EUROPEAN POLYGRAPH

PUBLISHED QUARTERLY

Volume 4 2010 Number 2 (12)

Andrzej Frycz Modrzewski Krakow University

EUROPEAN
POLYGRAPH

Volume 4 • 2010 • Number 2 (12)

Contents

Book reviews

EUROPEAN
POLYGRAPH

Volume 4 • 2010 • Number 2 (12)

James Allan Matte*
President of Matte Polygraph Service, Inc.
Williamsville, New York
USA

A Field Study of the Backster Zone Comparison Technique's Either-Or Rule and Scoring System Versus Two Other Scoring Systems When Relevant Question Elicits Strong Response

Background

In March 1974, Cleve Backster authored an article published in *Polygraph* entitled "Anticlimax Dampening Concept" (Backster 1974). In his article, Backster states "The anticlimax dampening concept is formulated on the well-validated psychological principle that a person's fears, anxieties, and apprehensions are channeled toward the situation which holds greatest immediate threat to his self-preservation or general well-being. A mother may sleep soundly as noisy freight trains pass her home yet quickly awaken at the slightest whim-

* JamesAllanMatte@mattepolygraph.com

per of her baby. This illustrates the ability within us to tune in that which may indicate trouble or danger by having our sense organs and attention set for a particular stimulus and oriented in a manner that will dampen any stimulus of lesser importance. The guilty suspect has his sense organs and attention set for that question which he feels will jeopardize his well-being." Backster continues that "By understanding anti-climax dampening effect we have many new avenues open to us for technique advancement." He concludes by stating that "we can carefully introduce certain questions or other stigma into a test structure which will be strong enough to be of concern to the innocent suspect, but will be strictly anticlimactic to the guilty suspect who is focused on the more intense relevant questions."

Backster theorized that by structuring a test which would offer both the Innocent and the Guilty examinee a narrow focal corridor comprising three exclusive control questions (Green Zone) and two neighboring relevant questions (Red Zone) dealing with the same act within the same issue, the responses to one Zone would ideally dampen potential responses to the other zone. Hence a Guilty examinee should respond strongly to both of the relevant questions, which should in turn dampen out any potential responses to the neighboring control questions due to the anti-climax dampening effect. It is important to note that the Backster Zone Comparison Technique is a true Single-Issue test in that both relevant questions deal with the same act within the same issue, whereas some other Zone Comparison Techniques refer to their technique as single-issue tests because their test questions deal with the same issue but not necessarily the same act, i.e. the inclusion of two separate sex acts allegedly perpetrated on the same victim during the same period, or the inclusion of an evidence-connecting question and/or a knowledge question. The reasoning given is that if the examinee is lying to one of the relevant questions he/she will most likely be lying to the other relevant question; however, inasmuch as they are separate acts, the test is not a true single-issue test as defined by Backster. In order to laterally score the charts, that is to combine the vertical scores attained from each relevant/control question pair, both relevant questions must deal with the same act, thus the same relevant questions worded differently, which provides not only internal reliability, but also a narrow focal corridor[1] for the Guilty examinee to focus his/her psychological set (Matte, Grove 2001). Conversely, the Innocent examinee's psychological set should

[1] Relevant questions are distinct, specific, narrow in scope and void of mental exercise, thus offering the examinee a clear choice in his/her answer to the relevant questions.

ideally be focused on the broader focal corridor[2] offered by the neighboring control (comparison) questions which are designed to be removed from the period encompassed by the relevant questions with the use of time bars that purposely render the control (comparison) questions structurally less intense than the relevant questions.

By positioning the stronger exclusive control question which deals with the period closest to the event being tested and the examinee's activities as an adult, immediately before the first quantified relevant question the Innocent examinee is given a chance to react strongly to this control question, and because his/her psychological set is focused onto that control question, it has the potential of dampening out concern about the neighboring relevant question that follows it. In fact, the Innocent examinee's psychological set should be focused squarely on the control questions, which should dampen out any concern for the neighboring relevant test questions, unless the Innocent examinee has a fear of error regarding the target issue.

However, the Guilty examinee should find the two relevant questions dealing with the same act an immediate and serious threat to his/her well being, far greater than his/her probable-lie to the three exclusive control questions which deal with potential issues that are significantly removed with time bars from the current period encompassing the offense.

It is in this context that we examine Backster's chart interpretation rule, which requires that when there is a strong (maximum) response on a relevant question it is compared with the exclusive control question that has the least or no response on either side of it, because that control question is functioning as designed. Should there be an equally strong (maximum) response to the exclusive control question that precedes the strong (maximum) response to the relevant question, that control question is viewed as defective.[3] That defective

---

[2] Exclusive Control (Comparison) questions are broad in scope, which invites mental exercise, known to cause autonomic arousal. But it does not embrace the period covered by the relevant questions, thus being structurally less intense.

[3] Abrams (1997) stated in his study that "He (Backster), however, strongly rejects the term weaker control indicating that it is actually a defective control. His method is to score to the defective (weaker) control…" Abrams admitted to this author (1999) that his above statement was semantically incorrect, and acknowledged that Backster considers the control question that elicits the *strongest* reaction is in fact the *defective* control versus the control question that elicits little or no reaction to be the effective control when there is also a strong response to the neighboring relevant question. This misunderstanding of Backster's concept is shared by a multitude of polygraphists and researchers.

control question may have time bars that are too close to the relevant issue and embrace unknown serious crimes, i.e. serial rapist, murderer, robber, etc., or a deliberate countermeasure was employed on the control question(s) to cause an inconclusive or false negative.

Backster "Either-Or" Rule

To arrive at an interim spot analysis tracing determination of (+2) or (–2) there must be a significant and timely tracing reaction in either the red zone or the green zone being compared (Backster 1989, Matte 2007).

(a) If the red zone indicates a lack-of-reaction it should be compared with the neighboring green zone containing the larger timely reaction.
(b) 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.

* * *

Sub-paragraph (b) of Backster's "Either-Or" Rule comprises the primary element studied in this field research.

The relevant questions deal with *known* acts that are specific in nature and composition. Therefore a strong (maximum) response to a relevant question, absent an artifact or non-addressed variable, can safely be interpreted as being a serious threat to the well-being of the examinee, and due to the psychological structure of the test, an inference of deception can be made to that relevant question. That there is an equally strong response to its neighboring exclusive control question does not alleviate the fact that there is an equally strong response to the relevant test question which embodies the reason for the examination. Thus a strong reaction to a relevant question cannot be ignored because there is an equally strong reaction to its neighboring control question. Furthermore, in the aforesaid situation, the strong reaction to the control question must be viewed as defective and remedied in accordance with Backster's Tri-Zone Reaction Combinations (Backster 1962, 1963, 1979). Hence, a relevant question that elicits a strong response is compared against the exclusive control question that expectedly elicits the least or no response.

However, Backster's rule articulated above has been criticized as being biased against the Innocent examinee (Raskin, 1986; Bell, Raskin, Honts, Kircher, 1999). Some techniques, such as the DoDPI Zone Comparison Technique, USACID Zone Comparison Technique, and USAF Zone Comparison Technique, compare the relevant question that elicits a strong response to the control question that elicits the greatest response. The Utah Zone Comparison Technique, which previously (Raskin 1979, Honts 1996) compared the relevant question(s) with the control question preceding it, now also compares the relevant question(s) against the neighboring control question that elicited the strongest response (Bell, Raskin, Honts, Kircher, 1999), but this change may be due to the fact that the control questions in the Utah ZCT are now directed-lies which are inherently weak (Matte 1998, 1999).

While a comparison of a strong response on a relevant question to the neighboring control question that manifests the least response appears to be selective, a comparison of a strong response on a relevant question to the neighboring control question that manifests the greatest response is equally selective. Hence a non-selective approach would require that each relevant question be compared with the neighboring control question preceding it, thus each control question would be paired with the relevant question immediately following it. Control questions usually precede the relevant questions to enable the structurally weaker control question to dampen potential responses to the relevant questions by the Innocent (as later verified) examinee. An example of this is the Matte Quadri-Track Zone Comparison Technique (Matte, 1978, 1980, 1989, 1996), which pairs each control-relevant question into a *Track*, and the polygraphist cannot jump the track (go outside of that track) to compare the relevant question with another control question located in another track. However, as is customary with both the Backster and Matte ZCT, each quantified relevant question is rotated in position with each chart conducted so that each relevant question is eventually compared with each control question in the same test.

However, several studies have been conducted in an attempt to resolve the issue of whether a comparison of a relevant question that displays a strong response should be made against the control question that elicited the strongest response, the control question that elicited the weakest response, or the control question preceding that relevant question for a pairing of control v. relevant questions. An examination of the literature on that subject is summarized below.

A Validation and Reliability Study of Counterintelligence Screening Test (CST) conducted by Gordon H. Barland (1991) revealed that when the relevant questions were compared to the single greatest control question reaction, it was very effective in identifying truthful subjects and truthful questions (89%); however "the greatest control method was unable to detect either the deceptive subjects or the deceptive questions at greater than chance levels. Therefore it should not be used in real-life situations unless future research is able to demonstrate that it is able to detect deception." However, the control questions used in Barland's study were Directed Lies, which have been shown (Matte, Reuss 1999) to be prone to false negatives. Barland's study did not address the effectiveness of comparison with the weakest control or the straight paired, but it did demonstrate the ineffectiveness of comparing the relevant question to the greatest control question reaction when using the CST.

