

Application of Computer Polygraph in Forensic Psychological Expert Research During Pre-Trial Investigation of Criminal Offenses*

Oleksandr Yevsiukov,¹ Yuliia Sukhomlyn,^{2**} Serhii Mytych,³

Borys Lukianchykov,⁴ Olena Riashko⁵

¹ Bokarius Forensic Science Institute of the Ministry of Justice of Ukraine

^{2,3} National Academy of Internal Affairs, Ukraine

⁴ National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"⁶ Kyiv, Ukraine, ⁵ Lviv State University of Internal Affairs, Ukraine



[10.15408/jch.v14i2.53584](https://doi.org/10.15408/jch.v14i2.53584)

Abstract

The article examines the methodological and procedural conditions under which the use of a computer polygraph within the framework of forensic psychological expert research at the stage of pre-trial investigation can be considered scientifically justified and procedurally acceptable. The study was implemented in the format of a sequential mixed design: in stage A, a quantitative content analysis of 60 depersonalized conclusions/reports of polygraph examinations in criminal proceedings was carried out, and in stage B, a two-wave structured survey of the expert community (N = 110; polygraph examiners, forensic psychologists/experts, investigators/prosecutors, lawyers) was conducted to form a consensus on a minimum standard. The quality of documentation was operationalized through the MPI and the PCI. The reports were found to be moderately transparent (MPI M = 44.2) and marginally compliant (PCI M = 38.4), with moderate correlation between the indices ($r = .54$; $p < .001$), indicating incomplete agreement between the methodological and procedural components. Systemic deficiencies in documentation were identified, including the lack of alternative explanations for responses, countermeasures, and evidence of test invalidity; CIT/GKT documents showed higher MPI and PCI than CQT. Phase B results showed an acceptable level of agreement (Kendall's $W = .74$), and a consensus minimum standard of 11 criteria was developed. A comparison of the stages revealed a significant gap between practice and the normatively expected standard, which outlines priorities for unifying documentation and increasing the evidentiary reliability of results.

Keywords: Computer polygraph; Forensic psychological examination; Pre-trial investigation

* Received: April 11, 2026; revised: May 21, 2026; accepted: May 22, 2026; Published: June 30, 2026.

¹**Oleksandr Yevsiukov.** Candidate of Pedagogical Sciences, Associate Professor, Leading Scientific Employee, Forensic Psychological Expert, National Scientific Center "Hon. Prof. M. S. Bokarius Forensic Science Institute" of the Ministry of Justice of Ukraine, Kharkiv, Ukraine, Email: a.evsiukov1971@gmail.com ORCID ID: <https://orcid.org/0000-0002-3554-8848>

²**Yuliia Sukhomlyn.** Ph.D. in Law, Associate Professor, Department of Criminal Procedure of the National Academy of Internal Affairs, Kyiv, Ukraine, Email: sjulia25@ukr.net ORCID ID: <https://orcid.org/0000-0002-5699-6914> (Corresponding author)

³**Serhii Mytych.** Ph.D. in Law, Associate Professor, Department of Criminal Procedure of the National Academy of Internal Affairs, Kyiv, Ukraine, Email: SergeyMytych@meta.ua ORCID ID: <https://orcid.org/0009-0004-4426-9543>

⁴**Borys Lukianchykov.** Ph.D. in Law, Associate Professor, Department of Information, Economic and Administrative Law, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"⁶ Kyiv, Ukraine, Email: boryn1971@gmail.com, ORCID ID: <https://orcid.org/0000-0002-4761-5980>

⁵**Olena Riashko.** Candidate of Law, Associate Professor, Department of Criminal Procedure and Criminalistics, Lviv State University of Internal Affairs, Lviv, Ukraine, Email: olena_ryashko@ukr.net ORCID ID: <https://orcid.org/0000-0001-5550-5223>

** Corresponding Author: sjulia25@ukr.net

A. INTRODUCTION

In modern criminal justice, the growth of the function of psychological and psychophysiological methods for evidence at the stage of pre-trial investigation is noticeable. One of the most controversial tools is the use of a computer polygraph, which is used to record and interpret a person's psychophysiological reactions to significant stimuli. Despite the active use of this tool in the practice of investigative bodies, its methodological status in the system of forensic psychological examination remains ambiguous.

The relevance of the problem can be explained by several modern trends. In particular, the digitalization of expert activity and the standardization of technical means of recording psychophysiological indicators have become noticeable (Baddeley et al., 2023; Funicelli et al., 2021). Such a situation can formally increase the accuracy of measurement; however, it will not solve the issue of interpretation and evidentiary significance of the results. On the other hand, in the practice of pre-trial investigation, the results of polygraph examinations are often used to make tactical decisions and evaluate testimony (Horsley, 2022; Iacono, 2024). However, the authors indicated that the level of documentation of such procedures, interpretation criteria, and procedural guarantees varies significantly (Leo et al., 2024; Malewski & Kraujalis, 2024). Finally, in the scientific field, there are conflicting data on the validity of various polygraph techniques (Wolsink et al., 2024). Therefore, the problem is of a scientific and practical nature and concerns the limits of the use of psychodiagnostic tools in evidentiary activities.

