



*Meta-Analytic Survey of Criterion Accuracy  
of Validated Polygraph Techniques. Report Prepared  
for the American Polygraph Association Board of Directors,  
“Polygraph” Special Edition, 2011, vol. 40, no. 4*

The report of the American National Research Council<sup>1</sup> published in 2003 contains a number of highly critical remarks concerning the theoretical premises and the practice of polygraph examinations. Among the results of this criticism was a strategic turn in the operations of the American Polygraph Association, focused on adaptation of the achievements of forensic science, and especially on “seeking to increase the level of science in our practices, to standardize our methodologies, to focus on continuous improvement, to upgrade our education, and broaden our vision to cover not only the interests of members, but to include protection of the public, as well”, as Pamela Shaw, the President of the American Polygraph Association, wrote in the introduction to the work in question. Again, it turned out that nothing has an equally positive impact on the development of a field of knowledge and the ensuing practice as a disinterested but also uncompromising criticism of dated solutions, legacy system weaknesses, and common errors.

---

<sup>1</sup> *The Polygraph and Lie Detection* (2003), Committee to Review the Scientific Evidence on the Polygraph, National Research Council of the National Academies, The National Academies Press, Washington DC.

The *Meta-Analytic Survey of Criterion Accuracy...* conforms to this current of modernisation of polygraph examinations, or – strictly speaking – to the effort to improve their scientific level **by standardisation of methodology of examinations conforming to modern requirements**. Such a point of view is most proper; the practical application of methods and techniques that have not been subjected to validation, with unknown accuracy and other factors influencing their diagnostic capacity, is simply **impermissible**, and experts (“experts”?) applying their own ideas for tests, own “modifications”, and own solutions, unfounded on thorough research, with delayed publication of results, should be situated beyond the realm of professionals. Still, it is about this that the authors of the report write clearly and openly, absolutely unambiguously defining one of the goals behind the origin of the report: “The goal is to eliminate the use of un-standardized, non-validated or experimental techniques in field settings where decisions may affect individual lives, community safety, professional integrity, and national security”. This sentence is all the more worthy of emphasis as it highlights the moral aspect of expert activity in polygraph testing, which in these days of ethical crisis in forensic sciences is especially material.

The report, as its authors state, is actually a comprehensive and, let us add, highly insightful study of the literature of the subject, conducted using well-defined criteria. Taken into account are both the analyses of results of experiments of laboratory studies and analysis of practical examinations, and – as it turned out – the differences in results acquired in their course are small or at least statistically insignificant, which may come as a surprise for many opponents of the use of the polygraph.

Attention was paid to the fact “that event-specific diagnostics examinations conducted for evidentiary purposes, for which it is expected that the results may be used as evidence in a judiciary proceeding, should be conducted using techniques that produce a criterion accuracy level of .900 or higher, excluding inconclusive, and with an inconclusive rate of .200 or lower. Diagnostic examinations conducted using the paired testing protocol can achieve a very high accuracy rate through the combination of results from examinations conducted with techniques that produce a mean criterion accuracy level of .860 or higher, excluding inconclusive, and with inconclusive rate of .200 or lower. Examinations conducted for investigative purposes should be conducted with techniques that produced a mean criterion accuracy level of .800 or higher, excluding inconclusive, and with inconclusive rate of .200 or lower.” (p. 206).

Furthermore, accounted for were only the analyses that concerned the same study techniques and were disclosed in at least two independent publications, while the analysed material was acquired in line with the requirements of the APA Standards of Practice. Altogether, the meta-analysis covered 38 publications that altogether investigated 3723 examinations, including 2015 corroborated opinions indicating examinee deception (DIs), and 1708 corroborated opinions indicating no examinee deception (NDIs) (p. 213). Compared to similar meta-analyses made in the past, these numbers are impressive and provide a testimony to the fact that the discussed work is a milestone that cannot be disregarded in the development of polygraph examinations.

Technically speaking, the research problems which the report was to find answers to (and did) include:

“1) which PDD examination techniques have published and replicated evidence of validity that satisfies the APA 2012 Standards of Practice requirement for decision accuracy and inconclusive rates

2) what is the overall accuracy of validated PDD techniques interpreted with the assumption of independence among the RQ stimuli

3) what is the accuracy level of PDD techniques interpreted with the assumption of non-independence, among the RQ stimuli

4) are there significant differences or outliers among any of the validated PDD techniques

5) are there any outlier results that are not accounted for by the presently available evidence.” (p. 213).

It is clear that these are elementary questions of the “what is the situation like” type, generally aiming to acquire a description of the phenomena that the researchers find interesting. There is nothing wrong with this “elementary character”; on the contrary, such a corroboration of the full picture of the basic facts serves directly the optimisation of practice and also points to the problems that should be elucidated in future, on the basis of further research.