Crowe, Chimarys and Schwartz (1988) conducted an analog study employing the General Question Test (GQT) format, which they labeled *a control question technique*. In this study, they compared the strongest control reaction to the reaction for each relevant question, and they also compared the weakest control reaction to the reaction for each relevant question. The results revealed that "Discounting inconclusive results, the strongest control was correct in six of seven deceptives (86%) and seven of seven nondeceptives (100%), for a total of 13 of 15 (93%). The strongest control resulted in an inconclusive rate of 16 of 30 (53%). The weakest control was correct in 17 of 17 deceptives (100%), but was correct for none of the nine nondeceptives (0%), for a total of 17 of 26 (65%). The weak control resulted in four inconclusive decisions (13%). However, the GQT traditionally and in this study employed disguised control questions, not Backster earlier-in-life control questions. In fact, only two Disguised Control Questions, i.e. "Do you intend to lie to any of the questions on this test?" and "Have you lied to me in any way since we have been talking today?" were used against at least double the number of relevant questions (Schwartz, 1999). This study did not replicate the format and psychological structure of the Zone Comparison Technique, hence its results cannot be factored into the analysis of the value of *greatest* versus *weakest* control question comparison when employed in the single-issue zone comparison technique such as the Backster ZCT.

"An Analysis of Zone Charts by Various Pairings of Control and Relevant Questions" conducted by Michael Koll (1979) involved the pairing of the Greater Control Question with neighboring relevant questions; the Greatest Control Question compared with all relevant questions, and the Straight Paired, which

compared each relevant question with the control question preceding it. Sixty charts equally divided into truthful and deceptive subjects from the files of the U.S. Army Polygraph School were evaluated. The conclusions of this study revealed that "The results of this study support the position that the greatest degree of accuracy with the lowest inconclusive rate for subjects in a mock crime can be attained in the Modified Zone Comparison Test with the Straight pairing CQ/RQ comparative evaluations." The Straight Pairs for the Truthful subjects, employing "the standard –/+ 6 for a decision, the straight-pair evaluations of two subjects (out of ten) were called inconclusive when they were actually truthful. No wrong calls were made." For the Deceptive subjects, there was one subject (out of ten) that was inconclusive when the straight pair evaluation was used, and no wrong calls. A comparison of validity and utility for the Greater control question, the Greatest control question and the Straight Pairs or paired control/relevant questions revealed that the straight pairs had a 100% validity with 85% utility and no errors; the Greater control question had a 100% validity with 75% utility and no errors, and the Greatest control question had a 93% validity with 70% utility and 5% error rate. The results of this study show that the Greatest control question had the least validity, accuracy and utility. The study favors the use of the straight pairs which is nonselective in its approach.

Capps and Ansley (1992) conducted a study where examiners were required to blind score forty sets of confirmed zone comparison charts that employed the DoDPI test format. Thirty-one of the examinations employed the zone comparison taught at DoDPI; the other nine used basically the same test format but replaced the symptomatic questions with irrelevant questions. Only the first relevant question within each test was used for comparison to the control questions (Abrams 1997, 1999). The examiners were instructed to "score against the strong control and separately score against the weak control." A fixed threshold of +/– 6 was used in the decision-making process, with lower scores classified as inconclusive. The results of Capps and Ansley's study revealed that "excluding inconclusives, the reviewers were correct in their analysis in 97.3% of the cases using the strong control procedure." An analysis of the same charts by the same reviewers using the weak control procedure revealed that "Excluding inconclusives, the reviewers were correct in 89.9% of the cases using the weak control scoring method." However, Capps and Ansley caution that "In fairness to Backster this approach should not be taken unless all the rules that are involved in the chart interpretation technique instituted by Backster are employed." Capps and Ansley further caution that "Our finding that use of the strong control is more accurate and reduces inconclusives is in

concert with the findings of others, but we are not yet prepared to recommend a change in the Backster system. We do believe our findings are sufficient to justify a full study on this aspect of scoring zone comparison charts."

The cautionary remarks by Capps and Ansley are well justified when we examine the methodology used in their research. Nine of the aforesaid examinations used in their study replaced the symptomatic questions with irrelevant questions. A study by Capps, Knill, Evans (1993) revealed that Symptomatic questions reduce inconclusives by two thirds, exactly as Cleve Backster had predicted. Furthermore, most tests examined used control questions that employed similar time bars within each test (Ansley 1999), as opposed to the Backster ZCT, which uses separate time bars within each test that completely divorces the time frame occupied by each control within the same test, i.e. 46. "Between the ages of 18 and 25, Do you remember ever engaging in an unnatural sex act?" and 47. "During the first 18 years of your life, do you remember ever doing anything sexually that you're ashamed of?" The Backster exclusive control questions, while in the same category, are independent of each other, thus less vulnerable to habituation and more likely to elicit individual attention and psychological set. The Capps and Ansley study employed a fixed threshold of –+6, which required the structurally weaker control questions to attain the same minimum score as the structurally stronger relevant questions. However, the Backster system requires a substantially lower score (threshold) for the Truthful, i.e. +3 for chart 1; +5 for 2 charts; +7 for 3 charts, versus the Deceptive which requires –5 for chart 1; –9 for 2 charts, –13 for 3 charts. Thus Backster's scoring system accepts lower scores (threshold) from the structurally weaker control questions to arrive at a Truthful decision. The use of only the first relevant question for comparison with the control questions by Capps and Ansley which was replicated by Abrams in his study skews the overall tally of the scores inasmuch as it omits the same comparison data from the other relevant question. Abrams (1997) points out that most of the confirmed cases used in the Capps & Ansley study were based on tests in which the stronger control procedure was used, and that "this could have given some advantage to the stronger control approach in their study, resulting in higher levels of accuracy for this procedure." Thus the results of the Capps & Ansley study "can only be generalized to tests administered in the same manner as in these studies."

Stanley Abrams (1997) conducted a study of twenty confirmed truthful and twenty confirmed deceptive charts which employed a single-issue zone comparison test format. Each examination consisted of two relevant questions and three control questions. Only the first relevant question was used for compari-

son in this study. Furthermore, Abrams used the fixed threshold of +–6 in his decision-making process. Abrams used the preceding control as the standard for ground truth in the evaluation of the effectiveness of comparisons to the stronger and weaker control questions. The results of Abrams' study revealed that, for the deceptive subjects, the stronger control was as accurate as the weaker control, to wit: 100 percent in each case; but the stronger control had a 10 percent inconclusive rate. For the truthful subjects, the stronger control attained a 100 percent accuracy with no inconclusives, while the weaker control attained a 92 percent accuracy with a 40 percent inconclusive rate. However, Abrams issued a cautionary remark regarding the results of the stronger control for the truthful, stating that "Since the preceding control was the standard used for ground truth, the complete accuracy with no inconclusives is misleading." Abrams recognized the limitations of his study using only the first relevant question for comparison with the various control questions (Abrams 1999), in that it skews the overall tally of the scores inasmuch as it omits the same comparison data from the other relevant question. As in the Capps and Ansley study, Abrams opined that his study can only be generalized to tests administered in the same manner as in his study.

Hence none of the aforementioned studies provide evidence that the Backster theory and rules regarding the comparison of relevant versus control questions are invalid and/or biased against the truthful (as later verified) examinee when specifically applied to the Backster "You Phase" Zone Comparison Technique.

Thus, of all of the aforementioned studies, only the Koll (1979) study comes close to replicating the Backster Zone Comparison format, but does not address comparison of a strong relevant question with the control question that elicited the least or no response.[4] However, for the same reason that the system of comparing the relevant question that elicited a strong reaction to the weakest control question can be criticized as being selective and biased in its comparison process against the truthful (as later verified) examinee, the system of comparing the same relevant question to the strongest control question can also be criticized as being selective and biased in its comparison process towards (for) the deceptive (as later verified) examinee.

---

[4] The Koll study, which most closely replicates the format of the Backster Zone Comparison Technique, supports the Matte Quadri-Track Zone Comparison Technique's use of tracks which pairs each control question with the relevant question that follows it, isolating them inside a track which prohibits jumping (leaving) the track for comparison with another control question. Inasmuch as each relevant question is switched in position with each chart conducted, each relevant question is eventually compared with each control question within the same test.

In the interpretation of physiological responses to relevant versus control questions, it must be recognized that a strong response to a control question may be a genuine response from an innocent (as later verified) examinee, or a fabricated response from a guilty (as later verified) examinee that employed a countermeasure. A third possibility is a strong response from a guilty (as later verified) examinee who finds the contents of the control question equally threatening due to a defect in its construction.

The initial focus in the interpretation of physiological responses must be on the relevant questions, because they are the reason for the examination, and are least likely to be faulty in their construction due to the required specificity, accuracy and completeness of the case information from which the relevant questions are formulated. On the other hand, the control questions are necessarily broad in nature and embrace unknown events which can include the commission of serious offenses, some of which may offer an equal or greater threat to the guilty (as later verified) examinee than the relevant questions.

Therefore, Backster's chart interpretation rule, which assumes that a relevant question which elicits a strong response is functioning as designed due its ideal formulation, and a neighboring control question which also elicits a strong response is defective, appears to have theoretical and face validity. None of the above cited research studies offer valid, applicable results which contradict Backster's concept and rule.

