EUROPEAN POLYGRAPH

PUBLISHED QUARTERLY

Volume 3 2009 Number 1 (7)

Andrzej Frycz Modrzewski Krakow University

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Contents

Book reviews

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Tuvia Shurany*
Liecatcher Polygraph Series
Jerusalem, ISRAEL
Einat Stein
Department of Psychology
Bar Llan University, ISRAEL
Eytan Brand
Israeli Security Agency
ISRAEL

A Field Study on the Validity of the Quadri-track Zone Comparison Technique

Background

This field study is the third published piece of field research on the validity of the Quadri-Track Zone Comparison Technique. The Quadri-Track ZCT was initially developed in 1977 by James Allan Matte as a result of field experiments designed to resolve the problem of false positives in psychophysiological

* tuvia@liecatcher.com

veracity (PV) examinations using the polygraph. Its theory and methodology were published in the American Polygraph Association's journal *Polygraph* in December 1978 and in several textbooks (Matte 1980, 1996, 2000 and 2002). The first field validation study on the Quadri-Track Zone Comparison Technique was published in 1989 in *Polygraph* (Matte 1989a: 4-18) from a doctoral dissertation (Matte & Reuss 1989b: 01452-1502). The second field study on the Quadri-Track Zone Comparison Technique was published in *Physiology & Behavior*, the official peer-reviewed journal of the International Behavioral Neuroscience Society (Mangan et al. 2008: 95-1-2). The results of this field study apply only to the Quadri-Track Zone Comparison Technique when used in its pure form without justifiable deviation. The Quadri-Track Zone Comparison Technique is a polygraph technique used exclusively for single-issue tests.

The Quadri-Track Zone Comparison Technique employs the basic test structure and quantification system of the Backster Zone Comparison Technique (Backster 1963/1979) with some refinements and the addition of a third spot (inside track) consisting of a control/relevant question pair to deal with an innocent examinee's *Fear of Error* and the guilty examinee's *Hope of Error* (See Table 1). Some of the differences between the two techniques include the application of Backster's "Either-Or" rule[**], wherein the Quadri-Track ZCT restricts the comparison of each relevant question to the control[***] question preceding it within the same track, hence non-selective, and the assignment of a *minus one* score rather than a zero (in the Pneumo and Cardio tracings only) when the relevant question elicits a significant reaction and its neighboring control question also elicits an equally significant reaction, inasmuch as Backster's "Either-Or" rule deems that control question to be defective. Furthermore, the increasing score threshold required for a decision of truth or deception does not diminish with the addition of charts collected and scored.

[**] Backster's "Either-Or" rule states that a significant reaction should be present in either the red zone (relevant question) or the green zone (control question) but not in both. If the red zone indicates a lack of reaction, it should be compared with the neighboring green zone containing the larger timely reaction. If the red zone indicates a timely and significant reaction, it should be compared with the neighboring green zone containing no reaction or the least reaction. A timely and significant reaction to both the red zone and green zone question being intercompared indicates a serious question defect in the green zone question (Backster 1963/1979).

[***] The term "comparison" question has replaced "control" question to comply with the general scientific literature. Nevertheless, this study continues to use the term "control" question to avoid confusion when the term is used in instances such as "comparison of the control and relevant test questions."

The *Fear of Error* by the innocent was recognized by the behavioral scientist Dr. Paul Ekman in his book "Telling Lies" (Ekman 1985). Ekman discusses the elements of "fear" in his book, which is devoted primarily to verbal and non-verbal behavior, which he relates to the polygraph test in his chapter entitled "Polygraph as lie catcher", in which he states "The severity of the punishment will influence the truthful person's fear of being misjudged just as much as the lying person's fear of being spotted – both suffer the same consequence." Ekman states that the polygraph examination, similar to behavioral clues to deceit, is susceptible to what he terms the "*Othello Error*", because Othello failed to recognize that Desdemona's fear might not be a guilty adulterer's anguish about being caught but could be a faithful wife's fear of a husband who would not believe her. Both cause an autonomic nervous response. The *Fear of Error* phenomenon was recognized and cited by the National Research Council of the National Academies' 2003 Report on the Polygraph and Lie Detection as a factor that could appreciably reduce the accuracy of field polygraph tests, and it also cited the use of countermeasures as another factor that offered a serious threat to the accuracy of field polygraph tests.

In the first field validation study (Matte 1989a: 4-18) of the Quadri-Track Zone Comparison Technique involving two separate entities and 122 confirmed cases, the technique's "Inside Track" containing the Fear and Hope of Error questions presented a 5% false positive error rate and a 2% false negative error, and also reduced the "inconclusives" from 34.5% to 6%. Furthermore, it correctly identified 91% of the innocent as truthful, with a 9% inconclusive rate and no errors. It correctly identified 97% of the guilty as deceptive with a 3% inconclusive rate and no errors. It must be recognized that the Quadri-Track Zone Comparison Technique's quantification system of assigning a "minus one" score rather than a "zero" when there is an equal strong reaction to both the relevant and control questions being intercompared in accordance with Backster's Either-Or rule provides a minimum total score that exceeds the threshold or minimum score required to render a decision of truth or deception. This has the effect of nullifying physical and mental countermeasures normally applied to control questions which under the aforementioned circumstance are deemed defective while the neighboring relevant question is considered ideally formulated, hence deserving at the very least a lean towards deception which translates into a minus one score.

It should be noted that when Matte, a Backster graduate, developed the Quadri-Track ZCT in 1977, the Backster Zone Comparison Technique's "You-Phase" single-issue test format employed two relevant questions for

comparison with neighboring control questions, which formed the basis of the Quadri-Track ZCT. At that time, the Backster Zone Comparison Technique's scoring system eliminated one of the three scores from each of the two spots (tracks), that did not follow the general trend or the weakest score, thus retaining only two scores from each of the two spots (tracks) for a total of 4 scores per chart. In addition, the Backster scoring system's conclusion table required a minimum score of minus or plus 5 for the first chart, minus or plus 9 for two charts, or minus or plus 13 for three charts (Backster 1969). The reduction of the threshold scores with the collection of each chart was to compensate for the examinee's habituation to the testing process. However, Matte observed that the deceptive examinee may habituate to the control questions but his anxiety level would remain constant throughout the collection of each succeeding chart on the relevant questions. Conversely, the truthful examinee may habituate to the relevant questions but his anxiety level would remain constant throughout the collection of each succeeding chart on the control questions. Hence, Matte saw no need to reduce the score threshold with each succeeding chart for the truthful or deceptive examinee. Therefore he adopted an increasing but non-diminishing score threshold of -5 for 1 chart, -10 for 2 charts, -15 for 3 charts, -20 for 4 charts for deception; +4 for 1 chart, +8 for 2 charts, +12 for 3 charts and +16 for 4 charts for the truthful.

Matte reduced the threshold score for the truthful by one point with the explanation that control questions are structurally less intense than the relevant questions, and should therefore require a slightly lower score to reach a decision of truthfulness. At least two charts had to be collected to make a decision of truth or deception, but if the scores were marginal, then more charts had to be collected.

Matte (1978: 7-4) explains that "the four highest scores left for evaluation and tally after elimination of the two weakest scores in each chart must contain at least one -2 (D) score reflecting a strong response. I don't believe that a finding of deception should be based on charts that produce only four -1 (d) scores which I classify as minimum deception scores and which Backster initially labeled as "lean toward deception", placing it in the indefinite category. Therefore, each chart used for evaluation should contain a minimum of one -2 (D) score plus a minimum total score of -3 from the other tracings in the same chart to reach a definite conclusion of "deception". The requirement for only one -2 (D) score on each chart is based upon the principle that the subject's psychological set may be focused upon only one of the relevant

questions, that which was the greatest threat to his well-being. That question may produce a -2 or even a -3 score; however, the other relevant questions may produce only minimal response as a result of the examinee's strong focus on the question which he feels most threatening." Matte further states that "The aforementioned required consistency and uniformity in the analysis and scoring of each chart is also applied in the truthful tally at the conclusion table. The lower score is justified on the basis that weaker responses are expected from control questions, and if each of the four remaining highest scores on average a +1 (t), each reflecting a mild response to those control questions as opposed to no response to the neighboring relevant question, it can be safely assumed that the results reflect truthfulness regarding the issue for which the examinee was tested. The +8 minimum score for two charts for a truthful conclusion is within the limits set forth in the Utah Study. The required minimum scores depicted in the aforementioned conclusion table are unaffected by the insertion or omission of control/relevant questions set N. 23/24 into the control-question technique because its primary role is to recoup response energy otherwise lost by the other preceding relevant/control question sets." (Matte 1978: 7-4. Utah Study, Raskin et al. 1977: 6-1)

In 1980, Matte and Backster discontinued the elimination of the lowest score or the score that does not follow the general trend, which did not affect the decision threshold of the Backster ZCT or the Matte Quadri-Track ZCT. At that time the Federal Polygraph School used a fixed score threshold of plus or minus 6 regardless of the number of charts collected (Matte 1980).

The results of the first field validation study on the Quadri-Track Zone Comparison Technique (Matte, Reuss 1989b: 01452-1502), produced statistical predictive tables for estimating error rates, which revealed that the potential error rate of 0.0 would be attained when the average minimum score per chart reached minus 5 for Deception and plus 3 for truthfulness. The Matte-Reuss study confirmed the existing threshold for deception and caused an immediate change in lowering the threshold for the truthful from plus 4 per chart to plus 3, without increasing the inconclusive or error rate, hence the lower score threshold for the truthful was adopted and factored into the Quadri-Track Zone Comparison Technique's quantification system.

