

Breakthrough in the validation and diagnostic value of test carried out in polygraph examination

Marcin GOLASZEWSKI

Polish Internal Security Agency (ABW)

Standardization of polygraph testing

In the last few years polygraphs both to Poland and abroad have been paying more and more attention to standardization and standards. Standard is a common and average model of something¹. A norm, on the other hand, according to the provisions of the Act of 13.09.2002 on standardization² is *a document adopted by common consent and approved of by an authorized organizational unit, setting rules, guidelines and providing specifications to be commonly and repeatedly used for a variety of activities in order to impose optimum order on a specific area*. The aims of standardization include: ensuring the quality and reliability of products, processes and services, as well as facilitating communication by defining terms, labels and symbols for common use. Whereas conditions of standardization include: unclassified nature and general availability of standards, respect for public interest, freedom of participation in the process of the development and use of standards, making it possible for all person interested to be involved in the process of developing standards, consensus as a bases for the determination f the content of standards, unification and coherence of the provisions of standards, use of effective scientific and technical achievements.

The following standards are applied to polygraph testing:

- Who may undergo polygraph testing, under what circumstances, who may the examiner be;

¹ Słownik Języka POLSKIEGO PWN, <http://sjp.pwn.pl/slownik/2576133/standard> [access: 21 Sept 2012].

² Act of 12.09.2002 on standardization (Dz.U. z 2002 r. "Nr 169. poz 1396).

- Types and methods of polygraph testing (including the preparation, pre-test interview, testing techniques, and data analysis);
- Reporting and drafting expert opinions;
- Quality control;
- Professional ethics.

Rules, guidelines and specification for the standards for polygraph testing listed above are developed based on:

- legal norms

In Poland they include merely premises for conducting polygraph testing and the obligatory consent of the examinee. Decisions taken by heads of Polish services cannot be considered standards due to the lack of their uniformity. By comparison, in the USA there is a quality control system for polygraph testing, as well as licensing requirement in many states);

- organizational norms (regulations adopted by professional organizations, e.g. American Polygraph Association (AOA), Association of Polish Polygraphers (SPP) and standardization bodies, e.g. American Society for Testing and Materials International (ASTM International³).

- Practice and long-standing conventions (e.g. there is a habit not to test pregnant women, although no regulations are to be found).

Why do we need standards?

Most practical areas of polygraph testing in Poland are not governed by any regulations. This poses a risk of violations and incompetence, and eventually loss of trust to examiners and the testing method. There was a similar problem in the 80s and 90s of the twentieth century, as there were no standards or regulations in place regarding the practice of polygraph testing. Together with cost cuts and severe competition on the polygraph testing market, it leads to a pathological situation, where a possibly large number of tests are performed at the expense of their reliability. The American society started, to defend itself against polygraph testing. The U.S. Congress introduced some restrictions on polygraph testing. Two legal acts deserve particular attention. The first was adopted in 1990 under the title Americans

³ The international nature of ASTM was enhanced by adding the Word International to its name in 2001. A similar idea was discussed during the 45 annual APA symposium in Myrtle Beach in 2010. However, the conservative approach received more enthusiasm, i.e. the belief that the old name is recognizable, which had been achieved within many years.

with Disabilities Act (ADA). This law forbids asking any questions about medical records, which also concerned alcohol consumption and taking medicines. In 1998 the Congress adopted Employee Polygraph Protection Act (EPPA) that restricted polygraph testing of employees. It excluded Government agencies and public order bodies, as well as employees of the pharmaceutical and nuclear industries.

The American experience shows that Poland should learn from the mistakes of others. Currently, the idea of deregulating gains more and more support among various professionals and specialists. However, in the area of polygraph testing there has been the opposite trend-an increasing need of introducing new regulations and standards.