The Backster Zone Comparison Technique employs an increasing threshold for both the Deceptive and the Truthful examinee. Other techniques (excluding the Matte Quadri-Track ZCT) employ a fixed threshold, usually a plus or minus 6, regardless of the number of charts conducted. With the Backster ZCT, the score required to attain a definite decision (threshold) of Truthfulness is significantly less than that required to attain a definite decision of Deception. For example, the Backster system requires a –5 for the first chart, –9 for 2 charts, and –13 for 3 charts in order to render a decision of Deception. Whereas a +3 for the first chart, +5 for 2 charts and +7 for 3 charts is required in order to render a Truthful decision. Hence the Innocent (as later verified) examinee is provided with a generously lower score threshold to attain a decision of truthfulness.

It therefore became apparent that, in order to determine the validity of Backster's "Either-Or" Rule, which dictates that when using the Backster "You Phase" Zone Comparison Technique, a relevant question that elicits a strong

response must be compared to the control question that elicited the least or no response, a field research study of actual cases that religiously employed the Backster ZCT with all of its rules needed to be conducted.

The cases that would have produced a strong response to the relevant questions should expectedly come from verified guilty examinees. Verified innocent cases would not be expected to produce strong responses to the relevant questions unless false positive physiological results occurred. The Backster ZCT requires that "If the red zone indicates a lack-of-reaction it should be compared with the neighboring green zone containing the larger timely reaction." Hence, regardless of the number of confirmed truthful cases found, the fact that they were confirmed truthful required that the control questions elicit and show a significantly greater response than its neighboring relevant questions, hence the Backster method as well as the others would compare the relevant question that elicit the least or no response to the control question that elicited the greatest response, thus negating our attempt to determine the validity of Backster's Either-Or rule when there is the presence of a strong reaction to the relevant question(s). In this instance, there were only five confirmed truthful cases found for the aforesaid period. Strong responses to the relevant questions from which a comparison could be made with the three methods of relevant/control comparison could only be expectedly found in verified guilty cases. In the instant study, all confirmed cases for the period 1 January 1998 to 1 September 1999 were reviewed and no confirmed false positives or negatives were found.

There are two other methods of relevant/control question comparison. One requires that a relevant question which elicits a strong response be compared with the control question that elicited the greatest response. The other requires that a relevant question, regardless of the strength of its response, be compared with the control question that preceded it. It therefore seemed logical to test all three methods of relevant/control question comparisons to determine the accuracy of each of the three procedures.

Method

The Commonwealth of Virginia Department of State Police assigned six experienced polygraphists under the leadership of Special Agent Gary M. Jenkins to review and score the polygraph charts of all confirmed guilty cases acquired from seven field offices that were conducted during the period 1 January 1998

to 1 September 1999. One hundred forty-six criminal cases, including the polygraph charts and related notepacks, were submitted for review. Twenty-three of those cases were rejected because they were not Backster "You Phase" Single-Issue Zone Comparison tests. Hence, 123 confirmed guilty cases comprising a total of 270 polygraph charts were used in this study. There were no errors found to have been committed by the original polygraphists who conducted the 123 confirmed guilty psychophysiological veracity (PV) examinations administered during the period 1 January 1998 to 1 September 1999. Test results were confirmed by confession (115) and conviction (8).[5] There were only five recorded confirmed truthful cases found for the aforesaid period and none of them contained a strong response to the relevant questions, hence were not candidates for the three comparisons in this study which required that the relevant question(s) elicited a strong response. In addition, eleven confirmed truthful cases were provided by Tuvia Shurany (2003), who had a team of three polygraphists analyze and score the charts in accordance with the requirements of this study. These 11 confirmed truthful cases revealed only 5 spots containing a significant deceptive score out of a total of 174 spots, of which 169 contained no significant reactions, and those 5 spots did not affect the accuracy of the results which were confirmed as truthful. Consistent with the previous 5 confirmed truthful cases that were not viable candidates for the three comparisons in this study, the 11 confirmed truthful cases from Shurany were also not viable candidates. It becomes quite apparent that in order to test the accuracy of the three methods of scoring charts when there is a strong response to the relevant question we use confirmed *deceptive* cases that are expected to produce significant responses to the relevant question which will test Backster's "Either-Or" rule in comparison with the other two methods of scoring charts.

---

[5] A study by C. Ronald Huff, Arye Rattner and Edward Sagarin (1986) estimated that the rate of wrongful convictions (irrespective of polygraph evidence) in the United States is one-half percent (0.5%). In a study by Gary D. Light and John R. Schwartz (1993 &1999) the authors point out that the argument against field studies of PV examinations that use confessions as ground truth is not substantiated by the results of their study which employed confessions as ground truth but the results were also confirmed by the findings of the other disciplines used in that same study. The authors state "While there can be no question that examinations verified by confession are a unique subset of PDD examinations, this study indicates that this bias has a minimal impact, and confession-based samples would accurately reflect the overall population." They further indicated that their assertion was further substantiated by another study (Mason 1988) in which 111 PV examinations were performed and ground truth was established by urinalysis. The validity of these PV examinations which were confirmed by biomedical tests was in excess of 95% and if confessions are used in conjunction with urinalysis examinations, the accuracy of that confession subset rises to over 98%.

The Backster Zone Comparison Technique employed in each of the aforesaid 123 cases used two relevant questions, numbered 33 and 35 for comparison to their neighboring control questions number 46, 47 and 48. Each of the relevant questions was compared in each tracing (Pneumograph, Electrodermal, Cardiograph) with the control question that elicited (a) the least or no reaction, (b) the greatest reaction, (c) the preceding (previous) control question, and a score was acquired for each comparison in each tracing, which was recorded on a data sheet. Inasmuch as the Backster ZCT is a true Single-Issue test wherein the two relevant test questions pertain to the same act, the scores from each relevant question can and were horizontally tallied for a total score in each chart quantified. It should be noted that each of the three methods of scoring was conducted by a different polygraphist blind to ground truth to avoid being influenced by the results of the other methods used in the same case (Jenkins 2002, Gibbs 2002).

Analysis and tabulation of the data indicated the following:

	Control with Least or No Reaction	**Control with Greatest Reaction**	**Previous Control (Control to Left)**
Total Inconclusives:	33 = 12.1%	95 = 35.3%	63 = 23.5%
Total False Negatives:	None	2 = 0.7%	None
Total Scores:	–1488	–668	–1100
Number of Charts:	270	269	268
Mean Number of Charts:	2.20	2.19	2.18
Average Score per Chart:	–5.51	–2.48	4.10
Average Score Per Case:	–12.12	–5.43	–8.94

The 123 cases representing 270 polygraph charts averaged 2.2 charts per case. Two false negatives would have occurred using the Greatest Reaction Control (GRC), and the GRC produced the greatest number of Inconclusives at 35.3%. The comparison of the relevant questions with the control question that elicited the least or no reaction produced the least number of inconclusives and with no errors, followed by the comparison of the relevant questions with the control question that preceded the particular relevant question with no errors.

The results of this field research study supports Backster's "Either-Or" Rule of comparison of the relevant question that elicits a strong reaction with the control question that elicits the least or no reaction, and refutes the contention that

its practice makes the Backster Zone Comparison Technique biased against the Innocent examinee. Furthermore, three field studies (Matte, Reuss 1989a, 1989b; Mangan, Armitage, Adams 2008; and Shurany, Stein, Brand 2009) on the Quadri-Track Zone Comparison Technique which employs Backster's "Either-Or" Rule (Backster 1989, Matte 2007), supports its validity.

In conclusion**,** this field study reveals that of the three methods of relevant/control question comparison, the comparison of a relevant question that elicited a strong response to the control question that elicited the least or no response was the most effective in terms of accuracy with the lowest inconclusive rate, which supports Backster's theoretical concept and chart interpretation rules, and refutes the notion that its methodology is biased against the innocent.

References

Abrams S. (1997), *Stronger v. Weaker v. Preceding Control*, Journal of the American Association of Police Polygraphists, 1997.

Abrams S. (1999, May 11), *Telephone Conversation with J. A. Matte Regarding Study Conducted by Abrams Entitled Stronger v. Weaker v. Preceding Control*, published in Journal of Association of Police Polygraphists, 1997.

Ansley N. (1999, May 11), Telephone conversation with J. A. Matte regarding Study conducted by Capps, M. H., and Ansley, N., entitled *Strong Control versus Weak Control*, published in Polygraph, 21 (4), p. 341–439.

Backster C. (1974), *Anticlimax Dampening Concept*, Polygraph, 3 (1), p. 48–51.

Backster C. (1964), *Spot Analysis vs. Relevant-Irrelevant Technique*, Law and Order, 12 (2), p. 3l–32.

Backster C. (1962), *Tri-Zone Reaction combinations: Backster Zone Comparison Test*, Backster School of Lie Detection, New York, NY.

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

Backster C. (1989, September), *Backster Zone Comparison Technique: Chart Analysis Rules*, Backster School of Lie Detection, San Diego, California.

Barland G. H. (1981, May 12), A Validation and Reliability Study of Counterintelligence Screening Test. Security Support Battalion, 902d Military Intelligence Group, Fort George G. Meade, MD. Unpublished manuscript.

Bell B. G., Raskin, D. C., Honts C. R., Kircher J. C. (1999), *The Utah Numerical Scoring System*, Polygraph, 28 (1), p. 1–9.