The second field study (Mangan et al. 2008: 06.004), involved 140 confirmed cases. Its inside track accurately increased the scores for the innocent by 43.6% and the guilty by 37.1%, thereby reducing the overall inconclusive rate from 19.5% to 1.4%. The Quadri-Track ZCT correctly identified 100% of the innocent as

truthful with no inconclusives and no errors. It further correctly identified 97.8% of the guilty as deceptive and 2.2% as inconclusive, with no errors. Inconclusive rates excluded, the Quadri-Track ZCT was 100% accurate in the identification of the innocent and the guilty. Inconclusives included, the utility rate was 98.6%. Blind scoring of polygraph charts showed extremely high correlations for the individual and total scores with a combined accuracy of 98.3%.

The Quadri-Track Zone Comparison Technique utilizes exclusive control questions that separate the period of time covered by the control questions from the period of time covered by the relevant questions to enable the Either-Or rule and facilitate the direction of the examinee's psychological set towards the type of questions (control or relevant) that offer the examinee the greatest threat to his/her well-being. The Fear of Error (control) question is presented to the examinee in a way that restrains an affirmative answer and produces a negative answer. Conversely, all examinees answer the Hope of Error (relevant) question in the negative and these two questions are compared and scored in the same manner as the other two control/relevant question pairs. The diagram of the Quadri-Track ZCT Test Structure in Table 1 shows the scores from all three tracks each containing a pair of control vs. relevant questions, which are added together for a total score, which is then associated to a conclusion table containing a score threshold (Table 2) that must be met or exceeded before a distinct conclusion of truth or deception can be rendered.

The "Fear/Hope of Error" question pair comprises the third track, also known as the inside track, which is located after the two traditional controls versus relevant question pairs or tracks. The "Fear of Error" question is a control question which is designed to determine the degree of fear that an examinee may have that an error will be made on the test regarding the target issue for which he is being tested, that only an innocent examinee should experience. Conversely, the "Hope of Error" question is a relevant question which is designed to determine whether or not the examinee is hoping that an error will be made on the test regarding the target issue which only a guilty examinee should experience.

Both the Fear of Error control question and the Hope of Error relevant question contain the suffix "regarding the target issue" which is thoroughly explained to the examinee during the pre-test interview and during the review of the test questions prior to the collection of the physiological data. Both questions contain the exact wording or meaning in cases of foreign translation or comprehension, except for the words "afraid" and "hoping." An example is as follows:

Are you afraid an error will be made on this test regarding the target issue?
Are you hoping an error will be made on this test regarding the target issue?

The above suffix removes the stigma of using an emotion-laden term associated with a sensitive crime or matter.

The "Fear of Error" question is designed to compensate for the ineffectiveness of the control questions in competing with threatening relevant questions which were caused by the "Fear of Error." Additionally, the inside track containing the Fear and Hope of Error questions provides the polygraphist with the means of determining whether a control question should be strengthened or weakened when there is an equal response to both control and neighboring relevant question or no response to either zone. This choice is not available to other zone comparison tests. It should be noted that both the Backster ZCT and the Quadri-Track ZCT mandate that once the test questions have been reviewed with the examinee, the collection of the data must not be interrupted with any language that would influence the examinee's psychological set towards the control or relevant questions (Matte 2007a: 36-2). The sole exception is when there is no response to either the relevant or the control questions. Then the control questions only are reviewed with the examinee, in accordance with Backster's Eight-Reaction Combination Guide (Backster 1963, 1979, 1983) or Matte's 23-Reaction-Combination Guide (Matte, 1981, 1996).[****]

The developer (Matte) of the Quadri-Track ZCT theorized that an innocent examinee's fear that an error be made on his PV examination will make the relevant questions exceedingly threatening, causing a physiological response that will compete with the control questions and bring about false positive or inconclusive results. This theory was subsequently advanced by the 2003 report of the National Research Council (NRC) of the National Academies, which stated "This theoretical argument also leaves open significant possibilities for misinterpretation of the polygraph results of certain examinees. It is plausible, for instance, that a belief that one might be wrongly accused of deceptive answers to relevant questions – or the experience of actually being wrongly accused of a deceptive answer to a relevant question – might produce large and repeatable physiological responses to relevant questions in nondeceptive

[****] Execution of the Backster or Matte Reaction Combination Guides, after starting the collection of the physiological data, which may influence or redirect the examinee's psychological set, requires the collection of at least two additional charts scored separately to remedy previous chart defects. The necessity to actually execute any of the remedies in the aforementioned guides has been found to be rare.

examinees that mimic the responses of deceptive ones." (NRC, page 74).

This report further expressed grave concern regarding the use of countermeasures that could seriously degrade the value of an otherwise valid test. The NRC stated that "Basic science and polygraph research give reason for concern that polygraph accuracy may be degraded by countermeasures, particularly when used by major security threats which have a strong incentive and sufficient resources to use them effectively. If these measures are effective, they could seriously undermine any value of polygraph security screening." (NRC, page 5) The NRC further stated that "Perhaps the most serious potential problem with the practical use of the polygraph is the possibility that examinees – particularly deceptive ones – might be able to decrease the test's accuracy by engaging in certain behavior, countermeasures, designed to produce nondeceptive test results." (NRC, page 139)

Unlike some other polygraph techniques that use a selective approach in the comparison of the relevant question to either the control question that elicited a large response or the least response, the Quadri-Track ZCT uses a non-selective approach, in that it confines each relevant question with the control question immediately preceding it into a track that restricts the comparison of each relevant question to the control question within that same track. Inasmuch as Backster's "Either-Or" rule dictates that when the relevant question and the control question against which it is being compared both contain significant physiological responses, the relevant question having been ideally formulated and based on solid facts is deemed effective whereas the control question must be defective, thus Backster will ignore the defective control question and compare the responsive relevant question to the other neighboring control question containing little or no response which is deemed effective. However, the use of countermeasures on all control questions would preclude the availability of a control question with "little or no response" against which to make a comparison which would result in zero scores throughout each relevant/control comparison for a final inconclusive decision.

The successful use of countermeasures requires the recognition of all control questions, which is easily obtained through readily available literature on polygraph tests. Therefore, the deceptive examinee will apply his physical or mental countermeasure to all control questions, which, if successful, will create a significant response to each control question. However, the deceptive examinee will not be able to suppress a significant response to the relevant questions to which he is being deceptive and, in the Quadri-Track ZCT, that deceptive

response to the relevant question will be compared with that reactive control question preceding it which will be deemed defective, hence a minimal deceptive score of minus one will be assigned to that track or question pair. Therefore, even with only those minimum scores which would tally at least a minus 6 score per chart, a deceptive result would occur in that the minimum score threshold for the Quadri-Track ZCT is an average of minus 5 per chart. As a result, regardless of the type of countermeasure employed, whether mental or physical, it will not effectively hamper the decision-making process and a valid and reliable result.

A Field Study of Three Methods of Comparison When the Relevant Question Elicits a Strong Response, (Matte, 2007b) which was presented at the Backster School of Lie Detection, tested Backster's "Either-Or" Rule and Anti-Climax Dampening Concept. The 123 confirmed guilty cases used in that study revealed that the use of Backster's "Either-Or" rule and concept produced the least number of inconclusives and no errors when compared with two other established scoring systems.

Table 1. Matte Quadri-Track ZCT Test Structure

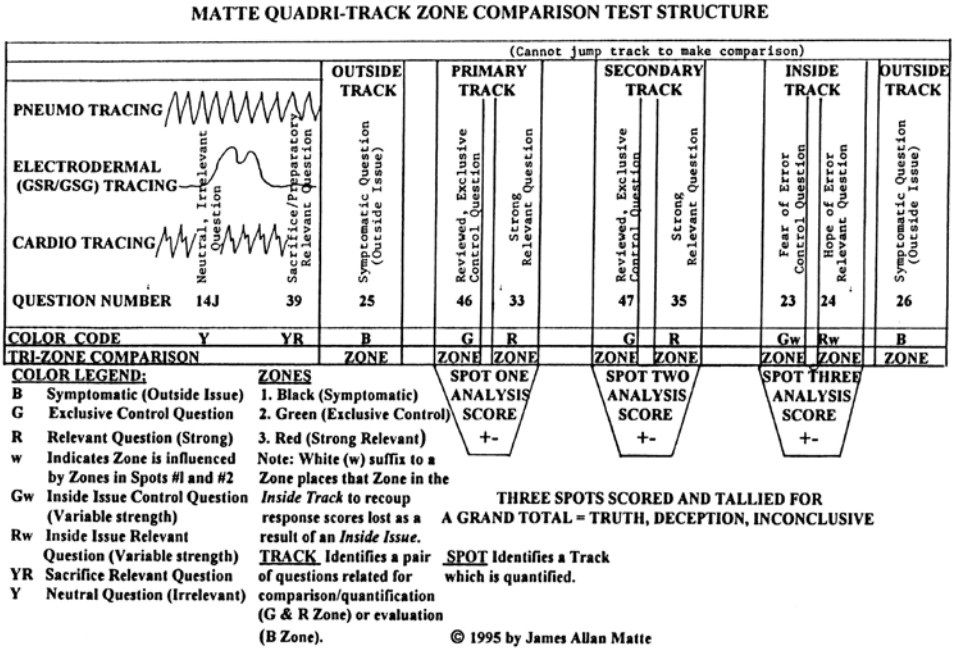


Figure V-3. Matte Quadri-Track Zone Comparison Test Structure.