Polish efforts to introduce uniform high quality standards of psycho-physiological examination based on solid scientific research

In Poland common standards for polygraph testing have been developed by specialists mainly through exchange of information at scientific seminars. One of the first such meeting was a national scientific symposium Polygraph Testing for Legal Purposes organized in 1976 in Toruń. In the following years there have not been any significant developments in this area, until the beginning of the 21 century. Every year the police and the military police take turns in organizing a series of polygraph seminars. In June 2010 an International Seminar of Polygraphers was held in Emów. It gave an impulse to some serious work on the standardization of polygraph testing in Poland.

The Association of Polish Polygraphers (SPP) that was established in 1994 provides another platform for work on uniform standards. After a few years of standstill the operation of the Association was reactivated in 2012. The Management Board of the Association provided the Minister of Interior, Jacek Cichocki, a Memorandum on the Possibilities of a More Comprehensive Use of Polygraph Testing in the Polish Special Services and Police Forces. It pointed out varied levels of the qualification of experts employed in state institutions, and, consequently, the need of regular professional training, including consultations with experienced trainers from the countries leading in polygraph testing (USA, Israel). It also expressed regret that many investigators among officers and employees of the judiciary are not fully aware how to use polygraph in investigations, so they fail to exhaust all its possibilities. They too should undergo the necessary training.

Moreover, the document indicates all areas, other than the common application, where polygraph testing might be of use. Polygraph testing might be more useful in psychophysiological testing for operational verification of sources of information and employment purposes (not only with, regard to candidates for service⁴ but also for promotion to sensitive posts, granting access to most important state secrets, as well as regular control of officers). As for new areas of application, it is suggested that psychophysiological testing⁵ might prove useful in handling crown witnesses, as well as persons who apply for security clearance, asylum or a refugee status. Some of these proposals would require drafting proper guidelines, regulations, or even reform of legal acts.

The memorandum also pointed out polygraph testing services offered on the private market. There has been some reasonable concern about unqualified persons being used as experts. In response to these concerns the SPP declared its readiness to present any interested bodies with a list of recommended persons, whose qualifications raise no doubt.

Generally, however, Polish polygraphers have bright future in front of them. The development of polygraph testing in Poland has been appreciated by the APA, an international organization that enjoys established reputation. According to its management, Poland might even assume a leading role in Europe. APA is ready to organize a serious scientific seminar like the one that regarded Asia and Pacific and was held in Singapore in the beginning of 2012. In the next few years n APA accredited office is expected to be established in Poland for the whole Europe.

Standards with international outreach. New higher APA standards of 2012

Standards for common polygraph testing can be found in obligatory regulations and recommended guidelines of American Polygraph Association, as well as the ASTM International standards. The first of those organizations was established in 1966. It gathers three thousand polygraphers. The latter is a standardizing organization set up in 1898.

⁴ It is astonishing that the candidates for service in the Government Protection Bureau do not undergo such examinations.

⁵ More and more often you can hear about a possibility of using polygraph testing for control and therapy of persons convicted of sex crimes. Such advanced programmes are carried out to the United States and Great Britain.

Although both of their names contain the word *American* they are in fact international and gather members from around the world.

On 1 January 2012 new APA standards of practice drafted were introduced. They laid down the following roles:

- obligatory use of movement detector;
- general requirement of using only those techniques that were stipulated in the validation (and are, therefore, confirmed scientifically);
- criteria for use of specific techniques for particular type of testing: in order to gather evidence, for the purposes of confrontation, detection (investigation) or selection;
- Requirement of calibrating the polygraph at least once every 6 months (Le. testing its performance).

It should be noted that *validus* means in Latin: strong, tough, and effective⁶.

Validation of the measuring method-according to PN-EN ISO.IEC 1725:2005⁷ - through examination and presenting objective evidence confirms that certain requirement have been met with regard to the intended use. In other words, the aim is to confirm that the test measures the parameters it was designed to measure. The analysis process carried out using a specific method has to be credible, provide reliable results and run efficiently, ensure coherence, i.e. the expected equivalency between tester (in the case of polygraph testing, the most important question is how accurate the polygraph is, when a random tester tests an average randomly selected subject from a random place around the world). The accuracy (correctness and precision) of the measuring methods and results is defined by the Polish Norm PN-ISO 5725-2; 2002⁸.

