EUROPEAN
POLYGRAPH

Volume 8 • 2014 • Number 4 (30)

Marek Leśniak*
Faculty of Law and Administration
University of Silesia in Katowice
Katowice
Poland

Selected Problems in Evaluation of Polygraph Examination Results

Key words: polygraph examination, results of polygraph examination, methodology of chart evaluation

Introduction

Standardisation is absolutely essential in methods of polygraph examination because every quality control is based on comparing features of a specific polygraph examination and established polygraph examination standards. In the case of polygraph examination, such a standardisation includes:

- the manner of carrying out a pre-test interview with the examinee
- the manner of selecting control questions
- the choice of control questions

* marek.lesniak@us.edu.pl

- the manner of presenting questions to an examinee
- the identification of countermeasures
- the measurement and quantification of physiological responses
- decision about the configuration of obtained data.

However, the application of standards for an evaluation of polygraph examination results may be connected with limits in obtaining more detailed information or may meet resistance. The objective of this study is to provide examples of such cases. In my study, I use data from observations derived from two of my empirical research projects in polygrapher conclusion accuracy.

1. The First Study

1.1 Methodology

The study included 86 participants recruited from among students of the Police School in Katowice. Some ("guilty subjects") read a disturbing text. They were instructed to deny having anything to do with the text, and to keep it in a pocket during polygraph examination. The participants were tested with a Lafayette Statement conventional polygraph. The interpreter learnt the actual role of each examinee after they made a decision about the role. Interpretations of polygraph charts were conducted manually with Backters numerical scoring scale. Every chart included three pairs of relevant-control questions. Values from -3 to +3 were assigned to each pair and to each reaction type. Thus, the chart contained nine possible measurement options, and the whole polygraph test – twenty seven, as each test was based on three charts. I drew conclusions after obtaining the global score from each particular test. I changed the size of the range of inconclusive results and reached four conclusions: they were sometimes different. I took four situations into account: without inconclusive results, and with the inclusion of inconclusive results in three ranges: from -5 to +5, from -10 to +10, and from -15 to +15. Then the accuracy of all polygraph test results was determined, separately with reference to each situation. Finally, I aggregated the data obtained to establish how the accuracy of polygraph test changed depending on the shifting of the inconclusive outcomes range. The findings are presented below.

1.2 Results

Test results (without inconclusive decisions)		
correct decisions	**wrong decisions**	
69.5%	30.5%	all subjects
72%	28%	"guilty"
67%	33%	"innocent"

Test results (with results from -5 to +5 considered inconclusive)			
correct decisions	**inconclusive decisions**	**wrong decisions**	
64.5%	23.5%	12.5%	all subjects
70%	24%	6%	"guilty"
58%	23%	19%	"innocent"

Test results (with results from -10 to +10 considered inconclusive)			
correct decisions	**inconclusive decisions**	**wrong decisions**	
59%	30.5%	10.5%	all subjects
66%	22%	12%	"guilty"
52%	39%	9%	"innocent"

Test results (with results from -15 to +15 considered inconclusive)			
correct decisions	**inconclusive decisions**	**wrong decisions**	
48%	51%	1%	all subjects
56%	44%	0%	"guilty"
40%	58%	2%	"innocent"

Test results ("guilty" subjects)				
polygrapher's decisions	**without inconclusive results**	**inconclusiveness range from -5 to +5**	**inconclusiveness range from -10 to +10**	**inconclusiveness range from -15 to +15**
correct decisions	72%	70%	66%	48%
wrong decisions	28%	24%	12%	0
inconclusive decisions	-	6	22	51

Test results ("innocent" subjects)				
polygrapher's decisions	**without inconclusive results**	**inconclusiveness range from -5 to +5**	**inconclusiveness range from -10 to +10**	**inconclusiveness range from -15 to +15**
correct decisions	67%	58%	52%	40%
wrong decisions	33%	19%	9%	2%
inconclusive decisions	-	23%	39%	58%

Test results (all subjects)				
polygrapher's decisions	**without inconclusive results**	**inconclusiveness range from -5 to +5**	**inconclusiveness range from -10 to +10**	**inconclusiveness range from -15 to +15**
correct decisions	69.5%	64.5%	52%	48%
wrong decisions	30.5%	12.5%	9%	1%
inconclusive decisions	-	23,5%	39%	51%

1.3 Discussion

I would like to remark that the accuracy of a polygraph test result varies depending on the width of the range earmarked to inconclusive decisions (obviously on condition that a polygrapher uses a numerical or a quasi-numerical scoring system). Thus, accuracy depends on decisions of authors of standards. It is a general rule that standards of a particular polygraph examinations method include such a range, which is determined explicitly in advance. Experts and persons using the results of polygraph tests do not receive information about different "accuracies" of polygraph test results depending on the assumed inconclusiveness range. They cannot obtain more detailed data.