Format of Quadri-Track Zone Comparison Technique

14 Neutral, Irrelevant Question.
39 Preparatory Sacrifice Relevant Question dealing precisely with single-issue covered by relevant questions #33 and #35.
25 Symptomatic (Outside Issue) Question.
46 Reviewed Exclusive Control Question.
33 Short and Direct Relevant Question.
47 Reviewed Exclusive Control Question.
35 More Descriptive Version of Relevant Question #33.
23 Fear of Error Control Question.
24 Hope of Error Relevant Question.
26 Symptomatic (Outside Issue) Question.

Table 2

The Quadri-Track ZCT Numerical Score Sheet and Conclusion Table

STIMULATION TEST DATA:	NUMBER SELECTED: CHART NUMBER:								

Quadri-Track Tri-Zone Quantification System Score Table

CHART 1	NDI	INDEF	DI			NDI	INDEF	DI			NDI	INDEF	DI			
PNE (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
EDA (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CAR (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CHART 2	NDI	INDEF	DI			NDI	INDEF	DI			NDI	INDEF	DI			
PNE (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
EDA (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CAR (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CHART 3	NDI	INDEF	DI			NDI	INDEF	DI			NDI	INDEF	DI			
PNE (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
EDA (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CAR (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CHART 4	NDI	INDEF	DI			NDI	INDEF	DI			NDI	INDEF	DI			
PNE (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
EDA (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
CAR (33)	+3 +2	+1 0 -1	-2 -3	= ( )	(35)	+3 +2	+1 0 -1	-2 -3	= ( )	(24)	+3 +2	+1 0 -1	-2 -3	= ( )		
TARGET ( )		TOTAL: ( )					TOTAL: ( )					TOTAL:		( )		

GRAND TOTAL: ( )
FOR ( ) CHARTS.

% Pop: _____
P.E.: _____

CONCLUSION TABLE

RESULTS FOR 1 CHART	CIRCLE APPROPRIATE NUMBER BELOW +27 to +3 TRUTH +2 to -4 INDEFINITE - 5 to -27 DECEPTION
RESULTS FOR 2 CHARTS	CIRCLE APPROPRIATE NUMBER BELOW +54 to +6 TRUTH +5 to -9 INDEFINITE -10 to -54 DECEPTION
RESULTS FOR 3 CHARTS	CIRCLE APPROPRIATE NUMBER BELOW +81 to +9 TRUTH +8 to -14 INDEFINITE -15 to -81 DECEPTION
RESULTS FOR 4 CHARTS	CIRCLE APPROPRIATE NUMBER BELOW +108 to +12 TRUTH +13 to -19 INDEFINITE - 20 to -108 DECEPTION

The purpose of this field study is to conduct an independent evaluation of the Quadri-Track Zone Comparison Technique (using confirmed polygraph examinations conducted by qualified polygraphists trained in the aforesaid technique who are employed by the Liecatcher Polygraph Services and Polygraph Center, Thailand).

A review of the existing literature (Ansley 1983, Matte 1996, 2000) on the validity of psychophysiological veracity PV examinations using the polygraph revealed that twice as many studies were conducted on the validity and reliability of PV examinations in a laboratory setting than those using real-life cases. Research conducted in a laboratory setting using mock paradigms lacks two very important elements that are present in real-life situations, namely "Fear of Detection" by the guilty examinee, and "Fear of Error" by the innocent examinee. Because the Quadri-Track Zone Comparison Technique particularly addresses the innocent examinee's "Fear of Error" and the guilty examinee's "Hope of Error" it was vital that this study use data acquired from real-life cases.
Thus, in this study we explore the validity of the Quadri-Track ZCT among field cases by comparing the decisions reached by the polygraphists using this technique with the results of confessions which solved these cases to determine the number of false positives, false negatives, and inconclusives, the latter as a measure of utility, not accuracy. We also compare the accuracy of the Quadri-Track ZCT with vs. without the use of the inside track's Fear/ Hope of Error questions.

Procedure

All specific-issue PV examinations conducted with the Quadri-Track ZCT by the Liecatcher Polygraph Services and Polygraph Center at Thailand from 1 January 2007 through 30 April 2008 were reviewed. There were 165 cases, of which 57 were later solved by confessions. Thus, 57 of the total of 165 available cases (34.5%) were subsequently solved, providing a base of confirmed cases for study. Of these 57 confirmed cases, 28 were confirmed as deceptive (49.1%) and 29 were confirmed as nondeceptive examinees (50.9%). The subject population of the 57 cases included 26 men and 31 women. There were 42 Thais, 4 Israelis, 4 Chinese, 2 Columbians, 1 American, 1 Vietnamese, 1 Burmese, 1 from the United Kingdom and 1 Australian. The age range was 19 to 61 and averaged 29.9. There were 5 crimes against property and 52 against people.

The polygraph instruments used in this study were Lafayette LX 4000 computerized polygraphs, which recorded thoracic and abdominal breathing patterns, electrodermal and cardiovascular activity, and covert movement.

Four polygraphists participated in this field research. All of them were formally trained in the use of the Quadri-Track ZCT. The 7-position scale scoring system was used in the analysis of the physiological data collected in each PV examination, by both the original polygraphists and quality control reviewers. The final decision as to truth or deception was made after a quality control review by another polygraphist not involved in the polygraph examination.

In this field research, we compared the final decisions reached by the polygraphists and quality control reviewers with the results of the field cases that were solved by confessions, to determine the number of false positives, false negatives, and the inconclusives rate, the latter as a measure of utility, not accuracy. We also collected the scores from each polygraph chart on each track where a comparison was made between a control and a relevant question, to determine the effect that the inside track (Fear/Hope of Error) had on the results of each polygraph test.

Academic arguments against the use of confessions as a criterion for ground truth in field examinations have been published (Iacono 2008: 06.001, Verschuere et al. 2008: 06-002) as objections to field studies that used confessions as ground truth (Mangan et al. 2008a: 03.001). Their objections were primarily based on the assumption that the confessions were coerced from the examinees confronted with the test results, which were allegedly not acquired independently of the confessions. It was also argued that the errors would most likely be found in the unconfirmed cases of examinees whose responsiveness was somehow different from examinees in the confirmed cases. Furthermore, guilty examinees whose test results showed no deception would not be subjected to an interrogation and subsequent confession, and thus would fall into the category of unconfirmed cases. These concerns by Iacono and Verschuere et al. would have some merit under past testing conditions, which they erroneously assumed still exist in all current polygraph techniques. Advances in instrumental technology, which includes motion sensors and the evolutionary progress in the psychological structure of test formats and protocol, have significantly improved the objectivity, accuracy and standardization of psychophysiological veracity examinations using the polygraph. The rebuttal (Mangan et al. 2008b: 06.004) to Iacono and Verschuere et al.'s objections to their use of confessions offers

compelling arguments including research studies (Light & Schwartz 1999: 28.3. Mason, 1991) that support the use of confessions as a criterion for ground truth. Mangan et al. point out that Iacono and Verschuere et al.'s objections presume that polygraph examinations conducted in their field study were conducted in a vacuum. "Unlike laboratory studies where there is no post-test connection, field studies of real-life cases are connected to post-test investigations and adjudications that can reveal errors or corroborate test results, which is another form of validity confirmation." Mangan et al. also pointed out in their rebuttal that they "calculated the average score for the unconfirmed and confirmed cases which revealed no significant difference in the reactivity of the subjects between the confirmed and unconfirmed cases, and there was no significant difference in the inconclusive rate, all of which indicates no significant difference in the examinees whose cases were unconfirmed, and the confirmed cases appear to be a representative sample of the total cases." They further pointed out that the results of all polygraph examinations conducted in their field study were entirely based on the analysis and numerical scores of the physiological data collected from each examinee in strict accordance with the technique's protocol, thus totally independent of any ensuing confessions. Furthermore, all polygraph examinations were audio-video recorded as required by the American Society for Testing and Material (ASTM) and American Polygraph Association (APA) standards of practice, which provided a quality control review that would expose any procedural violations that would invalidate the polygraph examination or the ensuing confession.

Further published research and arguments in support of confessions used as a criterion for ground truth in field research studies of PV examinations can also be found in Krapohl et al. 2003: 32-4, Hontz 1996: 123-4, Raskin et al. 1988: 85-IJ-CX-0040, Horvath 1977, 62-2). These studies refute Iacono's unsubstantiated claims of sampling bias in the use of confessions as a criterion for ground truth.

Results

The accuracy of the Quadri-Track ZCT with vs. without the use of the inside track's questions was compared among confirmed innocent and confirmed guilty cases. As can be seen in table 3, with the inside track's Fear/Hope of Error the QTZCT scoring system found 100% of the confirmed Innocent cases as truthful, with no errors and no inconclusives. Without the inside track's Fear/Hope of Error, the Quadri-Track ZCT scoring system would have found

69% of the innocent cases as truthful, with 31% inconclusives. Therefore the inside track's Fear/Hope of Error reduced the inconclusives from 31% to 0%, and increased in 31% the rate of accurate truthful decisions. Among the confirmed guilty cases, the Quadri-Track ZCT system with the inside track's Fear/Hope of Error found 92.9% as deceptive, 7.1% as truthful (i.e. false negative) and no inconclusives. Without the inside track's Fear/Hope of Error, the Quadri-Track ZCT system would have found 21.4% of the guilty as deceptive, 7.1% as truthful and 71.5% inconclusives. Therefore the inside track's Fear/Hope of Error questions reduced the inconclusives from 71.5% to 0% and increased in 71.5% the rate of accurate deceptive decisions. Overall, compared to ground truth, polygraph decisions using the Quadri-Track ZCT without the inside track's Fear/Hope of Error were accurate in 45.61% (26/57) of the confirmed cases, wrong in 3.5% (2 cases), and with inconclusive results in 50.89% of the cases. With the inside track's Fear/Hope of Error, polygraph decisions were accurate in 96.5% (55/57) of the confirmed cases, wrong in 3.5% (2 cases), and with inconclusive results in none of the cases.