Capps M. H., Ansley N. (1992), *Numerical Scoring of Polygraph Charts: What Examiners Really do*, Polygraph, 21 (4), p. 264–320.

Capps M. H., Knill B. L., Evans R. K. (1993), *Effectiveness of the Symptomatic Questions*, Polygraph, 22 (4), p. 285–298.

Crowe M. J., Chimarys M., Schwartz J. (1988), *The GQT Polygraph Test: Scoring and Validity*, Poster presentation at the 96th annual convention of the American Psychological Association.

Gibbs G. W. (2002, Jun 4), Personal communication with J.A. Matte.

Honts C. R. (1996, February 9), Letter of reply to letter of inquiry by J. A. Matte dated 13 Jan 96.

Huff C. R., Rattner A., Sagarin E. (1986), *Guilty Until Proved Innocent: Wrongful Conviction and Public Policy*, Crime & Delinquency, 52 (4), p. 518–544.

Jenkins G. (2002, May 17), Personal communication with J. A. Matte.

Koll M. (1979), *Analysis of Zone Charts by Various Pairings of Control and Relevant Questions*, Polygraph, 8 (2), p. 154–160.

Light G. D., Schwartz J. R. (1993, March), *The Relative Utility of the Forensic Disciplines*, Department of Defense Polygraph Institute, Ft McClellan, AL.

Light G. D., Schwartz J. R. (1999), *The Relative Utility of the Forensic Disciplines*, Polygraph, 28 (3), p. 240–258.

Mangan D. J., Armitage T. E., Adams G. C. (2008), *A Field Study on the Validity of the Quadri-Track Zone Comparison Technique*, Physiology & Behavior, 95, p. 17–23.

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

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

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

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

Matte J. A. (1998), An analysis of the psychodynamics of the directed-lie control question.

Matte J. A., Reuss R. M. (1989a), *A Field Validation Study of the Quadri-Zone Comparison Technique*, Polygraph, 18 (4), p. 187–202.

Matte J. A., Reuss R. M. (1989b), *Validation Study on the Quadri-zone Comparison Technique, Dissertation-Research Abstract*, LD 01452, Vol. 1502, 1989, Proquest Information & Learning (UMI).

Matte J. A., Reuss R. M. (1999), *Validation of Potential Response Elements in the Directed Lie Control Question*, Polygraph, 28 (2), p. 124–142.

Matte J. A., Grove R. N. (2001), *Psychological Set: It origin, Theory and Application*, In press in Polygraph, Journal of the American Polygraph Association.

Matte J. A. (2007), *Psychological Structure and Theoretical Concept of the Backster Zone Comparison Technique and the Quadri-Track Zone Comparison Technique*, Polygraph, 36 (2), p. 84–90.

Raskin D. C. (1979), *Nevada Supreme Court Rules that Stipulated Evidence is Admissible: Background to Case*, Polygraph Law Reporter, 2 (1), p. 10–14.

Raskin D. C. (1986), *The Polygraph in 1986: Scientific, Political, and Legal Issues Surrounding Application and Acceptance of Polygraph Evidence*, Utah Law Review, 1986, p. 29–74.

Schwartz J. R. (1999, April 29), Telephone conversation with J. A. Matte regarding the General Question Test format in the Crowe, Chimarys and Schwartz 1988 study.

Shurany T. (2003, June 6), Eleven data worksheets containing three types of scoring methods on confirmed truthful cases provided to J. A. Matte by mail.

Shurany T., Stein E., Brand E. (2009), *A Field Study on the Validity of the Quadri-Track Zone Comparison Technique*, European Polygraph, Vol. 1, No. 1 (7), p. 5–24.

The author wishes to express his sincere appreciation to the Commonwealth of Virginia Department of State Police, and specifically, Colonel M. Wayne Huggins, Superintendent of Virginia State Police, Lt. Colonel Darrel E. Stilwell, Director of Bureau of Criminal Investigation, Virginia State Police, Captain Terry A. Bowes, Commander, General Investigations Division, Virginia State Police, and Assistant Special Agent in Charge Gary M. Jenkins, Department of Virginia State Police Polygraph Quality Control Officer, who organized and supervised the acquisition of the data used in this field study. The author also wishes to express his sincere appreciation to the following Special Agents of the Virginia State Police who participated in this field study: ASAC George W. Gibbs, SSA Roy L. Kyle, SA Stanley P. Ogden, SA John E. McDowell, and SA Steven T. Oliver. The author further wishes to express his sincere appreciation to Cleve Backster and Donald Krapohl for their input relative to the methodology used in this field study, and Dr. Stan Abrams and John Schwartz for their input regarding the methodology used in their reported studies. Lastly but not least, the author wishes to thank Tuvia Shurany for having three examiners score confirmed truthful charts in accordance with the requirements of this research study.

EUROPEAN
POLYGRAPH

Volume 4 • 2010 • Number 2 (12)

Tuvia Shurany*
Liecatcher Polygraph Series
Jerusalem, ISRAEL
Fabiola Chaves
Polygraph Examiner at G4S Security Services
COSTA RICA

The Integrated Zone Comparison Technique and ASIT PolySuite Algorithm: A Field Validity Study

Abstract

This field study tested the validity of the Integrated Zone Comparison Technique[1] (IZCT) designed for specific issue testing and the ASIT PolySuite Algorithm for data analysis in a private commercial environment between April and December, 2009, at the G4S polygraph unit in Costa Rica. During this time period 27 cases were chosen to be tested with the IZCT. Out of these 27 cases, 21 were solved by confession. The 27 cases had a total of 113 suspects. Out of the 113 tests, 84 were confirmed results. Of these, there were 44 confirmed deceptive examinees and 40 confirmed innocent

* tuvia@liecatcher.com

[1] Gordon, N. *et al.*, *A Field Study of the Integrated Zone Comparison Technique.* Polygraph, 2000, Vol. 29, No. 3, pp. 220–225.

examinees. Data analysis was performed with the Academy for Scientific Investigative Training's ASIT PolySuite©, which is an examiner-controlled computerized algorithm, using the Horizontal Scoring System[2] (HSS), as well as an experienced examiners analysis using a 3-Point Scale[3]. IZCT, using ASIT PolySuite, had an overall accuracy of 92.9% with Inconclusives, and 98.73% accuracy excluding them. Manual 3-point scoring had an overall accuracy of 91.7% with Inconclusives, and 98.71% excluding them.

Running head: Integrated Zone Comparison Technique

Key Words: Integrated Zone Comparison Technique, Data Analysis, ASIT PolySuite Algorithm, Horizontal Scoring System, 3-Point Manual Scoring System, Validity

This field study is the third published research study[4] on the Integrated Zone Comparison Technique (IZCT) and the fourth study that included the Horizontal Scoring System (HSS). The theory and philosophy of the IZCT was first published in 1996, in the textbook *Forensic Psychophysiology using the Polygraph*[5].

The IZCT has been taught at the Academy for Scientific Investigative Training since 1987[6]. It is currently being used in the fields of law enforcement, intelligence, and private security in numerous countries around the world. It is a modification of the Backster Zone Comparison Technique[7] format, in a structure that closely resembles the zone technique validated at the University of Utah[8]. It is a flexible technique format, allowing it to be used for Single-issue, Multi-faceted and Multi-issue investigations.

IZCT format is a thirteen-question test consisting of two weak relevant questions (sacrifice relevant, countermeasure indicator), three flexible relevant

---

[2] Krapohl, D., *A Comparison of 3- and 7-position Scoring Scales with Laboratory Data*.

[3] Gordon, N. *et al.*, *A Field Study of the Integrated Zone Comparison Technique.* Polygraph,2000, Vol. 29, No. 3, pp. 220–225.

[4] Matte, J., *Forensic Psychophysiology using the Polygraph. JAM Publications,1996, Buffalo, NY.*

[5] Gordon, N. *et al.*, *A Field Study of the Integrated Zone Comparison Technique.* Polygraph, 2000, Vol. 29, No. 3, pp. 220–225.

[6] Backster, C., *Technique fundamentals of the Tri-Zone Polygraph Test.* New York, Backster Research Foundation, 1969.

[7] Matte, J., *Forensic Psychophysiology; Use of the Polygraph. JAM Publications, 1996, Buffalo, NY.*

[8] Gordon, N., *The Academy for Scientific Investigative Training's Horizontal Scoring System and Examiner's Algorithm for Chart Interpretation.* Polygraph, 1999, Vol. 28, No. 1.

questions, three probable lie comparison questions, one symptomatic question, and four irrelevant questions.

1.	Irrelevant:	Is today Sunday? (No)
2.	Symptomatic:	Do you understand I will only ask the questions I reviewed?
3.	Weak Relevant:	(Sacrifice) Do you intend to lie to any test question?
4.	Irrelevant:	Is today an actual day? (Yes)
5.	Exclusive Comparison:	During the first __ years of your life, …….?
6.	Flexible Relevant:	Primary or secondary relevant question, depending on type and facts of case
7.	Irrelevant:	Right now are you in the US? (Yes)
8.	Inclusive Comparison:	In your entire life did you ever ……..?
9.	Flexible Relevant:	Primary or secondary relevant question, depending on type and facts of case
10.	Irrelevant:	Right now are you in Switzerland? (No)
11.	Comparison:	Exclusive or Inclusive
12.	Flexible Relevant:	Primary or secondary relevant question, depending on type and facts of case
13.	Weak Relevant:	(Countermeasure indicator) Have you deliberately done anything to try and beat this test?