As Pamela Shaw (APA President 2011-2012) pointed out rightly, *the requirement to use validated testing methods is not a new idea, of course. Other fields such as medicine and psychology eventually came to the same conclusion, albeit many years after the fields were established. It has turned out to be a great thing for them. Try to imagine, if you can, what fields of medicine and psychology would be like if there were no*

⁶ W. KOPALIŃSKI, *Słownik wyrazów obcych i zwrotów obcojęzycznych*, Warszawa 2001, De Agostini, p. 532.

⁷ *Ogólne wymagania dotyczące kompetencji laboratoriów badawczych i wzorcujących*, Polski Komitet Normalizacyjny (Polish Committee for Standardization), Warszawa 2005.

⁸ *Dokładność (poprawność i precyzja) metod pomiarowych i wyników pomiarów*, Polski Komitet Normalizacyjny (Polish Committee for Standardization), Warszawa 2002.

*requirement to validate their methods. Validation serves a number of important functions, not the least of which is protecting the public from misuse, incompetence and quackery*⁹.

We should remember that the method consists not only in asking a sequence of test questions. It is also a set of rules to be implemented with regard to the pre-test interview, formulating and discussing the test questions, presentation of stimuli (questions) during the test, as well as the method of analysis of the test results. According to the APA guidelines, for a research method to be considered scientifically reliable, it needs to have to following features:

- the form of the test, needs to comply with research rules for the selection of aims, formulating questions and giving stimuli during the test;
- the form of the test needs to be based on scientifically reliable analysis method for test results;
- the studies on the new techniques should be published in Polygraph, or other Independent scientific magazines, government or university publications, at least twice.

The described guidelines also determined the criteria for the use of the method for particular type of testing: They are the following:

- **evidentiary techniques** (for judicial bodies); $\geq 90\%$ of accuracy (correct decisions) and $\leq 20\%$ of inconclusive results;
- **paired testing techniques** (two experts test two or more subjects that assert opposed information in a way that one of them surely lies) $\geq 86\%$ accuracy and $\leq$ inconclusive results;
- **investigative techniques** (for investigative purposes) $\geq 80\%$ accuracy (correct decisions) and $\leq 20\%$ of inconclusive results;
- **screening techniques** (verification): studies confirm a significantly higher level of accuracy than, a statistical chance together with the method of successive overcoming barriers that requires additional confirmed and more accurate tests, if the screening test does not end positively (i.e. there is some doubts as to the credibility of the subject),

APA members are subject to the new standards of practice developed by APA since 1 January 2012 (unless the national law stipulates otherwise). Another acknowledged organization, AAPP (American Association of Police Polygraphists) is planning on introducing such standards soon. Also other, state and national, polygraphist association are expected to do so. Polish Internal Security Agency has also been applying these standards.

⁹ APA, „Polygraph”, 2011.40(4), p. 194.

The above-mentioned standards that have recently been put in place, were developed in 2007 as a response to a 2003 report by an American scientific and research organization National Research Council¹⁰, which shapes public opinion. The conclusions of the report contained both negative and positive comments. However, the first prevailed:

- scientific basis for polygraph testing differ greatly from what is expected of the test, which has a huge impact on decisions regarding state security;
- a lot of literature on polygraph testing is purely theoretical;
- the possibilities of verification of the reliability of test results are highly limited, even taking into consideration the advancement of measuring and evaluation techniques;
- there are reasons for doubt as to the effectiveness of the use of disturbing measures during the examination.