Table 3. Accuracy of polygraph outcome compared to ground truth, using matte Quadri-Track ZCT with vs. Without inside track's fear/hope of error

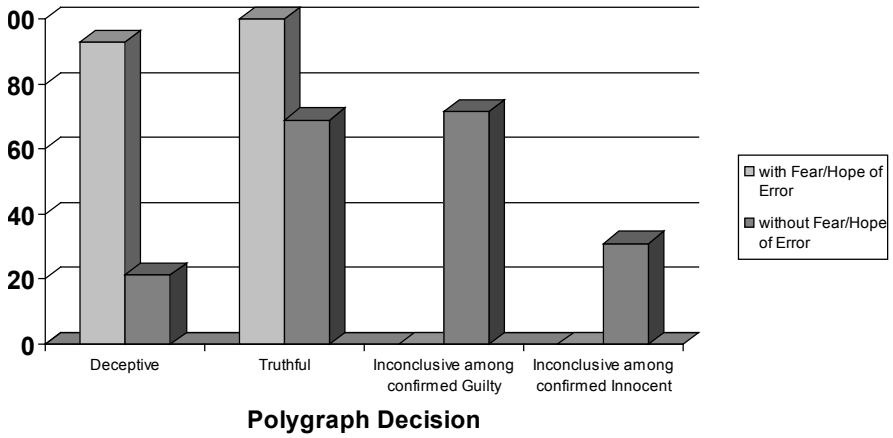
Outcome for the Polygraph Decisions separately for "innocent" and "guilty" cases compared to known confirmed cases. The Matte Quadri-Track ZCT was used to reach the decisions using original/current scoring method for the value of the inside track's Fear/Hope of Error in arriving at decisions.

GROUND TRUTH		POLYGRAPH DECISION				
		Truthful	Deceptive	Inconclusive	c^2	Eta^2
		With *Inside Track*				
Innocent	Number	29	0	0		
	Percentage	100	0	0		
Guilty	Number	2	26	0		
	Percentage	7.1	92.9	0	$c^2_{(1)}=49.28^{***}$	0.87
		Without *Inside Track*				
Innocent	Number	20	0	9		
	Percentage	69.0	0	31.0		
Guilty	Number	2	6	20		
	Percentage	7.1	21.4	71.5	$c^2_{(2)}=24.89^{***}$	0.44

*** P<.001

These comparisons show that the inside track is important in significantly reducing the number of inconclusives and increasing the number of accurate decisions when the Matte Quadri-Track ZCT is used. Moreover, the values of the ETAs presented in table 3 revealed a doubled effect size when the inside track is added for the Quadri-Track ZCT (eta^2=0.87) as compared to the effect size without it (eta^2=0.44). In other words, adding the Inside-Issue adjustment for the Quadri-Track ZCT has significantly increased its accuracy and doubled its utility rate. Figure 1 illustrates the additions in accuracy and utility as a result of adding the inside truck.

Figure 1. Percentage of accurate decisions and inconclusives with vs. Without inside track's fear/hope of error



Analysis shows that the inside track's Fear of Error control question generated an adjustment to the confirmed innocent scores by increasing their scores an average of +2 (59%) per case. The average total score per innocent case (i.e. the mean chart score) without the Fear of Error question adjustment was 3.39 and with this question 5.39. Among the confirmed guilty cases, the inside track's Hope of Error relevant question generated an adjustment to the scores by decreasing them (increasing its value) an average of –2.54 (71.75%) per case. The 2 false negative cases had no score adjustment due to the production of zero scores by the inside track. The average total score per guilty case without the Hope of Error question adjustment was -3.54, and with it was -6.08. These adjustments result in the significant reduction of inconclusives accompanied by increasing decision accuracy rate. This indicates that the "Fear/Hope of Error" factor as measured by the inside track

significantly fortifies the decision-making process, and cannot be ignored. Finally, we compared the results of the Matte Quadri-Track ZCT with the inside track, among the confirmed (n=57) vs. the unconfirmed (n=108) cases. The results show no significant difference between confirmed and unconfirmed cases, in the frequency of the three decisions (inconclusives: 0%, 0.9% respectively; "truthful": 54.4%, 52.8% respectively and "deceptive" 45.6%, 46.3% respectively; $c^2_{(2)}$ =0.55, p>.5). The results also show no significant difference in the average score per chart for the confirmed vs. not confirmed cases, both with and without adding the inside track ($t_{(163)}$ = .14, p>.5; $t_{(163)}$ = .20, p>.5 respectively). These data show that there is no significant difference in the reactivity or responsiveness of the examinees in the confirmed versus the unconfirmed cases. We therefore fail to see any difference in the examinees whose cases were unconfirmed, and the confirmed cases appear to be a representative sample of the total cases.

Conclusions

The data in this field study show that the Quadri-Track Zone Comparison Technique correctly identified 100% of the innocent as truthful and 92.9f% of the guilty as deceptive, with no inconclusive cases. Overall, the accuracy rate for the truthful and deceptive was 96.5%. According to the scientific literature pertaining to psychophysiological veracity (PV) examinations, there is a significantly greater likelihood of making errors against the innocent than against the guilty examinee (OTA 1983, Bersh 1969, Barland & Raskin 1975, NRC of National Academies 2003). In its 2003 report, the National Research Council of the National Academies expressed the belief that an innocent examinee's fear of error regarding the outcome of their PV examination could result in a false positive. Additionally, the NRC of National Academies indicated that PV examinations were susceptible to countermeasures and false negative results. There is no question that these issues merit serious consideration, and we believe that the Matte Quadri-Track Zone Comparison Technique has demonstrated through the Matte & Reuss 1989 field study, that of Mangan et al. (2008), and this current study that it is able to cope with and overcome the Othello error and countermeasures with a very high degree of accuracy.

References

Ansley, Norman (1983). A compendium on polygraph validity. *Polygraph*, 12/2: 53-62.

Backster, C. (1963/1979). Standardized Polygraph Notepack and Technique Guide: Backster Zone Comparison Technique. *Backster School of Lie Detection*, New York, N.Y.

Backster, C. (1969) Standardized Polygraph Notepack and Technique Guide, Backster Zone Comparison Technique.

Backster, C. (1969). Technique fundamentals of the Tri-Zone Polygraph Test. New York: *Backster Research Foundation*.

Backster, C. (1974). Anticlimax Dampening Concept. *Polygraph*, 3/1: 48-50.

Backster, C. (2003, December 15-19). 46th Annual Polygraph Examiner Work Conference. *The Backster School of Lie Detection*, San Diego, CA.

Barland, G. H., Raskin, D. C. (1975). An evaluation of field techniques in the detection of deception. *Psychophysiology*, 12: 321-230.

Bersh, P. J. (1969). A validation of polygraph examiner judgments. *Journal of Applied Psychology*, 53/5: 399-403.

Ekman, P. (1985). Telling *Lies – a How-To-Guide for all those who want to detect lies.* Berkley Books, New York: NY

Honts, C. R. (1996). Criterion development and validity of the CQT in field application. *Journal of General Psychology*, 123/4: 399-324.

Horvath, F. (1977). The effect of selected variables on interpretation of polygraph records. *Journal of Applied Psychology*, 62/2: 127-136.

Iacono, W. C. (1991). Can we determine the accuracy of polygraph tests? *Psychophysiology*, 4/20: 207.

Iacono, W. C. (2008). The accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Physiology & Behavior*, 95/1-2: 24-26.

Krapohl, D. J., Shull, K. W., Ryan, A. H. (2003). Does the confession criterion in case selection inflate polygraph accuracy estimates? *Polygraph*, 32/4: 234-244.

Light, C. C., Schwartz, J. R. (1999). The relative utility of the forensic disciplines. *Polygraph*, 28/3: 240-258.

Mason, P. (undated). Association between positive urinalysis drug tests and exculpatory examinations. Unpublished doctoral dissertation. United States Army Criminal Investigation Command. Baltimore: MD.

Mangan, D. J., Armitage, T. E., Adams, G. C. (2008a). A field study on the validity of the Quadri-Track Zone Comparison Technique. *Physiology & Behavior*, 95/1-2: 17-23.

Mangan, D. J. Armitage, T. E., Adams, G. C. (2008b). Rebuttal to objections by Iacono and Verschuere, et al. *Physiology & Behavior,* 95/1-2: 29-31.

Matte, J. A. (1976). A polygraph control question validation procedure. *Polygraph*, 5/2: 170-177.

Matte, J. A. (1978). Quadri-Zone Comparison Technique. *Polygraph*, 7/4: 266-280.

Matte, J. A. (1980). *The Art and Science of the Polygraph Technique*. Springfield, Illinois: Charles C. Thomas.

Matte, J. A. (1980). Polygraph Quadri-Zone reaction combination guide. *Polygraph*, 10/3: 186-193.

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

Matte, J. A. (2002). *2002 Supplement to Forensic Psychophysiology Using The Polygraph*. Williamsville, New York: J.A.M. Publications.