The thirteen questions in the IZCT structure are then reviewed with the examinee in the following order: (1, 4, 7, 10), (6, 9, 12), (5, 8, 11), 13, 3 and 2. The examiner then explains how the polygraph instrument works and as an anti-countermeasure procedure during this presentation surreptitiously records the examinee's respiration on a separate chart.

The first IZCT chart is collected as a Silent Answer Test (SAT), which is cognitively stimulated by instructing the examinee that during the test he/she is to remain silent and listen to the questions carefully to make sure he/she is comfortable with them, understands them, and most importantly, does not remember anything they have not told the examiner about, since this is their last opportunity to make changes in questions before their verbal answers are recorded. The SAT questions are asked in the following sequence: 1, 2, 3, 4, C5, R6, C8, R9, C11, R12, 13. Irrelevant questions 7 and 10 are not used, unless they are needed to re-establish a norm during the examination, or used due to an artifact committed by the examinee during the examination.

The sequence for the second chart is: 10, 2, C5, R12, C8, R6, C11, R9, 3 ("Did you lie to any test question?"), 13. To focus the examinee on their zone of threat, when the examiner begins this chart the examinee is instructed to make sure he/she answers each question truthfully, since the charts will be numerically evaluated and lying to any question in the test, no matter what it is about, could cause them to fail the entire examination.

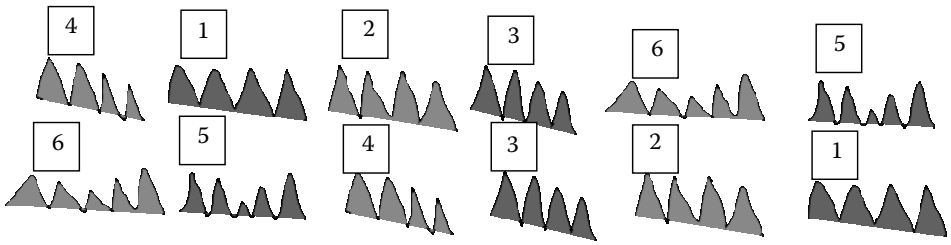
The third IZCT chart is administered with the relevant questions being asked before the comparison questions, and the relevant questions being rotated in the same manner. The sequence is: 1, 2, 3, R9, C5, R12, C8, R6, C11, 13. This allows for each relevant question to be asked paired with each comparison question once after three charts are administered.

If the need appears for additional data to be collected to reach a clear decision, or if there appear to be deliberate distortions, Chart 4 of the IZCT is used where all of the questions 1 to 13 are asked.
ASIT PolySuite combines the Horizontal Scoring System (HSS), with the Academy for Scientific Investigative Training's Algorithm for Manual Chart Interpretation[9] of polygraph data.

In the Horizontal Scoring System all four physiological channels of each relevant and comparison question are ranked horizontally from greatest to least, based on their significance in the chart. If the question format utilizes 3 comparison and 3 relevant questions, the most significant reaction in each channel is given a "6," and the least significant reaction is given a "1." If only 2 comparison and 2 relevant questions are used the channels are ranked from "4" to "1."

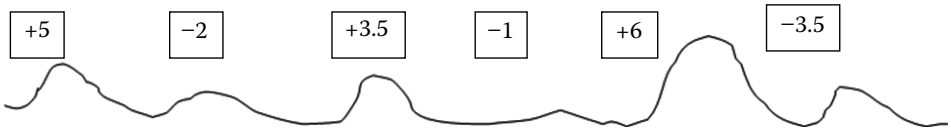
The below diagram shows Thoracic and Abdominal channels ranked horizontally from 6 to 1. Each question's abdominal and thoracic score is then averaged to ensure the pneumo tracings only account for 1/3 of the question's total score. Comparison question scores receive a positive numerical value and relevant question score receive a negative value.

[9] Honts, C., L. Driscoll, *An Evaluation and the Validity of Rank Order and Standard Numerical Scoring of Polygraph Charts.* Polygraph, 1987, Vol. 16, No. 4.

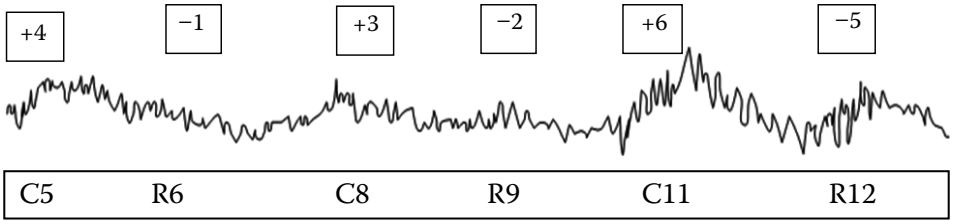


Avg. +5 **Avg. -3** **Avg. +3** **Avg. -3** **Avg. +4** **Avg. -3**

The electrodermal responses are ranked horizontally from 6 to 1. In case questions are equal in significance they are given the average of the rank positions they are competing for. In the electrodermal example below comparison question 8 and relevant question 12 are about equal. They are competing for the ranks of 4 and 3. Each question is given the average of those ranks, a 3.5.



The cardio responses are ranked horizontally from 6 to 1.



The average rank score for each question's pneumo channel can then be combined with the question's electrodermal and cardio ranks for a total question score. In the above example we have the following scores:

Average Pneumo					
+5	–3	+3	-3	+4	–3
EDA					
+5	–2	+3.5	–1	+6	–3.5
Cardio					
+4	–1	+3	–2	+5	–5
Total Question Scores					
C5	**R6**	**C8**	**R9**	**C11**	**R12**
+14	–6	+9.5	–6	+15	–11.5
SPOT SCORE:	**+8** (14–6)		+3.5 (9.5–6)		+3.5 (15–11.5)
SINGLE ISSUE CHART SCORE: **+15** (Combination of all Spot Scores)					

In the first two charts the rank of the relevant question is subtracted from the rank of the comparison preceding it. In the third chart we compare each relevant question with the comparison question that follows it [9].

The cut-offs using the Horizontal Scoring System (HSS) were established by Honts and Driscoll[10], who reported that accuracy for single issue tests, where three charts of data are collected consisting of 3 relevant and 3 comparison questions in each chart, would be above 90% when decisions of truth or deception were made using a ±13. Since that number reflected a total of 9 spot scores (13/9), decisions for Spot Scores for 3 charts of data are a $\pm$4.5. When four charts of data are analyzed single-issue cut-offs are a $\pm$18, and Spot Score cut-offs are a $\pm$6.

Method

In many studies critics say that the choosing of the examinees is selective, and therefore might influence the final result. As a result the first author decided

to be selective and chose to implement and run this study only in the 27 cases selected where all of the possible suspects involved were tested. These 27 cases had 113 examinees. Twenty-three of the cases involved theft, 2 involved falsifying receipts, 1 involved using a firearm without necessity, and 1 involved using a vehicle without permission.
In all of the examinations a multi-faceted type IZCT test format was used, having a primary relevant and two secondary relevant questions.

All data were first analyzed using the Academy's ASIT PolySuite, with cut-offs of a ±1.5 for each relevant question, for each chart administered. The spot cut-off for three charts was ±4.5. If any of the spots reached the negative cut-off the test was determined as a deceptive result regardless of the score reached by the other two spots. If all three spots reach the positive cut-off the final call for this test was truthful. With any other combination the final call for that test was inconclusive.

Many examiners in the world consider the 3-point scale the easier and less subjective scoring system to use, and therefore the authors decided to validate the format with a standard 3-point scale and to check if there are any significant differences between the results of the Horizontal Scoring System and the traditional vertical 3-point scale when implemented using the IZCT format. The manual scoring using the 3-point scale employed spot cut-offs of a +3 or higher for truthfulness, and a -3 or lower for deception.

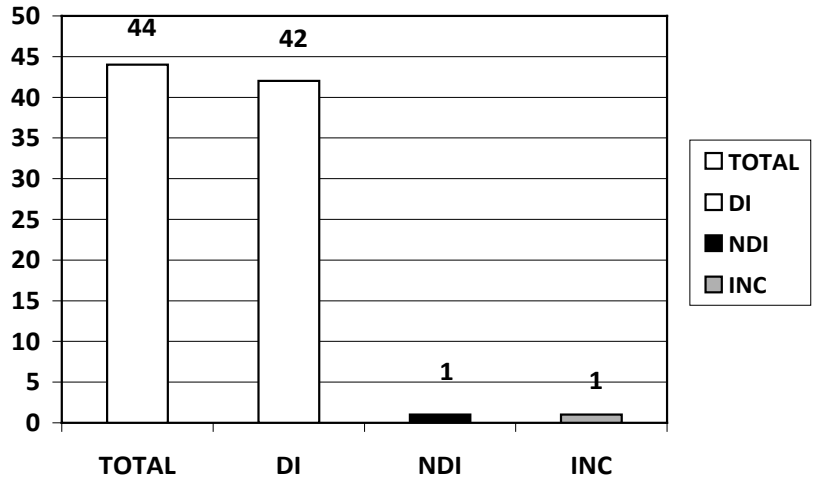
Results

A total of 27 cases were investigated using the IZCT during the period of April to December 2009. In these 27 cases there were 113 examinations conducted on all of the people who had any possibility of being involved. Forty-four of these suspects were deceptive, as later verified, and 40 were truthful, as later verified. All of the cases were verified by confession, and in some there was additional corroborating evidence of returned stolen items, or the showing of receipts for items paid for with stolen money.