The NCR report also contained only moderately positive perspectives for the future. According to the report although scientifically large limitations on the matching rate has been proven, it is possible to achieve such accuracy rate that they are useful in practice. A meta-analysis carried out by NCR the accuracy rate for specific purpose of tests is between 0.81 and 0.91 for 26 average values out of 52 sets of data examined¹¹. That means that, despite general critics, the high effectiveness of such studies has been confirmed. The results were obtained in 2003. Currently, the knowledge about polygraph testing is much wider.

APA meta-analysis of polygraph testing techniques

The crucial question is which of the available techniques meet the new higher standards (validation criteria) developed by APA in 2012. The answer has been provided by a meta-analysis, i.e. an independent scientific study consisting in secondary discovery of knowledge through a gradual review of a specific branch of knowledge.

A meta-analysis carried out by APA Ad-Hoc Committee on Validated Techniques encompassed: 37 studies (52 experiments and surveys), 289 scorers, 12665 scored results of 4283 confirmed exams (6597 scored results of 2300 confirmed deceptive exams and 6068 scored results of 1983 confirmed truthful exams). Studies that provided no statistical data of interest to the meta-analysis, as well as those whose testing method did not comply

¹⁰ National Research Council, The polygraph and lie detection, Washington 2003, The National Academies Press.

¹¹ See p. 148.

with the identifiable specified techniques and analysis methods for test data, were excluded. The meta-analysis produced the following results:

- **an average accuracy rate of all recognized polygraph testing techniques** of 87.1 %, with an average inconclusive rate (INC rate) of 12.7%;
- **an average accuracy rate of recognized single-issue techniques** of 92.1%, with an INC rate of 8.8%;
- **an average accuracy rate of recognized screening techniques** of 85%, with an INC rate of 12.5%
- **a list of techniques to be used for various types of polygraph testing** (in force since 01.01.2012.)

Table 1 List of techniques for polygraph testing according to APA standard 2012

Evidentiary techniques (for the purposes of judiciary bodies) Test data analysis method	Multiple subject testing methods (mainly paired testing methods) TDA method ¹²	Investigative techniques (the remaining tests)TDA method
US federal You-Phase/ESS ¹³ ■ accuracy: 90.4% ■ inconclusive (INC): 19.2% ■ sensitivity ¹⁴ :84.6% ■ specificity ¹⁵ 75.5	USAF MGQT ¹⁶ /ESS ■ accuracy: 87.5% ■ INC:17% ■ sensitivity:72,9% ■ specificity:70%	USAF MGQT/7-pos. scale ■ accuracy: 81.7% ■ INC: 19.7% ■ sensitivity: 78.3% ■ specificity: 53.8%
ZCT (Federal, Utah)/ESS ■ accuracy:92.1% ■ INC:9.8% ■ sensitivity:81,7% ■ specificity: 84.6%	Federal You-Phase/7-pos. scale ■ accuracy: 883% • INC:16,8% ■ sensitivity: 84.5% ■ specificity:75,5%	CIT (GKT)/Lykken system ■ accuracy:823% ■ INC:0>1% ■ sensitivity:81.5% ■ specificity: 83.2%
Utah ZCT (combined versions)/Utah ■ accuracy: 93% ■ INC: 10,7% ■ sensitivity: 85.3% ■ specificity: 80,9%	Federal ZCT/7 pos. ■ accuracy: 86% ■ INC: 17.1% ■ sensitivity: 5.8% ■ specificity: 58.1%	DLST(TES)/7po\$. ■ accuracy:84,4% ■ INC:8.8% ■ sensitivity:74.8% ■ specificity: 79,2%

¹² Paired testing - Am method of utilizing polygraph testing in situations in which two or more subjects assert contradictory accounts of a particular incident in such a way that at least one of the subjects must certainly be lying.

¹³ Empirical Scoring System-an evidence-based normative system for manual test data analysis of PDD examination data.

¹⁴ Sensitivity - the proportion of true positives a test can produce, this term describes how well a test identifies a person engaging in deception, ability to identify issues raising doubts that pose a threat to the subject (significant doubts), reduction of false negative scores (reactions of the subject that are incorrectly considered by the examiner as typical of truthful persons).