Matte, J. A. (2007a). Psychological structure and theoretical concept of the Backster Zone Comparison Technique and the Quadri-Track Zone Comparison Technique. *Polygraph,* 36/2: 84-90.

Matte, J. A. (2007b). A field study of three methods of comparison when the relevant question elicits a strong response. Data presented at the Backster School of Lie Detection. San Diego: CA.

Matte, J. A., Reuss, R. M. (1989b). Validation study on the polygraph Quadri-Zone Comparison Technique. *Research Dissertation Abstract*, LD 01452, Vol. 1502, 1989 Proquest Information and Learning (Formerly known as University Microfilm International).

Matte, J. A., Reuss, R. M. (1989a). A field validation study of the Quadri-Zone Comparison Technique. *Polygraph*, 18/4: 187-203.

Patrick, C. J., Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, 76/2: 229-238.

National Research Council of the National Academies. (2003). *The Polygraph and Lie Detection.* Washington, D. C.: The National Academies Press.

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

Office of Technology Assessment. (1983, November). *Scientific Validity of Polygraph Testing – A Research Review and Evaluation.* Washington, D.C.

Verschuere, B., Meijer, E. Merckelbach, H. (2008). The Quadri-Track Zone Comparison Technique: It's just not science. A critique to Mangan, Armitage, and Abrams. *Physiology & Behavior*, 95/1-2: 27-28.

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Vitas Saldžiūnas*
VIP Protection Department Ministry of the Interior
Vilnius, LITHUANIA
Aleksandras Kovalenko**
Police Department Ministry of the Interior
Vilnius, LITHUANIA
Aleksandr Soshnikov***
Polyconius Centre,
Moscow, RUSSIAN FEDERATION

Probability Assessment of the Value of Psychophysiological Stimuli

During a psychophysiological test, the relative extent of psychophysiological reactions is recorded by means of a polygraph. Using testing methodology, a polygraphologist assesses these values and makes a decision: is the test subject associated with the crime; and if so, in what way? Every polygraphologist assesses which stimulus was able to produce psychophysiological reactions.

Up until now, psychologists and psychophysiologists have intuitively considered the value of a stimulus to be its meaningfulness to an individual, its ability to

* vitas.saldziunas@vad.lt

** aleksandr.kovalenko@policija.lt

*** alexandr@polyconius.ru

attract his attention, its relevance to him, etc. Y. Kholodny (2006) believes that the value of a stimulus is the relationship of an information stimulus and a study subject in the specific situation of the content of the question to be decided. Particularly at present in applied psychophysiology (polygraph testing), this kind of interpretation of the value of a stimulus may be debatable.

Up until now, sometimes attempts have been made to assess the value of stimuli according to the extent of the psychophysiological reaction. Since the psychophysiological reaction being measured is by nature a complex phenomenon, not noted for its specificity or stability, any attempt to perform such a measurement becomes unreliable. This is first of all demonstrated by the fact that a stimulus of certain significance will every so often periodically cause the study subject to experience psychophysiological reactions of varying size and type. The extent of a psychophysiological reaction depends on many external and internal factors, which it is not possible to fully assess in the context of a specific psychophysiological test (V. Varlamov, G. Varlamov, 2000).

It is complicated to interpret the cause of psychophysiological reactions because of their nature and because of the consistency of psychophysiological reactions. Whereas in classic psychophysiological testing simple stimuli are usually used, in polygraph testing the stimulus is not a question in the usual sense (a word, a photo, an object, a diagram, a chart, a map, etc.) (Nakayama, 2002; Saldziunas, 2008), but rather the whole complex context resulting from the questions, answers and the general testing situation, which is conditioned by a broad range of external and internal factors that are linked one way or another with polygraph testing and influence the subject being tested.

Let us say that in a simple situation a study subject is affected by sound stimulus (S), for example, a question requiring a simple yes/no answer. Within a very short time the person experiences a psychophysiological reaction (R), which is recorded in some form by the polygraph. It is obvious that the psychophysiological reaction recorded (R) will be the function of several dimensions:

$$R = f(S, E, P, M) \quad (1)$$

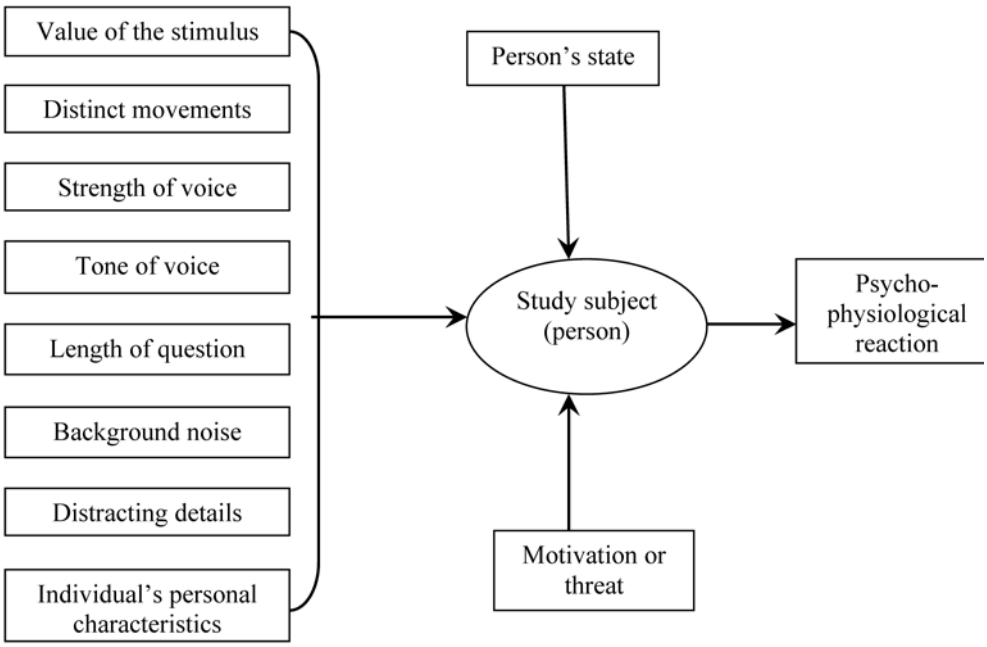
Where S = the nature of the stimulus,
E = environmental factors,
P = the individual's personal characteristics,
M = movements of the study subject.

Each one of these dimensions depends on other factors. For example, the nature of the stimulus (S) depends on the loudness (L) of the specialist administering the question, his tone of voice (W) and its duration (T). In addition, the nature of the stimulus (S) includes the subjective interest (I) that each person perceives in the stimulus. The subjective interest (I) of the stimulus depends in turn on the content of the question and the study subject's subjective view of it, including whether or not the study subject has answered this question before the testing takes place. Environmental factors (E) depend on background noise (N), details that distract attention (D) and so forth. It is very difficult to evaluate what an individual's personal characteristics depend on. It should be noted that P depends on the time, i.e. P(t). With regard to the factors identified above, the psychophysiological reaction may be expressed thus:

$$R = f(I,M,L,W,T,N,D,P(t)...) \quad (2)$$

This can be seen in the diagram in Figure 1.

Figure 1. Diagram of psychophysiological reactions measured by a polygraph.



The principal task of psychophysiological testing is to determine, during the measuring of reactions, to what extent the interest (I) of the stimulus to the patient influences the psychophysiological reaction (R). All polygraphologists know that it is not possible to totally exclude the effect of all other factors on the psychophysiological reaction. When organising polygraph testing it is possible to achieve minimal influence from the loudness of voice, tone of voice, length of question, background noise and distracting details. In a real work situation it is never possible to guarantee that any of the factors named above will not accidentally have an influence. It follows that, in repeated testing where the study subject is exposed to the same stimulus, we will not get a psychophysiological reaction (R) of the same size. It needs to be noted that all these accidental effects are of an arbitrary nature, and the psychophysiological reactions they cause are not systematic. Therefore, during the polygraph testing the polygraphologist has to decide: is the psychophysiological reaction recorded of an accidental nature or was it caused by the primary stimulus (I)? In classical methodology (Matte, 1997) it is recommended that the genuine psychophysiological reactions and the non-genuine ones (artefacts) be assessed by their extent time-wise in the polygram and by other non-systematic features.

Without doubt, doing this is a complex task that can only be achieved by an experienced polygraphologist. Using a polygraph of increased sensitivity will not help to resolve this issue. Two principles are offered for the solution of this problem:

- measurement of the psychophysiological reactions using ranking
- in order to determine convincingly whether the psychophysiological reactions are caused by the stimulus, the stimulus needs to be applied to the subject repeatedly.

A ranked evaluation simplifies the evaluation of the reactions before the next statistical or other processing of the data. This type of evaluation of physiological reactions is economic, reliable, responsive and sufficiently stable against arbitrary fluctuations. It is quite straightforward when used in real work situations.

On the basis of that which is explained above, measurement of the psychophysiological reaction (R) alone does not suffice as an assessment of the value of a stimulus. Keeping in mind that the nature of the incidental factors is arbitrary, the following principle may be formulated: the value of a

psychophysiological stimulus is proportional to the non-arbitrariness of the recorded psychophysiological reaction. That is to say, a stimulus is meaningful to a subject if the reaction caused by it is not accidental. This principle allows a quantitative evaluation to be made on the basis of probable values. It is on the basis of this principle that the ChanceCalc® algorithm used in the Diana-01 polygraph was created.