Finally, it is worth paying attention to the fact that two approaches are noticeable in the current scientific discourse. Some scientists have identified the polygraph as an important auxiliary psychophysiological tool without independent evidentiary value. Other scientists have considered the polygraph as a special method of expert research, which is aimed at making independent conclusions within the framework of forensic psychological examination. However, the analytical contradiction lies in the fact that forensic psychological examination provides clear criteria for reproducibility; however, the real practice of using the polygraph is determined by the variability of procedures and differences in documentation. Given the previous thoughts, the question remains unresolved:

What methodological and procedural conditions ensure the scientific validity and procedural acceptability of the use of a computer polygraph in forensic psychological expert research during pre-trial investigation?

Previous studies have also mainly focused on the technical characteristics of the polygraph. Other scientists have assessed the accuracy of individual test methods. Some authors have analyzed its effectiveness within operational activities, rather than forensic psychological expertise. This forms an important scientific gap it becomes noticeable that the authors have not systematized the level of actual documentation of procedures in practice. Therefore, there is a gap in the form of the absence of a systematic, empirically substantiated standard of documentation, interpretation, and procedural guarantees for the use of polygraph within forensic psychological expert research.

Accordingly, the aim of the article is to empirically substantiate the minimum methodological and procedural standard for the use of computer polygraph in forensic psychological expert research during pre-trial investigation.

RQ1. What level of methodological transparency and procedural compliance characterizes polygraph studies performed at the stage of pre-trial investigation, and what structural deficiencies in documentation are systemic?

RQ2. Does the type of polygraph methodology used (CIT/GKT vs CQT) affect the level of methodological transparency, interpretative certainty, and procedural compliance of documentation?

RQ3. What criteria for documentation, interpretation, and procedural guarantees are recognized by the expert community as mandatory for the minimum standard for the use of computer polygraph in forensic psychological expert research?

RQ4. Is there a gap between actual documentation practice (stage A) and the consensus minimum standard defined by the expert community (stage B), and which components constitute the area of greatest normative and methodological imbalance?

Literature Review

In modern scientific debate, the polygraph has been analyzed as a “lie detector” in the literal sense. That is, it was a tool for recording psychophysiological reactions, the interpretation of which depends on the testing procedure, evaluation criteria, and quality control (Yefimenko et al., 2023; Shalash et al., 2025). Generalizations of research on the problem of lie detection emphasized the limitations of universal solutions: the effectiveness of methods varied according to the individual context, task, and vulnerability to

intentional distortion of reactions (Quattrocolo, 2025). This situation immediately shifted the focus from “the accuracy of the instrument” to the methodological correctness of the procedure and the reproducibility of conclusions (Oswald, 2020).

Some of the most heated discussions revolved around the Comparison/Control Question Test (CQT). Meta-analytic works confirmed the existence of an evidence base in favor of the CQT in certain conditions, while at the same time pointing out the risks of systematic errors of assessment and overestimation of validity due to methodological errors of generalization (Mykityshyna, 2024). In this context, it is important that criticism often concerned the “effect”, transparency of scoring, rules of interpretation and control of countermeasures, which in real practice could be described incompletely or unsystematically (Lukács & Ansorge, 2023; Knappe et al., 2025).

At the same time, the Concealed Information Test / Guilty Knowledge Test (CIT/GKT) was more often described as a structurally more rigid procedure, where the logic of the stimuli and (potentially) quantitative criteria made the results more standardized and suitable for verification (Honts et al., 2021). In parallel, modifications of the CIT based on reaction time were developed, with analysis of applicability even in “field” samples. Researchers also turned to several psychophysiological approaches (ERP/P300), with meta-analytic generalizations that demonstrated promise (Gordon & Fleisher, 2019; Elton, 2017). However, this did not remove the requirements for procedural standardization and quality control.

A separate cluster of literature is devoted to the vulnerability of procedures to faking/countermeasures and the need for their control. For reaction-time versions of the CIT, it has been shown that artificial distortion is possible and requires special countermeasure protocols (Brennen & Magnussen, 2023; Ben-Shakhar & Iacono, 2021). For further studies, this approach would mean that the conclusion/report does not record how countermeasures were tested, how signs of test invalidity were identified, and whether alternative explanations for the responses were considered (Afzali et al., 2022). Accordingly, the scientific verifiability and legal stability of such results have naturally decreased (Goswami & Goswami, 2025). Interdisciplinary reviews have also been useful, emphasizing the complexity of cognitive and memory processes in legal contexts and the risks of simplistic interpretation of behavioral/psychophysiological indicators.