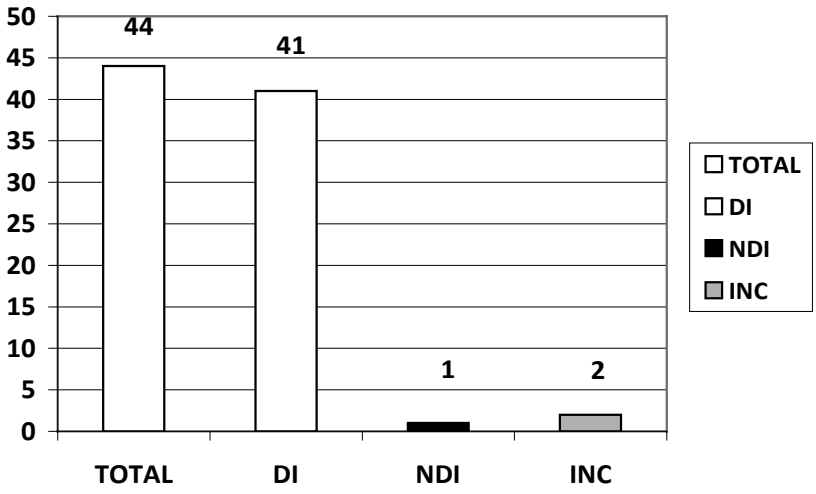
Of the 44 deceptive suspects, ASIT PolySuite correctly identified 43, had 1 False/Negative, and 1 Inconclusive outcome. Accuracy was 95.45% with Inconclusives and 97.72% without them. Of the 40 truthful suspects, ASIT PolySuite correctly identified 36 and had 4 Inconclusives. Accuracy was 90% with Inconclusives and 100% without them. There were no False/Positives.

Of the 44 deceptive suspects, the examiners using the 3-point scale correctly identified 41, had 1 False/Negative, and 2 Inconclusive outcomes. Accuracy was 93.18% with the Inconclusive, and 97.72% without it. Of the 40 truthful suspects, the examiners using the 3-point scale correctly identified 36 and had 4 Inconclusives. Accuracy was 90% with Inconclusive and 100% without them. There were no False/Positives.

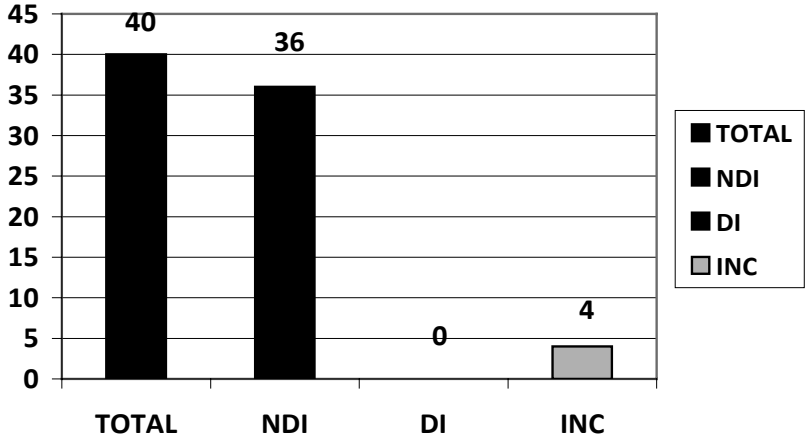
Accuracy of IZCT using ASIT PolySuite for DI Suspects:



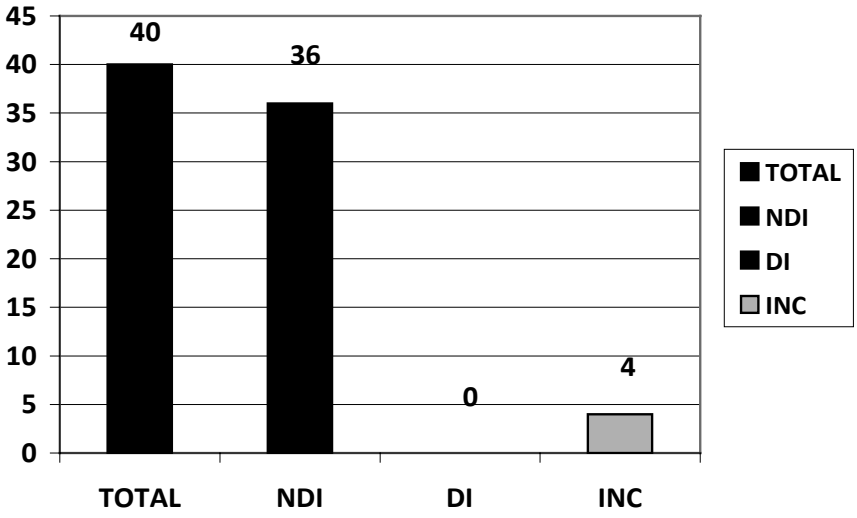
Accuracy of IZCT using 3-Point Scoring for DI Suspects:



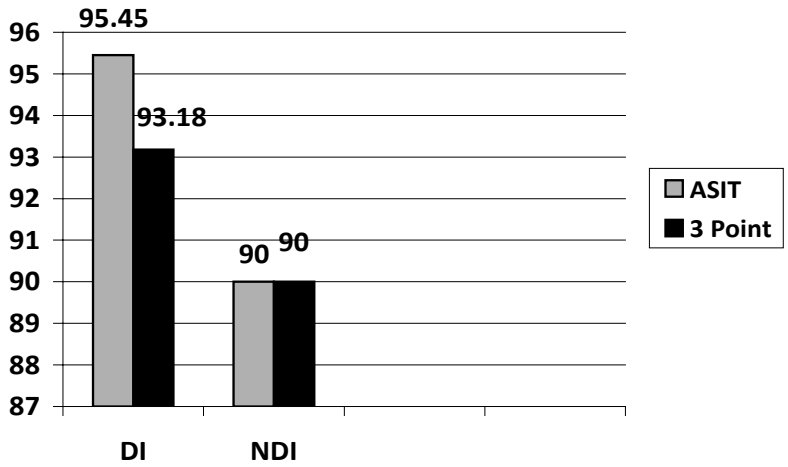
Accuracy of IZCT using ASIT PolySuite for NDI Suspects:



Accuracy of IZCT using 3-Point Scoring for NDI Suspects:



Overall Accuracy of IZCT using ASIT PolySuite vs. 3-Point Scoring



Conclusion

The result of this independent field validation study clearly demonstrates the efficacy of the IZCT for both truthful and deceptive suspects in multi-faceted law enforcement field investigations where the polygraph is employed as an investigative tool.

It should be noted that this study was consistent with the accuracy demonstrated in previous studies of the IZCT. All three studies performed to date have shown mean accuracy rates of truthful and deceptive examines at 90% or higher, which meets the industry standard of qualifying the IZCT to be used as a polygraph technique for both investigative and evidentiary cases.

EUROPEAN
POLYGRAPH

Volume 4 • 2010 • Number 2 (12)

Marek Leśniak*
Department of Criminalistics
Faculty of Law and Administration
University of Silesia in Katowice
POLAND
Magdalena Zubańska
Higher Police School in Szczytno
POLAND

A Comparison of Polygraph Examination Accuracy Rates Obtained Using the Seven-position Numerical Analysis Scale and the Objective Scoring System (A Study on the Polish Population)

Background

Different sorts of polygraph examinations have been conducted in Poland over the last sixty years (Krzyścin, 226, 227). It is curious that there are few empirical studies which concern the accuracy of such examinations in relation to the population of Poland. Hundreds of well-documented scientific studies on the accuracy of polygraph examination have been published all over the world. However, most concentrate on the American or Israeli population, so the problem of national differences should be taken into consideration, too. The nature of reactions to polygraph test questions usually justifies the transfer of

* marek.lesniak@us.edu.pl

the results of foreign studies onto Polish people. However, not having significant differences in this area should be confirmed by some empirical research. The current paradigm of empirical science demands replications of studies concerning the same variables – the most certain tool of scientific cognition is meta-analysis (Shaughnessy, Zechmeister, Zechmeister 2002, 271–273).

The objective of the present study is a comparison of psychophysiological detection of deception accuracy rates obtained by using the seven-position numerical analysis scale and the Objective Scoring System (this algorithm is sold together with Lafayette's polygraph software).

Both scoring systems rely on comparing the responses to the relevant question to responses to the appropriate comparison question (by individual physiological recorded parameters). The OSS was introduced by Donald Krapohl and Barry McManus in 1999 (Gordon, Mohamed, Faro, Platek, Ahmad, Williams, 253). Fifty per cent of decisions are generated from electro-dermal activity tracing, 25% from respiratory tracing and 25% from cardiovascular tracing. In contrast with a global evaluation (based on a form of a general impression) the numeral manual score allows adequately trained interpreters to reach extremely high reliability (Raskin, Honts, 18–19). Similarly to OSS features, during manually numerical scoring, changes in electro-dermal activity are more important than changes in respiration and cardiovascular activities. The reason for this is that the electro-dermal physiological parameter is only present where a ratio method is used solely in scoring the responses (Swinford, 18). Any expert's diagnosis based on numerical scoring may be checked easily by another expert. The polygraph technique used was the Utah Directed-Lie Test. This choice was driven both by the fact that this technique is practised in Poland and to avoid ethical (related to the characteristics of examinees) and technical difficulties in establishing effective comparison questions (Raskin, Honts, 24).