The core of the report is a discussion of the PDD techniques analysed successively. Thus, the reader may become familiar with the results of studies concerning the following solutions: AMFGQT (interpreted in two variants:

on a seven-point scale, and through ESS); Backster You-Phase; Concealed Information Test; Directed-lie Screening Test (also interpreted on a seven-point scale and through ESS); Federal You-Phase (interpreted analogously); Federal ZCT; Integrated Zone Comparison Technique; Matte Quadri-track Zone Comparison Technique; Utah ZCT – Probable Lie Test; Utah ZCT – Directed Lie Test; Utah ZCT – Canadian Police College/Royal Canadian Mounted Police Version; Utah ZCT – Combined PLT, DLT and RCMP; and the Event Specific ZCT.

Concise information concerning each of the techniques covers its general characteristics, the sources of data covered by the meta-analyses, calculations of standard deviation, and variance analyses (usually the two-way ANOVA), mentioning the potential statistical significance of the results (albeit at different levels of  $p$ ), and – in certain cases – the correlation coefficient specific for the given calculations technique. Every such description is also accompanied by a chart, in a layout “similar” to the Cartesian one, illustrating the results surveyed. A critical remark is due here: these charts are made in a non-professional manner, if not erroneously: in none is the horizontal axis described (and the vertical one is described imprecisely, so that the reader must guess what is being charted), yet the primary error is the use of lines in the charts, which suggest the presence of some form of continuity between the studies examined, which is obviously false. In this case, histograms should have been used for visual illustration.

It is also a pity that the report lacked space for quoting the formats of the discussed PDD methods, even despite the conclusion that they are not the most important (see below), as the volume of the publication would not have increased excessively, while the entire document would have gained significantly in self-sufficiency.

Nevertheless, the analyses quoted are highly interesting, and in certain cases contain comments that are found only sporadically in literature. Such precious remarks include, for example, the portrayal of differences in CIT precision, depending on the calculations accounting only for the diagnoses of people concealing information, and earlier, *tempore criminis*, behaviourally involved in the commitment of the crime, or accounting for the diagnoses of people who – although in the possession of key information – did not participate in the criminal activity. If the data is treated jointly, the precision of the CIT technique drops below a value of .800 (p. 218).

Deserving of a full and separate analysis is Appendix G of the report, which contains information explaining why some of the materials published in the past **could not be used** in the study. The arguments used are highly varied, and include antiquated methods of data interpretation, limiting the interpretation solely to the computer method, lack of statistical aggregation of the results, non-standard order of questions in the tests, failure to account for the results from some channels (e.g. cardio) in the results, the lack of representativeness of the analysed sample, divergence from the standards in the given method, and many others. There are no fewer than 39 materials published earlier (of which number 12 come from the 21st century), that were not accounted for in the report (with each being given the reason for rejection). Possibly, some of the negative decisions are disputable, yet the review of the rejects is certainly **a priceless collection of methodological guidelines** for all those planning and carrying out scientific projects covering empirical studies in polygraph examinations.

The report is complemented with carefully elaborated collections of statistical data used for analysis.

Finally, it is worth quoting the most important recommendations of the report: "Because no significant differences were found among the 14 PDD techniques included in this meta-analysis, no attempt should be made to describe these techniques in terms of a rank order regarding effectiveness. Available evidence does not support any PDD techniques as superior to others. Attempts at establishing any hierarchy of efficacy are therefore unwarranted. Instead, less attention should be given to named PDD techniques and meaningless differences in PDD test formats. More emphasis should be given to test construction details for which there is replicated evidence of their contribution to criterion accuracy. More emphasis should be given to the important practical and decision theoretic differences in PDD techniques for which the RQs are interpreted as independent or non-independent." (p. 257).

The authors also list specific problems which will have to be tackled in the future, mentioning among them the influence of variables including the characteristics of the examinees (juvenile, elderly, psychiatric patients, people with other medical problems), and recommending extensive use of statistical methods for the solution of difficult research problems (p. 257).

Certainly, the *Meta-Analytic Survey...* is an outstanding achievement of the APA analysts, while the authors do not show any semblances of triumphalism,

and the whole report is highly balanced, with the theses well defined and justified. Finally, it can be assumed that the document is evidence of a time in the development of polygraph examinations. If it were to be compared to similar works from previous years,<sup>2</sup> there would be visible progress providing grounds for deep satisfaction to all those dealing with polygraph examinations. Which is why it would be difficult to find more suitable words than those of Pamela Shaw from the Introduction to the report: “We are at a great time in polygraph history and we can be proud of the steps we are taking to move our profession forward. We must all grow with the knowledge in our field and the demands within our field to ensure our future success.”

Jerzy Konieczny\*

---

<sup>2</sup> E.g.: N. Ansley (1997), *The Validity and Reliability of Polygraph Testing*, Forensic Research Inc., Severna Park.

\* jerkonieczny@wp.pl