B. METHODS

The study was implemented in the form of a sequential mixed design, in particular, a quantitative content analysis of documents (stage A) and a two-wave structured expert survey (stage B) were combined. The choice of this sequential model is due to the logic of the research question: initially, the study was aimed at establishing the actual state of documentation and interpretation of polygraph examinations in real practice; later, based on the identified deficiencies and variability, a verified professional consensus on minimum standards was formed. Thus, Stage A answered the question “how is it actually documented?”; Stage B answered the question “how should this be documented from the point of view of the professional community?”.

The integration of the results was carried out at the level of forming an agreed minimum standard for the use of a computer polygraph in the structure of forensic psychological expert research during pre-trial investigation.

Content analysis was used as a method of systematic study of text documents to establish procedural and interpretative characteristics of polygraph studies. In contrast to assessing the accuracy of the polygraph as a psychophysiological tool, in this study the main emphasis is placed on the quality of documentation, standardization of procedures, and procedural transparency. These aspects are important for the expert status of the results. The unit of analysis was the structure of the expert document.

The unit of analysis was a depersonalized expert opinion or report of a polygraph study, which was prepared within the framework of criminal proceedings at the stage of pre-trial investigation. The study included 60 documents. This number was chosen considering aspects, ensuring statistical reliability of inter-coder consistency, the possibility of comparing subgroups (type of methodology, institution, etc.), and stability of index indicators.

A targeted document selection strategy was used, since the study involved the analysis of precisely those documents that meet the specified structural criteria. Inclusion criteria: 1. presence of a description of the applied test methodology; 2. indication of the conditions of conduct (time, environment); 3. presence of results and their interpretation; 4. structured design (sections, signatures, appendices). The following are the exclusion criteria: 1. absence of a description of the procedure; only a summary without a protocol; duplicate documents; materials that are not subject to depersonalization.

All documents were depersonalized before coding began. Coding was implemented according to a pre-formed codebook that included six domains. Each domain indicated a separate dimension of methodological and procedural quality.

D1. Methodology and protocol (procedural standardization). The following indicators were evaluated: 1. type of test procedure (CIT/GKT, CQT, mixed, not specified); 2. series structure; 3. presence of pre-test and post-test interviews; 4. wording of questions.

D2. Technical parameters (technical transparency): 1. specification of the model and software; 2. number of recording channels; 3. description of calibration; 4. control of artifacts.

D3. Analytical validity: 1. clarity of scoring criteria; 2. use of quantitative indicators; 3. availability of alternative explanations of reactions.

D4. Quality control (methodological stability): 1. verification of possible countermeasures; 2. signs of test invalidity; 3. repeated series to check stability.

D5. Procedural guarantees (legal correctness): 1. informed consent; 2. voluntary participation; 3. indication of the procedural status of the person.

D6. Data preservation and integrity: 1. recording of access; 2. indication of the location of file storage; 3. availability of graphs and raw records.

Each criterion was evaluated: 0 – absent; 1 – partially present; 2 – fully documented.

This assessment made it possible to avoid binary reduction and consider the degree of detail. The coding of 60 documents was carried out by two independent researchers after a preliminary pilot test (n=5), during which the formulation of the codebook criteria was clarified, and the ambiguity of interpretation was reduced. To assess inter-coder consistency, the following were used: Cohen's kappa (κ) – for nominal variables (type of methodology, presence/absence of procedural elements); Krippendorff's alpha (α) – for multilevel scales (0/1/2); Intraclass Correlation Coefficient (ICC) – for integral indices (MPI, PCI). A value of ≥ 0.70 was defined as acceptable consistency, ≥ 0.80 – as high, ≥ 0.90 – as very high. Three integral indicators were formed based on the code matrix.

Methodological Transparency Index (MPI). This index indicated the completeness and structure of the methodological description of the polygraph examination. It was formed based on domains: D1, D2, D3 and D4.

$$MPI = \frac{\sum_{D1-D4}}{Max(D1-D4)} \times 100$$

The MPI ranged from 0 to 100: 0–39 represented low methodological transparency, 40–69 represented a medium level, and 70–100 represented a high level of standardization.

Procedural Compliance Index (PCI). PCI indicated the compliance of documentation with procedural guarantees. Compiled based on: D5 and D6.

$$PCI = \frac{\sum_{D5-D6}}{Max(D5-D6)} \times 100$$

The interpretation scale is like the MPI. Given the availability of 60 documents, a separate analysis of the quality of documentation was carried out depending on the methodology used. For this purpose, the following were calculated: Average MPI for each methodology and Average PCI for each methodology. Frequency of key deficiencies within each methodology: lack of clear interpretation criteria; lack of description of countermeasures; lack of informed consent; lack of graphic appendices.