¹⁵ Specificity - the proportion of true negatives, the test verifies the honesty, identifies a problem, and reduces false positive scores (reactions of the subject that are incorrectly considered by the examiner as typical of untruthful persons).

¹⁶ Average data for AFMGQT both versions (1 and 2). LEPET and Utah MGQT are similar to this technique. Therefore, they have been deemed acceptable, as long as the same test data analysis methods are applied as with AFMGQT.

Utah ZCT DLC/Utah ■ accuracy:90.2% ■ INC: 7,3% ■ sensitivity:81,5% specificity: 85.7%	Federal ZCT/7 pos. evidentiary ¹⁷ ■ accuracy: 88% ■ INC: 8.5% ■ sensitivity: 80.4% specificity: 80.9%	DLST (TES)/ESS ■ accuracy: 85.8% - INC: 9% ■ sensitivity: 80.9% ■ specificity: 75.1%
Utah ZCT PLC/Utah ■ accuracy:93.1% ■ INC: 7.7% ■ sensitivity:86.7% ■ specificity: 833%	Backster You-Phase/ Backster ■ accuracy: 86.2% ■ INC: 19.6% ■ sensitivity: 83.6% specificity: 55.6%	
Utah ZCT RCMP (v.1)/Utah ■ accuracy: 93.9% »INC: 18.5% ■ sensitivity: 833% specificity: 70%		
IZCT/HSS* ■ accuracy: 99.4% ■ INC: 33% ■ sensitivity: 97.7% specificity: 94.6%		
MQTZCT/Matte* ■ accuracy: 99.4% ■ INC: 2.9% ■ sensitivity: 96.7% specificity: 963%		

Source: Own analysts based on the Committee report on Validated Technique APA "Polygraph" 2011, 40(4).

How to read the table? In the first column were placed such techniques that meet criteria ≥ 90 accuracy and $\leq 20\%$ inconclusive. In the second column-techniques with 86% accuracy and producing no more than 20% inconclusives. Techniques from the first column can be applied also in examinations specified in the second column, while in investigative examinations one can use techniques mentioned in all columns. Looking from the left to the right side of the table, criteria of the admissibility (the accuracy) are being lowered.

At two techniques-Integrated Zone Comparison Technique (IZCT) and Matte Quadri-Track Zone Comparison Technique (MQTZCT) - were marked as references with provisions. These techniques were listed in the table; however, it was indicated that statistical data are inconsistent with the distribution of results from all other techniques and are called outliers. Therefore, one ought to look at these data with large caution. All the more that IZCT and the MQTZCT were not verified by independent researchers. Furthermore, APA paid attention to some shortcomings in the validation process of these techniques.

¹⁷ In the evidentiary 7-pos, scale the decision threshold for the NDI decision is slightly lower than in the case of a traditional one and it is +4. For the DI decisions it remains unchanged (-6).

Also, the list does not contain those techniques that could not be confirmed by any studies that would confirm their reliability and effectiveness or if the determined accuracy rate or INC rate exceeded the APA requirement for evidentiary, paired or greening examinations* Those techniques include: US Army MGQT, Reid's test, POT-B, Marcy's technique or R/I¹⁸.

All those techniques not contained in the list should be marked as experimental Techniques, like R/I, POT-B or SAT may be used additionally, if they are not the core that an opinion on the subject's reliability is based on. The set of possible techniques is open. Any new techniques that meet the criteria of validity and minimum levels of accuracy, as well as the maximum INC rates may be added.

It is possible that some polygraphists have been utilizing some techniques, believing in its effectiveness. Therefore, it raises some concern that such a technique might not have been scientifically proven effective and has not been included in the list. In such a case they should adjust to the APA standards, since APA is the largest organisation gathering polygraphists from various geographic areas all around the world and has a great impact on the methodology and development of polygraph testing. In the last few years the organization has been putting particular effort in the promotion of scientific verification of the opinions produced based in polygraph testing. Consequently, such opinions would be perceived as scientifically proven by potential users of polygraph testing, its subjects and the judicial scientific circles.