Characteristics of the ChanceCalc® Algorithm:

- may be applied when working with practically all the tests known to be used today
- its high sensitivity minimises the possibility of the Othello or Brokau trap (Ekman, 1992) occurring
- enables the possibility of measuring psychophysiological reactions in an automated way and performing an expert evaluation
- enables the possibility of presenting the polygraph test results quantitatively with the likelihood of statistical error identified
- enables a maximally convincing result to be obtained. This will be explained further.

Figure 2. Conclusion Formulation Algorithm

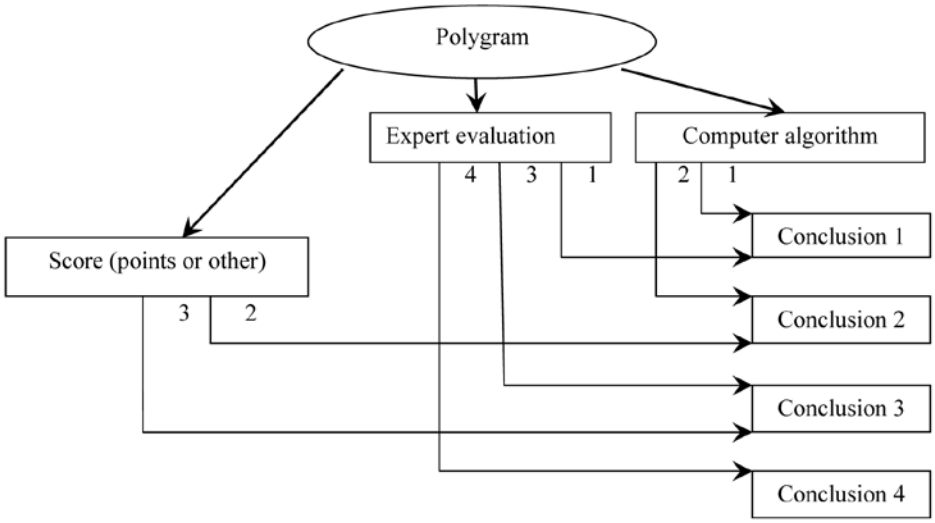


Figure 2 illustrates the algorithm for how conclusions are reached on the basis of a given polygram. This algorithm does not account for computer acoustic response signal analysis, including latent time scientific content analysis (SCAN) (Sapir, 1987), assessment of the study subject's non-verbal behaviour, etc. (Soshnikov, 2008). First of all the polygraphologist evaluates the curves of the measured psychophysiological reactions expertly or using a global approach (Kircher and Raskin, 2002) and a computer program. If the results of both evaluations coincide, the polygraphologist may write up his conclusion (Conclusion 1). If the results do not coincide, a numerical evaluation is performed. When the computer and numerical evaluations coincide, a conclusion is formulated (Conclusion 2). When the expert evaluation and the numerical evaluation coincide, Conclusion 3 is formulated. Conclusion 4 is used when no objective and convincing result could be obtained. This means that this is not a suitable case for polygraph testing, or a mistake has been made in the course of the analysis. Some possible errors are:

- not entirely accurate primary information about the event
- the questions and answers for the polygraph test were not formulated correctly
- the conditions were not appropriate for a polygraph test (various distractions)
- the study subject did not feel well or was not motivated, etc.

If Conclusion 4 is arrived at, the psychophysiological testing is either abandoned as unsuccessful or the methodological errors are fixed and the test is repeated after new questions are formulated.

From what has been stated above, we can conclude that a more reliable conclusion is obtained when the expert evaluation and the computer evaluation of the psychophysiological reaction curves are based on different principles; for example: the expert evaluation is based on the ranking principle and the computer evaluation on the probability principle.

References

Ekman P. (1992) Telling Lies. New York: Norton and Co.

Kholodny Y. (2006) Interrogation using polygraph and its scientific basis, Polygraph, 35(1), 1-21.

Kircher J.C. and Raskin D.C. (2002) Computer methods for the psychophysiological detection of deception. Handbook of polygraph testing. London: Academic Press.

Matte J. A. (1997) Forensic Psychophysiology Using The Polygraph. Scientific Truth Verification – Lie Detection. Willamsville, New York: Matte Polygraph Service.

Nakayama M. (2002) Practical use of the concealed information test for criminal investigation in Japan. Handbook of polygraph testing. London: Academic Press.

Saldziunas V. and Kovalenko A. (2008), The Event Knowledge Test (EKT) in Polygraph Examination (in case murder), European Polygraph, 2(4),137-142.

Soshnikov A. et al. (2008) Полиграф в практике расследования преступлений {*The polygraph in practice in the investigation of crimes}*, Moscow [text in Russian].

Varlamov V. and Varlamov G. (2000) Psychophysiology of Verification with Polygraph. Krasnodar: Prosvetschenije-Jug. [text in Russian].

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Jarosław Pietruszka

A renaissance in voice analysers as tools for detection of deception?

The current situation in the world, and especially the threat of terrorism, creates a demand for new technologies that allow observation of human psychophysiological reactions without entering into direct contact with the person and without that person's knowledge, to evaluate the person's honesty (deception), intentions, hidden information, and information provided.

This has resulted in a revival of interest in devices and techniques for investigating emotional changes in the sound of the human voice (Widacki 2007).

The fact that changes in the voice *are* a good indicator of emotional changes has long been known, both from general observation and from psychophysiological research.

In 1970, the American Dektor Counterintelligence and Security Company introduced to the market a device called the Psychological Stress Evaluator (PSE), used to register psychophysiological changes in the sound of the human voice (Hoddard, 2002).

The device was created by Charles McQuiston, Allen Bell, and Bill Ford, former officers of the US Army. McQuiston was a polygrapher in the American army, Bell a counterintelligence officer, and Ford an electronic engineer. One

may infer from this that the US Army conducted some research, or at least a study, of the potential to use changes in the voice to detect deception as early as the 1960s. Moreover, other sources prove that, even at that time, US army services made use of the study of emotional changes in the voice to test the honesty of their informants (Hopkins 2005).
Similar work was also conducted in the USSR in the 1960s (Kulicki 1994).
Since the introduction of the PSE to the market and efforts at its commercial use, many works have been published on this method of lie detection. The evaluation of efficiency achieved by this method has shown extreme divergence (see: Kubis 1973, Hopkins 2005, Chapman 1989, Damme 2001).
In 1998 voice analysers were disqualified by the US Equal Employment Opportunity Commission, and furthermore, the research conducted by the Polygraph Institute of the Department of Defence proved that there are no arguments suggesting that the validity of such tests exceeded statistical probability. Moreover, the Appellate Court of the State of Wisconsin recognized the results of such a test as non-credible in 2001 (Barland 2002, Horvath 2002).
Attention was turned at the time to the fact that all voice analysers register only a single physiological correlate of emotions (emotional changes of the voice), while the traditional polygraph registers at least three physiological correlates: changes in the respiratory activity, changes in the activity of the heart, and GSR (Widacki 1980, Widacki 2008).
In recent years, attention was again turned to voice analysers and the prospect of using them in detection of deception for forensic, counterintelligence, business, and human resources management purposes (Widacki 2007).
Currently available are both devices measuring only voice parameters and analysing its changes, and those integrated with polygraphs registering other psychophysiological reactions simultaneously.
Known and available on the market are the following independent devices and testing technologies related to them:

- Psychological Stress Evaluator (PSE)
- Voice Stress Analyzer (VSA)
- Layer Voice Stress Analyzer (LVA)
- Truster
- Vericator.

In recent years at the University of Warsaw a number of preliminary voice analyser trials have been conducted. The research was of the pilot trial type and covered voice analysers LVA 6.50, Ti Pi 6.40, and Truster Pro, available on the market and marketed as professional.

The voice analysers tested are software packages for computers equipped with a soundcard. The kits included a microphone, a connector for optional connection to a telephone, and hardware HASP protection against using more than a single copy of the software.
These analysers are the latest edition, advanced and developed variations that continue the line of the Truster Pro. The number of functions in the newer models is even greater.
The utterance is analysed automatically and divided into segments of approximately 2 seconds, with each separately evaluated by the device, with appropriate numerical values attached.
Each utterance of the subject undergoing the test may be analysed in real-time or afterwards, based on a tape or computer recording of the voice . The analysis covers the involuntary reactions of the organism manifested in the tone and frequency of vibrations in the vocal waves of the subject. The device makes use of the technology for analysing the stress level, manifested in the voice of the subject, calculated with a special algorithm analysing stress intensity, and measuring it and grading.
As has already been mentioned, testing psychophysiological changes of the voice can be conducted in various situations:

1) real-time testing, without the subject's awareness
2) forensic testing, akin to polygraph testing
3) analysis of sound recordings.

In case 1) the test is conducted in real-time, during the interview. The sound may be acquired from a microphone or telephone receiver, while the results of the test are displayed in real-time as simple messages or reports. The analysis may be conducted at the time while talking to the subject, or afterwards, based on the recording.
In case 2) the test is conducted like a classical polygraph test, with a voice analyser being used in place of the polygraph. The test is conducted according to one of the techniques used for polygraph testing. Analysis of emotional changes of the voice replaces the analysis of diagrams in polygraph testing.
Tested in case 3) are sound recordings acquired earlier.
The pilot research made use of the Peak of Tension (POT) test, known from routine polygraphic procedures.
The result of initial pilot testing encourages further experimental research and allows the following claims to be made:
1) without a doubt, the voice analyser detects and illustrates changes in psychophysiological reactions

2) there is a repeatability of results while using different types of voice analysers
3) deception, invoked both in laboratory conditions and in field studies, can be discovered with a voice analyser
4) marked changes occur in the recording of the voice caused by changes other than deception
5) it seems theoretically possible to single out reactions related to deception from among others portrayed by the voice analyser
6) potential voice changes related to deception look similar in laboratory studies to those conducted in real life as field studies
7) there is a broader field for using a voice analyser than in the case of a classical polygraph (for example, for an *ex post* analysis of a recorded utterance)
8) reactions recorded by the voice analyser in the POT-type tests are similar to the galvanic skin response (GSR) reactions in such classical polygraph tests
9) it was observed that answers concerning facts which the subject is not certain about result in a reaction, yet generate a smaller one than in the cases when the subject is consciously deceiving
10) manifestation of emotional changes in the voice with the voice analyser is possible also during a free conversation (e.g. while conducting negotiations), and does not involve a special test procedure, and therefore can be conducted without informing the subject
11) levels of capacity to manifest emotional changes in the voice differ from person to person
12) repetition of the test results – much like in polygraph testing – results in the weakening of emotional reactions in the voice, which is caused by habituation.

