were correctly identified compared to only 55.6% of the truthful cases.

It is also interesting to note the significant differences between the Investigative and Evidentiary Scoring Rules in the correct classification of truthful cases on an individual examiner basis. A review of the

data reveals better results with the Evidentiary Scoring Rules. As previously stated, the group average of correctly identified truthful cases with the Investigative Scoring Rules was 55.6%. However, individually, five of the ten examiners identified truthful cases at rates of 50% or lower. When the Evidentiary Scoring Rules were applied to their scores, all examiners correctly identified truthful cases at 74% or better. Moreover, eight of the ten examiners had inconclusive rates (for truthful cases) that were greater than 20% with Investigative Scoring Rules. When Evidentiary Decision Rules were applied, nine of ten had inconclusive rates below 10%, and half of the examiners had inconclusive rates below 5%. The lowest inconclusive rate for any examiner using Investigative Scoring Rules was 12%, equal to the highest inconclusive rate when Evidentiary Scoring Rules were applied.

In the Krapohl (2005) report, examiners were cautioned against employing the Evidentiary Decision Rules in any setting until replication could be completed. With publication of these results, one can have greater confidence that the Evidentiary Decision Rules have the effect of lowering inconclusives and balancing decision accuracy. Users should also be mindful of the modest but statistically significant increase in false negatives associated with these decision rules. The advantages of the Evidentiary Decision Rules are restricted to the narrow conditions outlined earlier in this report, and like the prevailing Investigative Decision Rules,

examiners should not blithely use them for every examination.

It is also worth repeating the type of data on which the Evidentiary Decision Rules were validated: the single-issue federal ZCT format that uses probable-lie comparison questions and exactly three charts. They should not be used with multiple-issue or multiple-facet examinations, nor with directed-lie techniques until those methods have been subjected to equivalent study.

Summary

This replication of Krapohl (2005) once again found that the standard Investigative Decision Rules (+/-6 cutting scores and -3 spot scores) did not deliver the balanced accuracy and low inconclusive rate of the Evidentiary Decision Rules (two-stage process: -6 and +4 cutting scores followed by -3 spot score for those cases that would have otherwise been found inconclusive). Evidentiary Decision Rules for federal single-issue ZCT examinations appear to provide advantages for courtroom and paired-testing polygraph examinations, with no significant costs detected for those contexts. Most of the improvement was attributed to significantly better detection of truthfulness and the reduction of inconclusives, though there was a small but statistically significant increase in false negatives over the standard Investigative Decision Rules. Evidentiary Decision Rules may not be appropriate in investigative settings where there is a high cost for false negative decisions.

References

- ASTM. (2005). E2324-04 Standard Guide for PDD Paired Testing. ASTM International.
- Blackwell, N. J. (1998). PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examination from actual criminal investigations. *Polygraph*, 28(2) 149-175.
- Bruning, J. L., & Kintz, B. L. (1997). *Computational Handbook of Statistics* (4th Ed.). Addison Wesley Longman: Reading, MA.
- Cullen, M. C., & Bradley, M. T. (2004). Positions of truthfully answered controls on control question tests with the polygraph. *Canadian Journal of Behavioural Science*, 36(3), 167-176.

- Franz, M. L. (1989). *Technical report: Relative contributions of physiological recordings to detect deception*. Contract Number MDA 904-88-M-6612.
- Krapohl, D. J. (2005). Polygraph decision rules for evidentiary and paired-testing (Marin Protocol) applications *Polygraph*, 34(3) 184-192.
- Krapohl, D. J., & Dutton, D. W. (2005). Differences in response profiles for the Zone Comparison Technique and the Army Modified General Question Technique. *Polygraph*, 34(1), 1-9.
- Krapohl, D. J. & McManus, B. (1999). An objective method for manually scoring polygraph data. *Polygraph*, 28(3), 209-222.
- Light, G. D. (1999). Numerical evaluation of the Army Zone Comparison Test. *Polygraph*, 28(1), 37-45.
- National Research Council (2003). *The Polygraph and Lie Detection*. National Academies Press: Washington, DC.
- Podlesny, J. A., & Truslow, C. M (1993). Validity of an expanded-issue (modified general question) polygraph technique in a simulated distributed-crime-roles context. *Journal of Applied Psychology*, 78(5), 788-797.
- Raskin, D. C., Kircher, J. C., Honts, C. R., & Horowitz, S. W. (1988). *A study of the validity of polygraph examinations in criminal investigations*. Final Report, National Institute of Justice, Grant No. 85-IJ-CX-0040.
- Senter, S. M. (2001, July). *New approaches for increasing polygraph accuracy*. Presentation given at the annual meeting of the American Polygraph Association, Indianapolis, IN.
- Senter, S. M. (2003). Modified General Question Test decision rule exploration. *Polygraph*, 32(4), 251-263.
- Veritas Center. (2005). *Polygraph Issues*. Retrieved February 21, 2006 from:
http://www.veritascenter.org/documents/Polygraph_FAQ.htm#PolygraphIssues

Appendix A. Instructions to participants.