Statistical analysis. For the processing of selected cases, t-test or Mann-Whitney U was used to compare MPI and PCI between CIT and CQT; χ^2 to compare the frequency of deficiencies; correlation analysis to determine the relationship between MPI and PCI; and cluster analysis.

Methodological logic of stage B. Stage B had a normative-consensus function and was aimed at substantiating the minimum standard for the use of a computer polygraph in the structure of forensic psychological expert research. For this purpose, a two-wave model of expert assessment with an expanded sample (N=110) was used, which made it possible to ensure statistical stability of the results and compare the positions of professional groups.

Sample type. The study used a quota target sample to represent the main professional groups involved in the criminal process. The sample structure included the following individuals: Polygraph examiners — 35 people; Forensic psychologists/forensic experts — 30 people; Investigators/prosecutors — 25 people; Lawyers — 20 people. Inclusion criteria included at least 2 years of professional experience, practical participation in criminal proceedings or expert research, and voluntary consent to participate.

Additional variables. The following indicators were recorded: length of service (years), number of cases using the polygraph, region of operation, and type of institution. These variables were used as controls in further analysis.

Instrument design. The questionnaire was formed on the basis of the results of content analysis (stage A); international standards of polygraph testing; principles of forensic psychological examination. Thus, the content validity of the instrument was ensured. Block 1. Assessment of minimum standard criteria. Each criterion was assessed on a 9-point scale according to four parameters: 1). Necessity for the minimum standard; 2). Impact on scientific validity; 3). Impact on procedural admissibility; 4). Feasibility in practice. Independent variables: professional group, experience, use of polygraph. Dependent variables: necessity rating, validity rating, admissibility rating, feasibility rating.

Red Flags (critical defects). Participants assessed whether a certain defect was critical (excludes admissibility); significant (reduces trust); or insignificant. This allowed the creation of a Critical Risk Index.

Vignettes (scenario analysis). 2–4 standardized scenarios were developed, varying: type of methodology (CIT vs CQT); presence of countermeasure control; completeness of documentation; presence of chain of custody. Dependent variables: acceptability (1–9), confidence level, risk of error, willingness to include in expert opinion

Two-wave procedure. Wave 1: initial assessment; fixing variability and identifying criteria with low agreement. Intermediate stage: participants were provided with aggregated statistical results; group differences were shown; Wave 2: re-assessment of criteria and reduction of polarization. This approach allowed to increase the stability of the results without losing analytical value. To assess consensus, the following were used: median; interquartile range (IQR); proportion of scores ≥ 7 ; Kendall's W (coefficient of concordance).

Statistical analysis. Since the scales were ordinal, the following were used: Kruskal–Wallis for between-group differences; Mann–Whitney for pairwise comparisons; Kendall's W for consistency. Multiple regression was also used to predict admissibility in vignettes and the effect size (r or η^2).

C. RESULT AND DISCUSSION

Analysis of 60 depersonalized polygraph reports collected during criminal proceedings at the pretrial investigation stage indicated variability in

the quality of documentation and the presence of systemic structural deficiencies.

The overall scores for both indices are presented in Table 1. The mean MPI score was 44.2 (SD = 12.8). This score generally corresponded to the average level of standardization (40–69). At the same time, the distribution of MPI values indicated a skewness to the left (skewness = 0.31). This indicated the absence of documents with a high level of methodological transparency (MPI $\geq$ 70) in the sample: 28.3% (n = 17) of documents received indicators below 40. This indicator indicated a low level. The average score of the Procedural Compliance Index (PCI) was 38.4 (SD = 13.5). This corresponded to the border between low and **medium** levels. The share of documents with a low level of procedural compliance (PCI < 40) had an indicator of 45.0% (n = 27). At the same time, the medium level was recorded in 55.0% (n = 33) of documents; no document reached a high PCI level ($\geq$ 70). A moderate positive correlation was recorded between MPI and PCI ($r = .54$, $p < .001$). This indicated an incomplete alignment between methodological transparency and procedural correctness of documentation.

Table 1.
Overall MPI and PCI indices (n = 60)

Indicator								
MPI (Methodological Transparency Index)	44.2	12 Aug	43.0	16	68	17 (28.3%)	43 (71.7%)	0 (0%)
PCI (Procedural Compliance Index)	38.4	13 May	37.5	9	66	27 (45.0%)	33 (55.0%)	0 (0%)

The author should explain the results of the research (what was discovered) in detail. The research results and discussion section contain the results of the research findings and their ensuing discussions. The findings acquired from the results of the research conducted should be written with the supplementary support of adequate data. The research results and findings should be able to resolve or provide explanations to the question stated in the introduction.