Results of the first study suggest emphasising that by expanding or narrowing the range of inconclusive results, the rate of false positives or false negatives may be increased. When expanding the range, we opt for obtaining more false negatives (type II errors) and narrowing it, we increase the occurrence of false positives (type I errors).

When the results of a polygraph examination are used as evidence before the court, type II errors should be preferred. Yet, whenever polygraph examinations are used for screening, there is nothing to bar the preference of type I errors. Such an approach to using a polygraph scoring system requires more detailed information on the accuracy of polygraph test results, as influenced by changing the range of inconclusive diagnoses.

I disagree with forcing lawyers to use only polygraph techniques whose level of accuracy reaches a precisely defined point (e.g., with probability of errors below 10%). Information about the accuracy of polygrapher opinion is obviously the foundation for the decision whether to use it as evidence. Therefore, the purpose of polygraph examination may require various sizes of the inconclusive diagnose range. Accepting only one such range for a particular scoring system results in polygraph result users losing valuable information.

2. The Second Study

2.1 Methodology

The research covered 18 participants recruited from among students of the Silesian University in Katowice. Some ("guilty subjects") took a note out of the professor's cabinet. They were instructed to deny having anything to do with

the note and to keep it in a pocket throughout the polygraph examination. The participants were tested with a Lafayette LX–4000 computer polygraph. While making the decision, the interpreter did not know the actual role of the examinees. Standards of the Utah Directed-Lie Test were applied. Interpretations of polygraph charts were conducted manually with both a numerical scoring scale and computer scoring applications: OSS 2 and OSS 3. These algorithms are sold bundled with Lafayette polygraph software. Much like in the first study, every chart included three pairs of relevant-control questions and every test was based on three charts. After obtaining the global score of a particular test, I compared the accuracy of using of a particular scoring system with the accuracies of others. The findings are presented below.

2.2 Results

Manual numerical scoring diagnoses		
correct decisions	**wrong decisions**	
83%	27%	all subjects
90%	10%	"guilty" subjects
75%	25%	"innocent" subjects

OSS 2-supported diagnoses		
correct decisions	**wrong decisions**	
61%	39%	all subjects
40%	60%	"guilty" subjects
87.5%	12.5%	"innocent" subjects

OSS 3-supported diagnoses		
correct decisions	**wrong decisions**	
55.5%	34.5%	all subjects
87.5%	12.5%	"guilty" subjects
55.5%	34.5%	"innocent" subjects

2.3 Discussion

Results of the research are open to the following questions:

- The applied application-based algorithms evaluated "innocent" subjects with high accuracy. However, the evaluations of "guilty" subjects showed poor accuracy. This means that the designers of these computer-based scoring systems impose the preference of the type II error to the type I error on the users.
- Compared to the OSS 3, the OSS 2 scoring system less often decided about the inconclusiveness of a result.
- The manual numerical scoring system proved more effective than the computer scoring systems. This may suggest (among other things) that the difference lies in the cultural background.

In my opinion, there is a need for subsequent research of the issues listed above. They are very important for using polygraph test results in real cases in Poland. For example, there is a risk inherent in only a computer scoring system (OSS 2 or OSS 3) being used by an inexperienced polygrapher. In line with the findings presented above, in such a case, the level of false negatives reaches approximately 50%.

Conclusion

Regrettably, in Polish practice examiners very often rely on an overall evaluation of polygraph charts (without using numerical or quasi-numerical scoring systems). Such an evaluation is based on an expert's subjective experience and, as a matter of fact, it is beyond quality control. That is why there is certainly a need for using standards. However, as explained above, their use may cause some problems.

It should be noted that what was examined in both the studies were only results of laboratory experiments, and their results should be taken with a pinch of salt with translated into real life situations. The presented studies are not free from flaws either. For example, the population recruited for the second polygraph examination is not relatively large. Moreover, such research requires application of more sophisticated statistics. Nevertheless, the findings are validated by the result of other studies. That is why the issues presented should be further investigated in future research carried on a reference population for Poland.

References

Widacki M. 2014, *Aktualny standard badań poligraficznych a praktyka* polska, [in:] J. Widacki (ed.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM, Kraków, p. 193.

Konieczny J. 2010, *Polygraph examination as scientific evidence*, European Polygraph 3 (13), pp. 129–135.

Leśniak M., Zubańska M., *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)*, European Polygraph 2 (12), pp. 84–85.