Method

Ultimately the research includes 43 participants (six polygraph outcomes were rejected on account of their very poor quality) – 23 men and 20 women. Individuals were recruited from students of the University of Silesia in Katowice. Some of them (called the "guilty subjects" group) had taken a 10-zloty note out of the cabinet in room number 3.53. Others (called the "innocent subjects" group) did not open the cabinet and did not see the note. Members of the

"guilty" group had been instructed to keep the note in their pocket until the time of the polygraph examination and during this examination too. The polygraph examinations were carried out by eight experts; they were certified as having completed skills training in this area. The experimenters assumed that the experts had adequate training requirements. The participants were tested with a Lafayette LX-4000 polygraph that monitored thoracic and abdominal breathing, electro-dermal activity and cardiovascular activity (only a standard blood pressure cuff). The experimenters did not conduct the polygraph examination; they measured the size of particular reactions on charts and interpreted the results. The interpreters got to know the actual role of each participant only when they had finished interpreting the polygraph outcomes. Polygraph examinations were conducted according to the rules that were described by David C. Raskin and Charles R. Honts (22–24). During the manual numerical evaluations, the interpreters followed Jimmie Swinford's instructions (1999) precisely. The OSS scoring system was utilized in version number 2.

The test questions for participants were as follows:

1. Are you aware of the fact that I will ask only the questions we have discussed?
2. Do you intend to answer untruthfully all of the questions about the taking of a ten-zloty note?
3. Is it Wednesday today?
4. Prior to 2009, did you ever take something that did not belong to you?
5. Did you take the 10-zloty note out of the cabinet from room number 3.53?
6. Did you see the 10-zloty note inside the cabinet in room number 3.53?
7. Prior to 2009, did you ever do anything that was dishonest or illegal?
8. Did you open the cabinet which the 10-zloty note was put in?
9. Are you holding the 10-zloty note in your pocket now?
10. Before the age of 18, did you ever lie to get out of trouble?
11. Are we in Katowice now?

Results

The results of the accuracy estimation are presented below in two ways: the first shows inconclusive results as errors, and according to the second inconclusive results are excluded (Gordon, Mohamed, Faro, Platek, Ahmad, Williams, 253).

Concerning the OSS interpretations within the "guilty subjects" group the percentage of participants correctly classified as "guilty" was 26% altogether (38% – did not include 32% IC diagnoses). Within the "innocent subjects" group the percentage of participants correctly classified as "innocent" was 83% altogether (95 – did not include 3% IC diagnoses).

Concerning the manually numerical score within the "guilty subjects" group the percentage of participants correctly classified as "guilty" was 58% altogether (100% – did not include 42% IC diagnoses). Within the "innocent subjects" group the percentage of participants correctly classified as "innocent" was 67% altogether (80% – did not include 25% IC diagnoses).

As far as all participants are concerned the percentage of participants correctly classified as "guilty" or "innocent" by means of computer scoring was 58% altogether (78% – did not include 21% IC conclusions).

As far as all participants are concerned the percentage of participants correctly classified as "guilty" or "innocent" by means of the manually numerical score was 61% altogether (93% – did not include 35% IC conclusions).
The scoresheets shown below reveal a range of divergence between computer scoring and manually scoring (the figures presented in the scoresheets reflect the number of conclusions).

Within the "guilty" subjects group:

	DI computer scoring diagnoses	NDI computer scoring diagnoses	IC computer scoring diagnoses
DI manually numerical scoring diagnoses	3	4	3
NDI manually numerical scoring diagnoses	0	0	0
IC manually numerical scoring diagnoses	2	4	3

Within the "innocent" subjects group:

	DI computer scoring diagnoses	NDI computer scoring diagnoses	IC computer scoring diagnoses
DI manually numerical scoring diagnoses	0	0	2
NDI manually numerical scoring diagnoses	0	16	0
IC manually numerical scoring diagnoses	1	4	1

Discussion

According to Jerzy Konieczny's recommendation (61, 62), a criterion for acceptability of a polygraph examination technique accuracy was taken from the American Society for Testing and Materials. Among other criteria, the method of forensic identification is accepted if its accuracy reaches 90 per cent (when its outcomes may be used as evidence in court) or 80 per cent (for investigation purposes). Simultaneously, inconclusive decisions may not exceed the 20-per cent limit. The results of the present research do not meet this criterion. Presumably, the primary cause for this is the low level of participant motivation to deceive the polygrapher. As a source of reaction the experimenters could depend only on the subject's feeling of discomfort (when lying is found as something against his or her internalized moral standards) or a participant's joy in misinforming an expert. The research thus faced a typical laboratory study problem.

It is necessary to mention that the quality of recording physiological changes did not belong to the highest. Six subject's charts were rejected on the account of their poor quality. Amongst the accepted charts, improper polygrapher activity sometimes disrupted the course of recording. During a manual evaluation such spots received the number "0". One of the many merits of the numerical scoring systems is the requirement for experts to base on the highest quality record.

An explanation for the divergences received between results of manual numerical evaluations and computer algorithm diagnosis can be found in the characteristics of applied estimation systems. Both put special emphasis on electro-dermal changes. But the diagnostic importance of other recorded changes (concerning the respiration and cardiovascular activity) is more sig-

nificant in the manual numerical evaluation. The electro-dermal activity component is the most responsive of all the parameters recorded by the polygraph (Swinford, 17, 18). At the same time it is more susceptible to coincidental disturbances. It is significant that the mentioned divergences are serious in case of relevant participants. In this instance, some subjects' motivation and the related reactivity to relevant questions were not strong enough.

The accuracy rates received were lower than the results of published reports about laboratory studies into Directed-Lie Polygraph Tests (Raskin, Honts, 25; Kircher, Packardt, Bell, Bernhardt, 35). The explanation can be found in the above-mentioned circumstances: problems with participants' motivation and a lack of skills on the part of the expert.

The results of the present study confirm the statement that a computer diagnosis should not be a single basis for a polygrapher's interpretation decision, but may play a major part in this kind of decision-making (Konieczny, 178–181).

References

Gordon N. J., Mohammed F. B., Faro S. H., Platek S. M., Ahmad H., Williams J. M. (2006), *Integrated Zone Comparison Polygraph Technique Accuracy with Scoring Algorithms*, Physiology and Behaviour, 87 (2).

Kircher J. C., Packard T., Bell B. G., Bernhardt P.C. (2010), *Effects of Prior Demonstrations of Polygraph Accuracy on Outcomes of Probably-Lie and Directed-Lie Polygraph Tests*, Polygraph, 39 (1).

Konieczny J. (2009), *Badania poligraficzne. Podręcznik dla zawodowców*, Wydawnictwa Akademickie i Profesjonalne, Warszawa.

Krzyścin A. (2000), *The Debate over Polygraph in Poland*, Polygraph, 29 (3).

Raskin D. C., Honts Ch. R. (2002), *The Comparison Question test*, [in:] *Handbook of Polygraph Testing*, ed. by M. Kleiner, Academic Press, San Diego.

Shaghnessy J. J., Zechmeister E. B., Zechmeister J. S. (2002), *Metody badawcze w psychologii*, Gdańskie Wydawnictwo Pscyhologiczne, Gdańsk.

Swinford J. (1999), *Manually Scoring Polygraph Charts Utilizing the Seven-position Numerical Analysis Scale at the Department of Defense Polygraph Institute*, Polygraph, 28 (1).

Book reviews

EUROPEAN
POLYGRAPH

Volume 4 • 2010 • Number 2 (12)

A. Vrij

Wykrywanie kłamstw i oszukiwania. Psychologia kłamania i konsekwencje dla praktyki zawodowej,

Wydawnictwo Uniwersytetu Jagiellon'skiego, Kraków 2009

A. Vrij

Detecting Lies and Deceit: Pitfalls and Opportunities,

Second Edition, Wiley Series in the Psychology of Crime, Policing and Law, John Wiley & Sons 2008.