Note: M is the mean; SD is the standard deviation; Mdn is the median. MPI distribution: skewness = 0.31. MPI–PCI correlation: $r = .54$, $p < .001$.

A detailed analysis of the level of documentation for all 20 components of the codebook (domains D1–D6) is shown in Table 2. In particular, the highest indicators of completeness of documentation were recorded in domain D1 (Methodology and protocol). It is noticeable that the type of test procedure was fully described in 56.7% (n = 34) of documents, in 15.0% (n = 9) - partially, which

in total is 71.7% of documents with at least partial information about the methodology. A satisfactory level was also demonstrated by the component of indicating the procedural status of the person (D5). It is noticeable that full presence was recorded in 63.3% (n = 38) of documents, partial — in 20.0% (n = 12). Together, this is 83.3%. In contrast, domains D3 and D4 had a low level of documentation: alternative explanations of reactions were completely absent in 80.0% (n = 48) of documents, and verification of countermeasures was absent in 78.3% (n = 47). However, Domain D2 (Technical parameters) was determined by discrepancies between components. While the model of the device and software was indicated in 75.0% (n = 45) of documents (fully or partially), the description of calibration was absent in 73.3% (n = 44).

Table 2.
Level of documentation by codebook domains (n = 60)

Domain	Component				
D1. Methodology and protocol	Type of test procedure (CIT/CQT)	34 (56.7%)	9 (15.0%)	17 (28.3%)	43 (71.7%)
D1. Methodology and protocol	Structure of stimulus series	18 (30.0%)	21 (35.0%)	21 (35.0%)	39 (65.0%)
D1. Methodology and protocol	Pre-test and post-test interviews	8 (13.3%)	11 (18.3%)	41 (68.3%)	19 (31.7%)
D1. Methodology and protocol	Question/stimulus formulation	22 (36.7%)	17 (28.3%)	21 (35.0%)	39 (65.0%)
D2. Technical parameters	Device and software model	31 (51.7%)	14 (23.3%)	15 (25.0%)	45 (75.0%)
D2. Technical parameters	Number of recording channels	19 (31.7%)	12 (20.0%)	29 (48.3%)	31 (51.7%)
D2. Technical parameters	Calibration description	7 (11.7%)	9 (15.0%)	44 (73.3%)	16 (26.7%)
D2. Technical parameters	Artifact control	6 (10.0%)	11 (18.3%)	43 (71.7%)	17 (28.3%)
D3. Analytical validity	Clear scoring criteria	11 (18.3%)	11 (18.3%)	38 (63.3%)	22 (36.7%)
D3. Analytical validity	Quantitative evaluation indicators	14 (23.3%)	9 (15.0%)	37 (61.7%)	23 (38.3%)
D3. Analytical validity	Alternative explanations of responses	4 (6.7%)	8 (13.3%)	48 (80.0%)	12 (20.0%)
D4. Quality control	Checking countermeasures	3 (5.0%)	10 (16.7%)	47 (78.3%)	13 (21.7%)
D4. Quality control	Signs of test invalidity	5 (8.3%)	8 (13.3%)	47 (78.3%)	13 (21.7%)
D4. Quality control	Repeated series/stability	9 (15.0%)	14 (23.3%)	37 (61.7%)	23 (38.3%)
D5. Procedural safeguards	Informed consent	15 (25.0%)	13 (21.7%)	32 (53.3%)	28 (46.7%)
D5. Procedural safeguards	Voluntariness of participation	21 (35.0%)	14 (23.3%)	25 (41.7%)	35 (58.3%)

D5. Procedural safeguards	Procedural status of the individual	38 (63.3%)	12 (20.0%)	10 (16.7%)	50 (83.3%)
D6. Data retention	Recording of access to materials	9 (15.0%)	12 (20.0%)	39 (65.0%)	21 (35.0%)
D6. Data retention	File storage location	12 (20.0%)	11 (18.3%)	37 (61.7%)	23 (38.3%)
D6. Data retention	Graphs/raw recording records	10 (16.7%)	13 (21.7%)	37 (61.7%)	23 (38.3%)

Missing $\geq 70\%$ — critical deficiency Missing 50–69% — significant deficiency

Note: Verification: fully + partially + missing = 60 for each row. Components missing in $\geq 50\%$ of the documents were systemically missing.