It seems that the testing of emotional changes of the voice as the method for the detection of deception requires further experimental research. Although the research conducted so far does not allow classical polygraph testing to be replaced with voice analysis, as the pioneers of the method wanted, they allow the assumption that analysis of emotional changes of the voice may firstly significantly complement classical polygraph testing, and can secondly be taken into consideration as one of the range of methods (for example together with the method for analysing eyeball movements and/or testing changes of facial temperature) incorporated into the new polygraph to allow detection of deception without direct contact with the subject and without the subject's informed consent. Thus it can be taken into account as a constituent

of the new polygraph, testing physiological correlates of emotions, other than the ones tested so far.

Literature

Barland G ,(2002): Use of Voice Changes In the Detection of Deception, Polygraph, 2002, 31, 2;
Chapman J., (1989) The Psychological Stress Evaluator as a Tool for Elicting Confessions, Criminal Justice Department, Corning Community College, New York 1989;
Damme G.v., (2001): Forensic Criminology and Psychophisiology. Truth Verification Tools, with a Special Study of Truster Pro, Durban Westville 2001;
Haddad D., Walter R., Ratley R., Smith M., (2002): Investigation and Evaluation of Voice Stress Analysis Technology: Final Report, National Criminal Reference Service NCJ, No 193832;
Hopkins C., Ratley R.,Benincasa D., Grieco J, (2005): Evaluation of Stress Analysis Technology, proceedings of 38 th Hawaii International Conference on System Sciences, 2005.
Horvath F., (2002): Experimental Comparison of the Psychological Stress Evaluator and the Galvanic Skin Response in Detection of Deception, Polygraph 2002, 31,2;
Kubis J., (1973): Comparison of Voice Analysis and Polygraph as Lie Detection Procedures, N. York 1973;
Kulicki M., (1994): Kryminalistyka. Wybrane zagadnienia teorii I praktyki śledczo-sądowej, Toruń 1994;
Lynch B., Henry D., (2002): Validity Study of the Psychological Stress Evaluator, Polygraph 2002, 31, 2;
Widacki J., (1980): Kryminalistyka, II, Katowice 1980;
Widacki J., (2007): From Forensic psychophisiology to Forensic Neurophisiology. New Trends in Examinations in the Detection of Deception, European Polygraph, 2007, 1,2;
Widacki J.,(2008): Kryminalistyka, Warszawa 2008.

Book reviews

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Antagonistic polygraph examination by Ryszard Jaworski,

Wydawnictwo Uniwersytetu Wrocławskiego
(Wrocław University Publishers)
Wrocław 2008; 63 pp.

This new book by Ryszard Jaworski provides an interesting and important overview of cases in which polygraph examination has played a crucial investigative and evidentiary role. The author selected 12 cases from the period 1991-2004, almost exclusively homicides. In each of the cases there were two suspects, and consequently two polygraph examinations. In such situations, in the words of the author, "hypotheses of perpetration may be assumed in relation to both subjects. Thus, in relation to one of them the hypothesis will be wrong, i.e. the person will be innocent. Will they display differences in physiological parameters?" (p. 8). The author calls examinations conducted under such circumstances "antagonistic examinations". The assumption of such circumstances, and the question cited above, define the boundaries of the book.

The approach of the book is casuistic. The book therefore consists of an introduction, a general introductory chapter ("Basic information on the polygraph examination"), 12 case studies and the conclusions.

The examinations presented (carried out by the author himself) used almost exclusively the Reid technique, with CQ tests being supplemented on occasion by POT tests. The book contains over 50 high-quality images of authentic polygrams or fragments thereof, all derived from the examinations

in question. The final section of each case description presents the litigation outcome of the case, the decision of the court, and the attitude of the court towards polygraph examination as evidence.

Let us begin a closer analysis of the book with a number of comments of a critical nature. First of all, I find the key purpose of the book, expressed in the question cited above – "Will they [the two subjects in an "antagonistic" situation – JK] display differences in physiological parameters?" – to be incomprehensible. This is in fact not a research question. It is a question aimed at the very core of the theory that underlies polygraph examinations. If the reader of this review has been feeling the suspense increase, please breathe freely again: the answer is "Yes", a polygraph examination does in fact enable a differentiation between a perpetrator and a non-perpetrator (see page 60, points 1 and 2). But enough irony. The error of the author appears to lie in the fact that he was searching for interpersonal differences in the intensity (and more generally in the reflection) of the physiopsychological responses of two different categories of people, but was using an inappropriate tool for it. Namely, he was using tests that were developed to evaluate interpersonal reactions to different types of impulses: in the case of RCQT tests, in responses to relevant and comparison questions, and in the case of POT tests, to padding and key (critical) questions. Comparison of responses – even to the same questions – in analogous tests but conducted on different subjects, makes no sense. Relativisation and assessment of the intensity of a response is only sensible within one examination, one test and one pair of neighbouring relevant/comparison questions. If one attempts to investigate interpersonal differences, then any conclusions will necessarily be as unspecific as those listed by the author (page 60). These conclusions have, though, been known for a long time, and have formed the core of the theory of polygraph examination.

Another reason, no less important, for the weaknesses of Jaworski's work, is the fact that he limits his interpretation of the tests to the visual (global) method, shying away from the semi-objective method using the 7-position numerical scale. The reasons for this decision are difficult to understand, since there were no obstacles to employing this technique. Had the author made the effort of interpreting the RCQT tests using this method, his results would have been immeasurably richer, would lend themselves to statistical analysis, and would open avenues for arriving at conclusions of much greater importance, conclusions that would possibly be truly innovative. This opportunity is not yet wasted – after all, the empirical material presented in the book is undoubtedly still safe in the author's possession. Possibly, and hopefully, he will make use of it again in the future.

Let us now turn to the strong points. They are more numerous than the weak points, and they make the book valuable.

First of all, the book presents empirical material, which is rare and thus rather impressive. The research is very well documented. Due to the large number of polygrams that made their way into the book, a competent reader may easily compare the author's view with his/her own interpretation of the tests. In the great majority of cases, the author also provided a list of the questions asked. Therefore, the reader has the opportunity to examine Jaworski's empirical material in depth. This is something worthy of extra praise.

The book itself is a great tribute to Reid's technique or, more generally, to the comparative techniques of examination. Objections to the author's not using in his current practice more modern tests are justifiable. However, in the early 1990s and even later UZCT tests were very uncommon, and moreover, today still Reid's technique constitutes a fully acceptable examination technique. Furthermore, there is still a good reason to dispute the usefulness of comparison question techniques. For example, in Poland a group of polygraphists remains who fail to recognize such techniques' right to exist, and much less to be used in actual examinations; they believe that only CIT tests should be used. Jaworski's work clearly demonstrates the absurdity of such an approach.

The final strength of the book – important even if not large in volume – is in the comments it includes with respect to the legal outcomes of the cases discussed. There are also excerpts from the justifications to the courts' decisions, which shed light on the courts' views on polygraph examinations (it is valuable for the foreign reader to note here that the rulings were issued after the transformation in Central and Eastern Europe, and thus were issued on behalf of the sovereign Republic of Poland). It is somewhat regretful that the author paid limited attention only to these matters.

This book is quite modest in its intentions. However, it should become obligatory reading for all polygraphists, whom it will certainly inspire and for whom it will provide valuable food for thought. For law enforcement officers and justice officers, the book offers a rich source of information. What makes it attractive is that this information comes accompanied by presentations of real-life, actual cases. Finally, theoreticians of criminal justice and forensic sciences will be satisfied with the solid amount of thought-provoking material contained.

Jerzy Konieczny

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Ewa Gruza
Ocena zeznań i wyjaśnień
[in:] E. Gruza, M. Goc, J. Moszczyński
Kryminalistyka – czyli rzecz o metodach śledczych
Wydawnictwa Akademickie i profesjonalne, Warszawa 2008

This modern and perfectly illustrated manual of forensic studies, which will definitely serve well as a compendium of knowledge for both police officers and prosecutors, and judges includes a chapter by Ewa Gruza entitled "Ocena zeznań i wyjaśnień", which can be translated literally as "assessment of evidence and explanations". Unfortunately, the level of the chapter stands out drastically from the entire manual, contains numerous material errors that may only misinform the user of the manual: all in all, it may bring harm to investigations and the judiciary.

In the chapter, the author included information on polygraphs tests (for which she uses the Polish word "variographic"). The very inclusion of polygraph tests in the chapter is a misunderstanding, as they serve neither the assessment of evidence nor of explanations. Moreover, such a use of the tests is clearly forbidden in Poland by the Code of Penal Procedure.