In his everyday practice, Aldert Vrij (b. 1960) deals with the teaching of applied social psychology at the Institute of Psychology of the University of Portsmouth, and is an expert witness in court cases. His chief scientific interest focuses on research concerning lying and deceit, its cognitive, emotional and physiological correlates, and especially on examining the connection between the lie and non-verbal behaviour, a precise description of which may be found in his publications. He is the author of over 300 articles and six books, which primarily refer to the questions of lying and the potential of its detection. The most popular include *Detecting Lies and Deceit: Pitfalls and Opportunities* and *Detecting Lies and Deceit: Pitfalls and Opportunities, Second Edition*, an expansion and continuation of the first publication, especially enriched with state-of-the-art research with the use of innovative methods of lie and deceit detection. The goal of the first monographic work, entirely devoted to the question of deceit and the possibility of its detection, is to present research and theories concerning the connections between deceit and its three most

significant aspects: verbal behaviour (speech content), non-verbal behaviour, and physiological reactions – the aspects that play a significant role in detecting deceit. The author focuses especially on the factors that influence the behaviour of the person communicating a lie, finding them both in individual differences (personality, age, sex) and in the circumstances accompanying the deceit (complexity of the lie, motivation of the deceiver). The publication is composed of three principal parts. In the first, the author discusses the relations between deceit and non-verbal behaviour, emphasising especially the importance of non-verbal behavioural hints for discovering deceit. This part also includes a review and a thorough description of the most important types of non-verbal behaviours that may accompany deceit, as well as a discussion of the difficulties that are inherent in the interpretation of these behaviours in detecting lies. A doubtless advantage is that the author points both to the objective (i.e. confirmed by research) and non-verbal behavioural hints and to the subjective ones that we usually intuitively assume in relation to the potential liar. This is true both for those who deal professionally with the problems of deceit, including investigating police officers and customs officers, and for ordinary people without any special training in this scope. The author points out the mistakes committed most often while detecting lies, considering the capacity of people – both those trained in lie detection and ordinary ones – on the grounds of empirical studies. The publication also embarks on the question of the possibility of training people, especially those professionally dealing with lie detection, in its more accurate detection.

The second part of the work is devoted to the presentation of a relatively new scope of problems concerning the connection between deceit and verbal behaviour. The author presents differences in the content of speech of deceiving and truthful people, and verbal criteria which may be expected among people communicating an untruth. Moreover, he discusses in detail two verbal methods that are used to evaluate veracity: Statement Validity Analysis (SVA), focusing primarily on the Criteria-Based Content Analysis (CBCA), and also on the Reality Monitoring technique, at the same time evaluating the applicability of these methods, and presenting studies of their accuracy in detecting deceit.

The third and last part of the book describes the question of physiological reactions that accompany deceit and at the same time play a significant role in its detection through polygraph examinations. Vrij discusses the principle of the operation of the polygraph, the description of which is enriched with cases from authentic investigations where the tool was used. Further, the author

discusses in detail the two most frequently used polygraphic tests: Control Questions Test (CQT) and Guilty Knowledge Test (GKT). The presentation is accompanied by criticism and results of studies of their accuracy conducted in both experimental and natural conditions. Moreover, the author describes the factors that disrupt the examination and may influence the lowering of efficiency of the method, including those that focus on the phenomenon of "counteracting" that takes place when the person examined produces physiological reactions aimed at deceiving the person conducting the examination during the examination. Moreover, the author points to the profound role of the examiners conducting the test and stresses the need for their highest level of skills and qualifications concerning conducting such examinations. Closing, Vrij compares the aspects of lie detection mentioned above and considers the possibility of combining guidelines coming from both verbal and non-verbal behaviours as well as physiological reactions to increase the efficiency of detecting lies. He also formulates specific guidelines as to how to see through the liar on the grounds of his verbal and non-verbal behaviours. A principal merit of the publication is the multifaceted and honest description of all three aspects of lie detection: verbal behaviour, non-verbal behaviour, and physiological reactions accompanying deceit, as well as disclosing both the problems and the potential in using each of these aspects for lie detection. Additionally, the author focuses precisely on the individual differences in the scope of behaviours connected to deceit that result in the absence of a single, typical standard or a pattern of behaviour suggesting deceit, which makes lie detection a difficult art. Yet, one that, according to the author, can be fine-tuned. A particular innovation is the special focus on personality correlates in both verbal and non-verbal behaviours connected to a lie, and pointing to the fact that these hints may differ one from another depending on the personality type. This provides an opportunity to send studies concerning the lie in a new direction, focusing more on differences than on the search for general regularities related to deceit, as these have so far not brought any satisfactory results. Thus, this book combines scientific and practical values, being on the one hand a thorough description of research and theory concerning mechanisms connected to the lie, while on the other hand providing at the same time a review of state-of-the-art knowledge of the subject. The author refers to works by pioneers in the scope, including works by Paul Ekman, considered a leading researcher of non-verbal behaviour, and provides hints verified through research, telling how people professionally dealing with lie detection may improve the efficiency of lie ascertainment. The author not only provides specific guidelines about the criteria and symptoms of behaviours that should be paid attention to while verifying veracity, but also points out the most frequent errors and stereotypic

assumptions towards the people who lie, proving that the elimination of these mistaken assumptions could help to improve efficiency of detecting deceit, both by people trained in the area and by ordinary people.

The publication is further enriched with additional interesting life study cases and actual investigations, thanks to which it also contains the practical experience of the author and precise description of the experiments and studies conducted, which makes the presentation of issues contained in this monographic work more attractive for the reader.

Detecting Lies and Deceit: Pitfalls and Opportunities, Second Edition, published as part of the Wiley Series in the Psychology of Crime, Policing and Law by John Wiley & Sons in 2008, is a precious expansion and continuation of the issues studied in the preceding publication. Much like the previous book by the same author, it focuses on the discussion of non-verbal, verbal, and physiological markers of deceit, and the capacity to detect the lie on the grounds of these factors. This publication is specifically enriched with the precise description and analysis of modern and promising methods of lie detection, including Functional Magnetic Resonance Imaging (fMRI) and electroencephalography (EEG-P300).

The author supports the controversial thesis that deceiving others is a part of everyday life. We must say that we enjoy a present, even if it is not so, and we also have to praise the food served by hosts even if we do not like its taste. In the first chapter, the author presents the most important definitions of a lie. Mitchell defines a lie as "false communication, which is meant to act for the benefit of the sender of the communication" (p. 12). From this point of view, the number of deceivers includes, besides people, many other living creatures, even plants (for example orchids deceive wasps with false pheromones to entice them to pollinate). The author then presents precisely various types of deceit, enriching his descriptions with interesting examples. Moreover, he offers reasons why people lie, evaluates how often they do so and lists individual differences in the manner of deceiving, which provides a continuation of the question undertaken in the previous publication. Furthermore, Vrij proves that – on the grounds of empirical studies – people far more often deceive those distant to them than those closely related (especially spouses), and justifies this not as much with loyalty (even though it can be taken into account too) as with the fact that a closely related person already knows so much about us that he or she is simply hard to deceive. In the second chapter, the author proves that the lie is in fact a common part of everyday life and that it has

two main faces: on the one hand it harms the people who are lied to, and on the other it may bring them benefits, as – to quote the author's term, lies are a social "lubricant".

The third chapter touches upon the relationship between non-verbal behaviour and deceit. The author develops the questions posed in the previous book, presenting new theories, for which non-verbal symptoms may point to the presence of deception. Here, his grounds include studies aimed at the observation of behaviour while telling the truth and deceiving, the experience of police inspectors, and cases of lies observed among politicians. The author presents the conclusion here that the majority of non-verbal hints seem to be entirely unconnected to deceit, or the connection is very weak. He also points to the relationship between non-verbal behaviour and deceit being very complex, among other reasons because individual people manifest different behaviours connected to the lie, and also because the behaviour of the liar depends on the context of the situation.

In the successive chapters, the author presents the basic behavioural hints of deceit and describes in detail the tools that serve assessment of veracity that are also described in his previous publication: Statement Validity Analysis (SVA), and Reality Monitoring (RM). He devotes separate chapters to careful analysis and definition of accuracy in lie detection in two interesting methods, namely Behavioural Analysis Interview (BAI), focusing mostly on the discovery of non-verbal lie indicators (even though it also accounts for certain verbal aspects) and the Scientific Content Analysis (SCAN), which analyses the verbal activity from the point of view of veracity assessment.

Chapters 11 to 13 present the most important physiological deceit markers. Besides the classical polygraph, described already in detail in the previous publication, the book presents innovative methods that give hope for increasing efficiency in lie detection. These methods include the Thermal Imaging Technique and Voice Stress Analysis. Chapter 12 focuses on the description and assessment of usefulness of the method based on electroencephalography analysing the P-300 signal in the revelation of veracity in testimonies (EEG-P300). Chapter 13 presents the technology that is most expensive to use and most advanced: neuroimaging with functional magnetic resonance (FMRI). Chapters 14 and 15 present the reasons why people are so often incapable of detecting lies, while the last chapter deals with the question of improving and developing the skills of deceit detection, pointing at the same time to the need to cooperate with representatives of other fields, for example psychology.

Beyond doubt, this book is both a continuation and primarily an update of the questions tackled in *Detecting Lies and Deceit: Pitfalls and Opportunities*, the reason behind this being the great progress in the research and also emergence of new, highly promising methods serving lie detection without which the description of state-of-the-art knowledge would be impossible. The book focuses especially on the precise analysis of individual techniques used in lie detection, with each of these techniques having a separate chapter devoted to it by the author, who presents both an insightful criticism of them and the potential that the application of these methods brings to improving the efficiency of lie detection. The author presents the question of lies and deceit in all its complexity, and the book meticulously exhausts the subject on which it embarks.

Thanks to the holistic look at the issues, both books assume a significant place among the available literature devoted to the phenomenon of the lie and the potential for its detection. Beyond doubt they provide a significant contribution to expanding knowledge of the subject. These books are precious both for experts dealing professionally with the question (lawyers, investigating officers, psychologists), and anyone interested in improving their efficiency in everyday lie detection.

Natalia Mirska*

* e-mail

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
ul. Gustawa Herlinga-Grudzińskiego 1
30-705 Kraków (Poland)

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