Table 3 shows the components that were found to be systemically missing (in $\geq 50\%$ of the documents analyzed). A total of 12 system documentation deficiencies were identified, of which three were classified as critical: lack of alternative explanations for reactions (80.0%, $n = 48$), lack of verification of countermeasures (78.3%, $n = 47$), and lack of signs of test invalidity (78.3%, $n = 47$). Significant deficiencies included the lack of a description of calibration, artifact control, pre-test and post-test interviews, clear scoring criteria, recording access to materials, graphs and raw registration records, repeated series to check stability, and quantitative assessment indicators.

Table 3.
Systematic documentation deficiencies
(absent in $\geq 50\%$ of documents), sorted by frequency

№	Lack of documentation	Domen	n of absense	% absense	Qualification
1	Alternative explanations for reactions	D3	48	80.0%	Critical
2	Checking for countermeasures	D4	47	78.3%	Critical
3	Signs of test invalidity	D4	47	78.3%	Critical
4	Calibration description	D2	44	73.3%	Substantial
5	Control of artifacts	D2	43	71.7%	Substantial
6	Pre-test and post-test interviews	D1	41	68.3%	Substantial
7	Recording access to materials	D6	39	65.0%	Substantial
8	Clear scoring criteria	D3	38	63.3%	Substantial

№	Lack of documentation	Domen	n of absense	% absense	Qualification
9	Graphs/raw registration records	D6	37	61.7%	Substantial
10	Repeated runs/stability	D4	37	61.7%	Substantial
11	Quantitative assessment metrics	D3	37	61.7%	Substantial
12	Informed consent	D5	32	53.3%	Qualification

Note: Critical — a deficiency that made it impossible to reproduce or verify the results. Significant — a deficiency that reduced validity or procedural acceptability. Deficits cover all 6 domains of the codebook.

Intercoder agreement indicators confirmed the reliability of the data shown. Cohen's κ for nominal variables was 0.76–0.84, and Krippendorff's α for multilevel scales was 0.71–0.79. All indicators exceeded the acceptable threshold of 0.70, which confirms sufficient reliability of the coding procedure.

To determine the influence of the method on the MPI and PCI indicators, a comparative analysis was performed. Documents in which the type of method was not specified, or a mixed procedure was used were excluded from the comparative analysis ($n = 6$; 10.0%). The distribution of documents by type of methodology is presented in Table 5.

Table 5.
Distribution of documents by type of polygraph methodology ($n = 60$)

Method type	n	%	Included in the comparative analysis RQ2
CIT / GKT (Concealed Information Test / Guilty Knowledge Test)	26	43.3%	Yes
CQT (Control Question Test)	28	46.7%	Yes
Mixed procedure	4	6.7%	No
Type not specified	2	3.3%	No
Total	60	100%	—

Note: The comparative analysis of RQ2 included $n = 54$ documents (CIT/GKT + CQT). Documents with mixed procedures and unspecified type ($n =$

6) were excluded.

The comparison of the integral indices showed some differences. Documents prepared using CIT/GKT had a higher mean MPI score ($M = 49.1$, $SD = 11.4$) compared to CQT documents ($M = 40.3$, $SD = 13.1$); the difference was significant ($U = 218.5$, $p = .003$, $r = .42$). At the same time, we recorded a similar trend for PCI: the mean score for CIT/GKT was 43.2 ($SD = 12.7$) versus 34.6 ($SD = 13.9$) for CQT ($U = 231.0$, $p = .008$, $r = .38$).

A detailed analysis of the frequency of key deficits by methodology is presented in Table 6. Significant differences were found between the CIT/GKT and CQT groups for four of the eight components analyzed. The largest difference was recorded in the presence of clear scoring criteria. In the CIT/GKT group this component was absent in 50.0% ($n = 13$) of documents, while in the CQT group it was absent in 75.0% ($n = 21$); $\chi^2(1) = 4.08$, $p = .043$. A significant difference was also found in the presence of quantitative assessment indicators. Regarding the verification of countermeasures and signs of test invalidity, the differences between the groups were statistically insignificant ($p > .05$).

Table 6.
Comparison of the frequency of key deficits between the CIT/GKT and CQT ($n = 54$)

Lack of documentation	CIT/GKT absence, n (%)	CQT absence, n (%)	χ^2	p	Significance
Clear scoring criteria	13 (50.0%)	21 (75.0%)	4.08	.043	✓ Significant
Quantitative assessment indicators	12 (46.2%)	20 (71.4%)	3.94	.047	✓ Significant
Pre-test and post-test interviews	15 (57.7%)	22 (78.6%)	3.02	.082	— Trend
Alternative explanations for reactions	20 (76.9%)	24 (85.7%)	0.77	.381	— n/a
Checking for countermeasures	18 (69.2%)	24 (85.7%)	2.18	.140	— n/a
Signs of test invalidity	18 (69.2%)	24 (85.7%)	2.18	.140	— n/a
Informed consent	13 (50.0%)	16 (57.1%)	0.27	.612	— n/a
Graphs/raw registration records	15 (57.7%)	19 (67.9%)	0.65	.421	— n/a

Note: n/a — statistically insignificant ($p > .05$). χ^2 with Yates correction for 2x2 tables. Check sums: CIT/GKT n=26, CQT n=28 for each row ✓

Therefore, the type of polygraph technique used is related to the level of methodological transparency and procedural compliance of documentation. As can be seen from the data, CIT/GKT documents had higher quality according to integral indices.