It goes without saying that polygraph tests for the use of an investigation or a criminal procedure may be performed only as a part of an professional study by an expert witness. In the light of today's criminal studies, such a test is a method of identifying emotional traces, and so it has been treated for at least five decades. If in this period (that is for the last 50 years) the

name "lie-detector" has been in use, it was only in the colloquial sense, with its name "lie-detector" as a rule placed within inverted commas. At least since the 1960s, the scientific literature has rather used the term "detection of Deception", with the deception defined as conscious telling of lies or withholding certain information possessed. During the last at least 20 years, the term "forensic psychophysiology" has been used to render the essence of polygraph tests. Hence the author's considerations ending in the conclusion that a polygraph test is not a light detection is strongly anachronic, belated if not by 50 then at least by 40 years.

The historic information on polygraph testing provided by the author is quite doubtful and at least challengeable. Why does she consider Benussi – and not for example Marston, Munsterberg, Mosso or Mackenzie – the precursor of such tests remains unknown. Primacy of Luria as the pioneer is also highly doubtful. The individual physiological correlates of emotions were measured and described at least a few decades earlier (see e.g. Mosso, Fere, Tarchanoff, Mackenzie, and Marston), while they were observed already in ancient times.

Describing techniques of polygraph testing, the author disqualifies Control Questions Technique (CQT), stating that "today's knowledge, especially in the field of psychology, negates this approach to variographic tests", which is an obvious deception. Today CQT are techniques commonly used and developed wherever polygraphic testing is performed. Yet the author follows here her erroneous assumption that the use of the CQT is tantamount to considering the polygraph a "lie detector".

Ignorant of foreign literature, the author claims that the technique of "ustalania wiedzy o czynie" (literally "determining the knowledge on a fact" – competitive for the CQT – was proposed by M. Kulicki in 1976. Yet in 1976, Kulicki – who much like the author of the work reviewed here did not know foreign literature – believed that he invented the Peak of Tension tests, known and applied in practice since the 1920s! Having polygraphic tests based solely on this type of testing was proposed by David T. Lykken well over a decade before Kulicki (see: D.T. Lykken: *The GSR in detection of guilt*, Journal of Applied Psychology 1959, 43, 6; by the same: *The validity of guilty knowledge test*, Journal of Applied Psychology 1960, 44, 4; by the same: *Guilty Knowledge Test – the right way to use lie-detector*, Psychology Today, 1975, 8, 10;). To make matters even more ridiculous, Lykken's works had been known in Poland much before Kulicki's "discovery". Thus, it is evident that what is lacking here is the knowledge of not only foreign but also Polish literature!

Let us provide the information withheld by the author by saying that the techniques she calls "determining the knowledge on a fact" is used today at

par with the CQT under the names of CIT (concealed information test) and GKT (guilty knowledge test).

Information on the psychological stress evaluator (*"psychologiczny analizator głosu"*), and also on hypnosis and Narco Analysis contained in the work reviewed, ignores the achievements of science in the last 25 years, and is therefore much behind the times and strongly out of date. It definitely does not present the state-of-the-art knowledge in those fields.

In turn, treatment of thermovision as a method of "assessment of evidence and explanations" at par with polygraphic testing, moreover, based on the most general information from a work by H. Kołecki from exactly 30 years ago without a reference to contemporary research is as bizarre as absolutely unacceptable.

Concluding, the reviewed chapter of the manual not only fails to provide a reliable source of information concerning today's forensic psychophysiology, but constitutes a major source of disinformation, and as such is simply harmful.

Jan Widacki

EUROPEAN
POLYGRAPH

Volume 1 • 2009 • Number 1 (7)

Ryszard Jaworski

Multi-Subject Polygraph Examination

Wydawnictwo Uniwersytetu Wrocławskiego

The field of the polygraph is a unique one that occupies an interesting and dynamic place in our society. Ryszard Jaworski provides his readers with a comprehensive overview of real case scenarios concerning a variety of murder cases.

Before we proceed in digesting this well-written unique book, let us examine the use of the polygraph. The polygraph is an exceptional investigative tool. It records a person's heart rate, pulse, and breathing as they respond to questions concerning a particular issue, in this instance: murder, rape, and theft. The polygraph examiner focuses the subject on the test issues to determine innocence or guilt. This is the process of instilling the "fear of detection" or focusing the subject's "psychological set" on the polygraph examination. Ryszard Jaworski discusses the accuracy of the polygraph in relation to multi-subject polygraph examinations.

The length of each chapter makes the book an easy read and allows the reader to stay focused on the content. Do not be misled though. The book provides a wealth of knowledge to the novice and the experienced examiner alike. The descriptive title of each chapter draws you into the case and examination, while the case study method is a valid teaching method. The case study methodology for each chapter is a unique way to describe the investigation. The ability to examine the various polygraph examination question techniques, review segments of the polygraph charts and read the overall evaluation of the case is an exceptional way to learn the position of the author.

The overall evaluation of each case by showing the case facts, prosecutors' hypotheses, polygraph examination, results, conclusion and comments adds to the value of this book's style. And, unsurprisingly to me in my experiences, it is an excellent way to learn. The description of the subject's behavior in the review and the methodology for the polygraph examination provides learning opportunities. The discussion of why a technique was or was not used (i.e. Chapter 10, biting lips, clearing throat, and subject's complaints of cuff pressure) further provides insight into the subject's use of countermeasures.
The author exposes the audience to the mindset of the perpetrator. Jaworski states that in multi-subject examinations, the actual perpetrator perceives the polygraph examination as a greater threat than when only the individual perpetrator is examined.
Jaworski concludes that multi-subject polygraph examinations constitute a new methodological point of view in polygraph examinations. Using the results from the perpetrator's uncertainty as to the results of polygraph examination of other people who are aware of their role in a particular crime (confessions of co-perpetrators or accomplices, truthful testimonies of witness), this contributes to stronger reactions to relevant questions. This is an interesting methodology that makes the reader want to learn more. It reminds me of Peak of Tension tests, where the main object or crime is placed amongst other less important but relevant objects.
The superbly presented case studies by a polygraph examiner in the field address the technical and psychological aspects of polygraph or 'lie detector' testing. The book comprises interesting case studies including polygraph charts and results. This is especially a good learning tool for new examiners and experienced examiners, or for those who must increasingly use or come into contact with polygraph testing in the criminal arena. The polygraph field is always open to case study reviews as a valid method of learning and teaching, and this book is a welcome addition to the polygraph field. It functions very well as an introduction to polygraph case studies for the novice or the experienced polygrapher. The structure of the book is very useful, in that it balances very detailed case background information with applicable test techniques and results.
Furthermore, the book provides information on polygraph examinations, question techniques, and chart scoring, and gives the results. The main objective of the book from my point of view is to highlight polygraph test techniques and results in criminal investigations.
As an expert polygrapher, one often talks about what polygraphs should be used for. Using John Reid's MGQT is excellent. I was trained in this method, and thought it interesting the way the Peak of Tension Test and

Guilty Knowledge Test were used to focus in on the guilty subject. Today, the focus of examinations is mainly on Relevant and Irrelevant testing, which in screening examinations has its place.

This book is ideal for novice and experienced forensic psychophysiologists. Attorneys, judges, law enforcement personnel, probation and parole officers, defendants, litigants, and others involved in the criminal justice system, plus psychologists, researchers, historians, business persons, employers and employees, and various educational institutions, will each for their own reasons find it singularly comprehensive, authoritative and helpful in understanding and/or utilizing the various polygraph techniques and instruments described.

Overall, this is a very balanced, teaching-based book providing varied case studies for students of all levels of experience. I would recommend it for any library with a legal, criminology or law enforcement user community. Any psychology collection would also benefit from this book.

The basic information for Authors

To publication will be accepts unpublished research papers as well as review article, case reports, book reviews and reports connected with polygraph examinations.

Submitted manuscripts must be written in English.

All papers are assessed by referees (usually from Editorial Board), and after a positive opinion are published.

Texts for publication should be submitted in the form of normalized printout (1800 characters per page) and in electronic form (diskette, CD), or sent by e-mail to Editorial Office.

The total length of research papers and review article should not exceed 12 pages, case reports – 6 pages, and other texts (book review, report) – 5 pages.

The first page of paper should contain: the title, the full name of the author (authors), the name of institution where the paper was written, the town and country.

Figures should be submitted both in printed form (laser print, the best) and electronic form.

Tables should be numbered in Roman numerals and figures in Arabic ones.

Figures, tables, titles of figures and titles of tables should be included on a separate page. The places in the text where they are to be included should be indicated.

The references should be arranged in the alphabetical order according to the surnames of the authors.

The references should be after the text.

Each reference should include: the surname (surnames) of the author (authors), the first letter of author's first name, the title of the book, year and place of the publication, the name of publisher, or the title of the paper, the full title of the journal, the year, the volume, the number and the first page of the paper.

For example (in references):

Reid J., Inbau F. (1966), *Truth and Deception: the Polygraph ("Lie-detector") Techniques*, Williams & Wilkins, Baltimore.

Abrams S. (1973), *Polygraph Validity and Reliability – a Review*, Journal of Forensic Sciences, 18, 4, 313.

and (Reid, Inbau, 1966), (Abrams, 1973) inside text.

Texts for publication in "European Polygraph" should be mail to:

"European Polygraph"
Andrzej Frycz Modrzewski Krakow University College
ul. Gustawa Herlinga-Grudzińskiego 1
30-705 Kraków (Poland)

Or e-mail: margerita.krasnowolska@kte.pl