Use of polygraph evidence (legal regulations in Poland and the USA)

In 1920 there was precedence in the United States, when a defense counsel of James Frye, who had been accused of homicide, lodged a motion to put his client to a polygraph test. Then the court formulated an idea of *general acceptance*, which means that it is possible to admit scientific evidence, when it is the high value of the evidence is generally recognized by experts in this particular area. Most experts decided that the development polytrophic testing is still in progress and their reliability is insufficient.

¹⁸ The use of the latter two techniques is temporarily acceptable until 2012. This is additional time for validation. However, the efforts of the supporters of these techniques seem unlikely to succeed.

In 1989 the United States Court of Appeal, Eleventh Circuit in the case *United States vs. Piccinonna*, ruled that per se that the exclusion of polygraph evidence is no longer guaranteed¹⁹.

In 1993 the US Supreme Court in the case *Daubert vs. Merrell Dow Pharmaceuticals Inc.* ruled that the Fry standard was too restrictive. The court determined the condition of admissibility of scientific evidence that was later referred to as Daubert standard:

- possibility of testing the technique;
- determination of a possible rate of error for the technique;
- available peer review and publication of studies on the technique;
- the level of acceptance of the technique within the relevant scientific community;
- setting standards for the correct, and acceptable use of the technique.

Although in Poland the Daubert standard is not officially applied the conditions were mentioned as a reference point that should be considered by all Polish polygraphists. As for admitting polygraph evidence is court the most applicable provisions in the Polish law are contained in Articles 192a, 193 and 199a of the Polish Code of Criminal Procedure²⁰. Article 193 lays down general rules with respect to experts saying that *if the determination of material facts having essential bearing upon the resolution of the case some special knowledge, the court shall consult an expert or experts*. However, Article 192a stipulates that *in order to reduce the circle of suspects or determine the probative value of identified traces fingerprints, the cheek mucosa, hair, saliva, sample letters, or fragrance can be collected, a photograph of the person can be taken or to their voice can be recorded. (...) With a consent of the examined person, an expert may also apply technical measures in order to control unconscious reactions of the person's body*. This provision is addressed to a body that carries out preliminary proceedings in a given case. Article 199a complements the provisions of Article 192a par. 2 and has the following wording; *The use of technical*

¹⁹ M. HANDLER, Ch. R. HONTIS, D. J. KRAPOHL, R. NELSON, S. GRIFFIN, Integration of pre-employment polygraph screening into the police selection process. *Polygraph* 2009, 38(4), p. 242.

²⁰ Act of 6 June 1997 – Code of Criminal Procedure (Dz.U. z 1997 r. iNr 89, poz. 555 z późn. zm).

measures by the expert in their examination in order to control the unconscious reactions of the person's body is possible only after receiving the consent of the person. What does this mean? An indirect conclusion may be that based on this article polygraph testing may be carried out also in the *in personam* stage of the preliminary proceeding with regard to the subject, or even in the court proceedings, with respect to the defendant²¹. Polygraph may also be used on a witness (Article 192 par. 1). Article 199a repeals the evidence prohibition specified in Article 199 (i.e. the inadmissibility of the defendant's statement as evidence for this being guilty of committing the act he was accused of committing that could be submitted by the expert).

However, as Tadeusz Tomaszewski (Ph.D.) rightly concluded “The largest problem does not concern the provisions of the codes (although of course their interpretation is incorrect), but the evaluation of evidence. It is known that the Polish criminal court procedure is based on the principle of the freedom of evaluation of evidence. Therefore, it has always been common practice that the provisions allowed the admission of polygraph evidence. It was only at the discretion of the court to evaluate whether the method was sufficiently reliable to be applied. The court decides, whether to use the expert opinion based on polygraph testing in its ruling. Unfortunately, lawyers lack proper training and expertise on polygraph testing. Therefore, many solicitors do not understand, how to use the polygraph properly”²².