The results of the expert survey regarding the minimum standard of computer polygraph use in forensic psychological expert research are presented in Tables 7 and 8. In particular, according to data from wave 1, the variability of assessments was highest for the criteria of analytical validity (IQR = 3.0–4.0) and quality control (IQR = 2.5–3.5). However, procedural guarantees received relatively stable assessments already at the first stage (IQR = 1.5–2.5). After an interim stage, in which participants were provided with the aggregated results of wave 1 and shown intergroup differences, the variability of ratings in wave 2 significantly decreased for most criteria. The overall Kendall's W concordance coefficient increased from 0.52 (wave 1) to 0.74 (wave 2).

The ratings of the criteria by evaluation parameters, in particular, the indicators of necessity, validity, acceptability, feasibility) are presented in Table 8. The minimum standard includes criteria that received a median of ≥ 7 for the parameter "necessity" and the share of ratings ≥ 7 of at least 70%. At these thresholds, 11 criteria out of 20 proposed were included in the consensus minimum standard. The highest ratings of necessity were received by: indication of the procedural status of the person, availability of consent, description of the applied test methodology, and the presence of clear scoring criteria. The lowest feasibility ratings at high need were recorded for the calibration description (Mdn need = 8.0; Mdn feasibility = 5.0).

Table 7.
Minimum standard criteria ratings based on wave 2 results (N = 110)

Criterion	Domen	Necessary Mdn (IQR)	Validity Mdn (IQR)	Admissibility Mdn (IQR)	Realisation Mdn (IQR)	Marks ≥ 7 , %
Procedural status of the individual	D5	9.0 (1.0)	8.0 (1.5)	9.0 (1.0)	8.0 (1.5)	94.5%
Informed consent	D5	9.0 (1.5)	8.0 (2.0)	9.0 (1.0)	8.0 (2.0)	91.8%
Description of test methodology	D1	8.0 (1.5)	8.0 (1.5)	8.0 (2.0)	8.0 (1.5)	88.2%
Clear scoring criteria	D3	8.0 (2.0)	8.0 (2.0)	7.0 (2.0)	7.0 (2.5)	83.6%

Voluntariness of participation	D5	8.0 (1.5)	7.0 (2.0)	8.0 (1.5)	8.0 (1.5)	82.7%
Question/stimulus formulation	D1	7.0 (2.0)	8.0 (1.5)	7.0 (2.0)	7.0 (2.0)	79.1%
Quantitative assessment indicators	D3	7.0 (2.0)	8.0 (2.0)	7.0 (2.5)	6.0 (2.5)	76.4%
Alternative explanations for responses	D3	7.0 (3.0)	8.0 (2.0)	7.0 (2.5)	6.0 (3.0)	74.5%
Checking countermeasures	D4	7.0 (2.5)	7.0 (2.5)	7.0 (2.0)	6.0 (2.5)	72.7%
Graphs/raw recordings	D6	7.0 (2.0)	7.0 (2.0)	7.0 (2.0)	7.0 (2.0)	71.8%

Note: Scale 1–9. Mdn — median of wave 2; IQR — interquartile range of wave 2. To standard (✓): Mdn of necessity ≥ 7 and scores ≥ 7 at least 70%. Kendall's W wave 1 = .52; wave 2 = .74.

Intergroup differences in the assessments of the necessity of the criteria between the four professional groups were tested using the Kruskal–Wallis test. Statistically significant differences were found for five of the eleven criteria included in the minimum standard. The largest discrepancies were recorded for the criterion "availability of alternative explanations of reactions": polygraph examiners rated it significantly lower (Mdn = 6.0) compared to forensic psychologists (Mdn = 8.0) and lawyers (Mdn = 9.0); $H(3) = 18.4$, $p < .001$, $\eta^2 = .17$. Significant differences were also found for the criterion "checking countermeasures": investigators and prosecutors provided lower assessments of necessity (Mdn = 6.0) compared to polygraph examiners (Mdn = 8.0) and forensic psychologists (Mdn = 8.0); $H(3) = 12.7$, $p = .005$, $\eta^2 = .12$. However, for procedural guarantees, informed consent, and procedural status of the individual, the intergroup differences were statistically insignificant ($p > .05$).