Dear Examiner:

This disk contains 100 field cases where ground truth was independently confirmed. They are numbered between 001 and 999, and are in PDF format. To open the files, you will need Adobe Reader version 6.0 or higher, which can be downloaded for free at <http://www.adobe.com/products/acrobat/readstep2.html>. Once the PDF file is opened, you can enlarge the chart as desired.

Also on this disk is a standard score sheet for 7-position scoring. You must make 100 copies of this template, one for each of the 100 cases. You are not obligated to score the cases, and may make your decisions using other scoring systems or by global analysis. However, you must return the 100 hard copy score sheets to the address below to record your decisions.

You are encouraged to use 7-position scoring, as it will permit us to apply a specially developed set of decision rules that may improve your decision accuracy. The minimum accuracy established by ASTM for the Marin Protocol is 86% without inconclusives and no more than 20% inconclusives. If you submit the 7-position scores, the higher accuracy of either your own decision rules or the specially developed decision rules will be recorded. In the case that only the specially developed decision rules achieve the minimum accuracy, you will be obligated to use those rules if you choose to perform examinations under the Marin Protocol.

The 100 cases were all conducted using the federal Zone Comparison Technique. The relevant questions are 5, 7 and 10. The comparison questions are all exclusionary probable-lies, and are labeled as 4, 6 and 9.

The time needed to complete all cases depends on your individual pace, but the average time is about eight hours. For best performance, we encourage you to spread the project over several days to avoid errors from fatigue. Because the project was devised to assess individual performance, you must not consult with any other examiner regarding the cases. Safeguards are in place to discourage collaborations.

All further questions can be directed to dkrapohl@aol.com. With your inquiries be sure to state that you have Packet (A - E), and identify the specific cases by the three-digit file number.

The mailing address to send your score sheets and to return the disk is:

Krapohl
PO Box 10411
Ft. Jackson, SC 29207

Good luck!

Instructions to Authors

Scope

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

Manuscript Submission

Manuscripts should be in English, and submitted, along with a cover letter, to Editor, American Polygraph Association, PO Box 10342, Ft. Jackson, South Carolina 29207 (USA). The cover letter should include a telephone number, return address, and e-mail address. Authors should also state clearly in the cover letter if they wish to submit their manuscript to a formal peer-review. The preferred method of manuscript submission is as an email attachment (MS Word, WordPerfect, or PDF format) with the cover letter included in the body of the email. Send to the Editor at:

DSenter@sc.rr.com

Authors without Internet access may also submit manuscripts on computer disk along with 5 paper copies to the editorial address above. As a condition for publication, authors shall be required to sign a statement that all text, figures, or other content in the submitted manuscript is correctly cited, and that the work, all or in part, is not under consideration for publication elsewhere.

Manuscript Organization and Style

All manuscripts must be complete, balanced, and accurate. All authors should follow guidelines in the *Publication Manual of the American Psychological Association* (4th edition). The manual can be found in most public and university libraries, and can be ordered from: American Psychological Association Publications, 1200 17th Street, N.W., Washington, DC 20036, USA. Authors are responsible for assuring their

work includes correct citations. Consistent with the ethical standards of the discipline, the American Polygraph Association considers quotation of another's work without proper citation a grievous offense. The standard for nomenclature shall be the *Terminology Reference for the Science of Psychophysiological Detection of Deception* included in this volume. Legal case citations should follow the *West* system.

Manuscript Review

A single Associate Editor will handle papers, and the author may, at the discretion of the Associate Editor, communicate directly with him or her. For all submissions, every effort will be made to provide the author a review within 12 weeks of receipt of manuscript. Articles submitted for publication are evaluated according to several criteria including significance of the contribution to the polygraph field, clarity, accuracy, and consistency.

Copy-editing

The Editor reserves the right to copy-edit manuscripts. All changes will be coordinated with the principal author.

Copyright

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

Professional Copies

The senior author will receive ten (10) copies of the journal issue in which the article appears.