Significant progress in the production of test results. Solution to some problems concerning practice

Until recently, the only technique that made it possible to specify the probability of error for the measured results was CIT²³. This is, in fact, a slightly expanded techniques that includes a series of tests regarding the set:

²¹ T. GRZEGORCZYK, *Kodeks postępowania karnego. Komentarz*, Kraków 2005, Zakamycze, p. 513.

²² T. TOMASZEWSKI, *O biegłych dla biegłych - wykład dla ekspertów*, w: *Normy prawne i standard branżowe w zakresie badań poligraficznych w wybranych krajach*, material from the International Symposium of Polygraphers, Central Training Centre ABW, Emów 21-24.06.2010, p. 90.

²³ Concealed Information Test.

GKT²⁴ or POT-A²⁵ that uses Lykken evaluation system²⁶. Its disadvantage undoubtedly is the use of scenario that is only in a few percent verifiable in polygraph testing. Since 2009 experts also have the Empirical Scoring System (ESS) at their disposal, which might be used in most recognized polygraph techniques. Its nature is almost revolutionary. ESS is a tool that significantly improves the diagnostic value (in terms of evidence) of an expert opinion based on polygraph testing. It makes it possible to determine the actual statistical value (probability of error) of test results. The system is simple and easy to use. It takes into account standard data (representative data that may be used as reference points and compared with another data), both for evidentiary and screening tests. Therefore, its application is extensive. It also provides the sensitivity, specificity rates, as well as the rate of false positive results, false negative results, and inconclusive with respect to subjects who lie and tell the truth.

In many cases polygraphists draw a conclusion from test results that there is “a high probability of the subjects misleading statement or the opposite. For example, in the decision made by the Commander -in-Chief of the Polish Border Guard²⁷ presented a version of an expert opinion that might be drafted based on a polygraph test: *as a result of the analysis of polygraph materials, the explanations made by the subject and their behavior during the examination, there is a high probability that the subject is hiding relevant facts regarding questions (no of the question), which may mean they are not an adequate candidate for a Border Guard officer.* However, such conclusion has a serious fault. It raises a question about what “a high probability” means. What was the adopted error tolerance and what is the probability of error (p value) for the results of particular tests carried

²⁴ Guilty Knowledge Test, in Poland known as a test of knowledge about the crime-conventional reconnaissance test.

²⁵ Peak of Tension A, with available answers, (the test includes the key).

²⁶ In the Lykken system only electrodermal activity is taken into consideration. If the strongest physiological reaction accompanies the key question, the score is 2. If the second largest EDA/GSR reaction accompanies the key question, the score is 1. All other variants are scored 0. The final score may vary from 0 to a double CIT number. The decision threshold is the number equal to the number of the CIT series that have been carried out. A special table specifies the probability of identifying the key test element by the subject for each test score, e.g. a total number of 7 for 5 series of questions means 3% probability that the subject does not know the details of the event.

²⁷ Regulation nr 4 of the Commander-in-Chief of the Border Guard of 11 January 2012 (Journal of Laws of the Border Guard, poz. 3, 1 February 2012).

out in the examination? Of course, taking into consideration that the average diagnostic value of all polygraph tests is about 90%, that together with the experience of the examiner that has direct contact with the subject, the great majority of expert opinions about the honesty of the subject will be accurate. At least you may assume that it is highly probable. However, an analysis of a hypothetical scenario in a courtroom might be useful²⁸.

Scene 1

Judicial body to an expert: You have issued an NDI opinion that means that no deception was indicated. Are you completely certain that reactions to questions regarding the case recorded during the polygraph test you performed are typical of a truthful person?

Expert: Yes.

Judicial body: Yes? I will repeat. Are you absolutely sure that the recorded on the polygraph diagrams regarding the relevant questions might show only in persons who are 100% truthful. Expert: You know well that almost no forensic analysis method gives a 100% certainty. However, in this particular case there is an extremely high probability that the subject²⁹ showed reactions that are typically associated with a truthful person. In other words, there is only a slight likelihood that the reactions recorded would be shown by a deceitful person."

Judicial body: You confirmed that the method does not give 10% certainty. At the same time you are claiming that there is a slight probability that the opinion you issued is incorrect. In such a case, please present the statistical significant or probability of error with regard to the results of your test.

Expert: I'm sorry. Could you repeat the question? I don't believe I understand....

Judicial body: What is the value or the probability of error of your test in numbers?

Expert: Unfortunately, I cannot answer that. I do not have such data.

Judicial body: Then I have no further questions, Your Honor."

In this scenario, the court is unlikely to use the expert opinion while drafting its ruling.

²⁸ Compare R. NELSON, M. HANDLER, P. SHAW, M. GOUGLER, B. BLALOCK, Ch. RUSSELL, B. CUSHMAN, M. OELRICH, *Using the Empirical Scoring System, Polygraph*, 2011, 40 (2).

²⁹ Defendant or witness.

That situation might have taken a different turn. The above-mentioned evaluation system, ESS, might be useful in solving the problem. Let's assume that a federal ZCT³⁰ was used in the scenario (single issue). The test is evaluated based on the ESS system and the total result is +12. According to the ESS rules for comparing zones, when the total result is $\geq +2$, an NDI opinion is issued (no deception indicated). Next, you consult the appropriate probability consulted, and find the +12 values in the column showing the truthfulness of the subject.

Table 2. Table showing the probability of error for various results of a test data analysis according to the ESS system in examinations including all ZCT tests with three relevant questions.

DECISION THRESHOLDS- TRUTHFULLNESS		DECISION THRESHOLDS- DECEPTION	
Based on the distribution of total scores indicating deception			
decision threshold	p value	decision threshold	p value
-1	0.159	1	0.159
0	0.130	0	0.127
1	0.106	-1	0.099
2	0.085	-2	0.077
3	0.067	-3	0.058
4	0.052	-4	0.043
5	0.040	-5	0.032
6	0.030	-6	0.023
7	0.023	-7	0.016
8	0.017	-8	0.011
9	0.012	-9	0.008
10	0.008	-10	0.005
11	0.006	-11	0.003
12	0.004	-12	0.002
13	0.003	-13	0.001
14	0.002	-14	<0.001
15	0.001		
16	<0.001		

Source: R. Nelson, *ML Handler, Empirical Scoring System: NPC Quick Reference, Lafayette, Instrument 20120.*

³⁰ Comparison test of the areas of the U.S. Federal Government.

According to statistical data contained in the Table, in the case of an NDI opinion based on a total score of +12 means that the statistical significance for this test result is 0.004. Therefore, there is probability of deception during the test is merely 0.4%. In other words, the probability of truthfulness is as much as 99.6%, If the expert had had such data at hand in the courtroom, the scene would finish differently.

Scene 2

Judicial body: What is the statistical significance or probability of error with respect to the score of the test performed manually?

Expert: On the basis of standard data and the Empirical scoring system supported, by scientific evidence of the numerical test data analysis for the psychophysiological detection of deception, the statistical significance of probability of error (p value) for this test would be 0.004. In other words, the chance of such a score being achieved by a deceitful person is 4 to 1000, i.e. 1 in 250 people (0.4%)”

Judicial body: I have no further questions, Your Honour.

This example shows how accurate the opinions of a contemporary expert in polygraph testing can be. An expert has an array of recognized techniques and systems of polygraph scores evaluation at hand, which give them a possibility of effective verification of numerous problems with various complexity levels. It needs to be noted that while using scientific proofs gives, polygraphists can confirm the validity of and reliability their scores in the eyes of the receivers. Let us hope it becomes a Polish standard.