Comparison of actual practice with the consensus standard indicated a gap between what the expert community considers mandatory and what is actually recorded in polygraph reports. Of the 11 criteria included in the minimum standard (Mdn of necessity ≥ 7 ; scores ≥ 7 at least 70%), none reached the level of actual availability of 90% or higher. Only two criteria demonstrated a relatively acceptable level of compliance: procedural status of the person (actual presence 83.3% with a consensus necessity Mdn = 9.0) and informed consent (actual presence 46.7% with a Mdn = 9.0). The gap for informed consent is 44.2 percentage points, the largest among the procedural criteria. The largest overall gap was recorded for the criteria of analytical validity (domain D3) and quality control (domain D4).

Imbalance zone	Criterion	Domen	Actual presence, % (stage A, n=60)	Consensus: scores ≥ 7 , % (stage B, N=110)	Gap, pp.	Elimination priority
Critical (≥ 50 pp)	Alternative explanations for reactions	D3	20.0%	74.5%	54.5	Highest
Critical (≥ 50 pp)	Checking for countermeasures	D4	21.7%	72.7%	51.0	Highest
Significant (30–49 pp)	Clear scoring criteria	D3	36.7%	83.6%	46.9	High
Significant (30–49 pp)	Informed consent	D5	46.7%	91.8%	44.2	High
Significant (30–49 pp)	Calibration description	D2	26.7%	70.9%	43.6	High
Significant (30–49 pp)	Quantitative assessment indicators	D3	38.3%	76.4%	38.1	High
Significant (30–49 pp)	Indications of test invalidity*	D4	21.7%	58.2%	36.5	High
Significant (30–49 pp)	Graphs/raw registration records	D6	38.3%	71.8%	33.5	High
Acceptable (<30 pp)	Voluntariness of participation	D5	58.3%	82.7%	24.4	High
Acceptable (<30 pp)	Description of test methodology	D1	71.7%	88.2%	17.2	Average
Acceptable (<30 pp)	Procedural status of the individual	D5	83.3%	94.5%	10.7	Average

Thus, there is a noticeable systemic gap between the actual practice of documenting polygraph examinations and the consensus minimum standard defined by the expert community. The area of greatest normative and methodological imbalance contains the components of analytical validity and quality control. Thus, the results obtained indicated that polygraph reports conducted at the pre-trial investigation stage were characterized by an average level of methodological transparency (MPI M = 44.2) and a marginal level of procedural compliance (PCI M = 38.4). Such data can generally be compared with the conclusions of previous studies, which recorded significant variability

in the quality of documentation of polygraph procedures. For example, similar trends were described by other authors within the framework of the analysis of forensic practice of countries with a continental legal tradition (Blikhar et al., 2022; Husieva et al., 2022). The authors also indicated that polygraph testing functions mainly as an operational mechanism that has lower requirements for standardization of documentation (Kotsoglou & Biedermann, 2024). At the same time, the data obtained expanded the existing knowledge, since previous studies mostly focused on the technical accuracy of polygraph techniques.

A particularly valuable result is that only a moderate correlation was recorded between MPI and PCI ($r = .54$). This indicator indicated a structural disconnect between the methodological and procedural components of documentation. Thus, even technically more detailed reports may not necessarily provide the appropriate level of procedural guarantees. We can explain this gap by differences in the training of specialists. As other scientists have proven, polygraph examiners have undergone specialized technical training, but their legal training regarding the procedural requirements of forensic psychological examination is significantly less (Lukács et al., 2020; Mergaerts & Van Daele, 2023).

Among the identified systemic problems, the lack of alternative explanations for reactions and verification of countermeasures had the greatest theoretical and practical significance. The problem of the lack of alternative explanations has also become important from a methodological point of view, since psychophysiological reactions recorded by a polygraph are not specific indicators of deception. At the same time, the process of ignoring alternative interpretations in documents can affect the violation of the basic principle of scientific validity of conclusions and reduce their evidentiary value. Similarly, the lack of verification of countermeasures will make it impossible to assess the reliability of the testing procedure itself. In addition, the revealed advantage of CIT/GKT documents over CQT documents in terms of MPI indicators was consistent with the theoretical differences between these methods. Modern scientists have indicated that CIT/GKT serves as a structurally more rigid procedure with a clearly defined logic of stimuli and quantitative criteria for evaluating reactions (Nagale & Khandare, 2025; Ukhno, 2021). According to such logic, this can facilitate standardized documentation. However, on the other hand, CQT requires more variability in the formulation of questions and interpretation of reactions (Yukhno & Turenko, 2025). This situation may result in lower reproducibility and, consequently, is more difficult to document in a uniform manner. Such data have also been shown in other works.

At the same time, it is fundamentally important that the largest differences were found between the groups in the components of analytical validity (D3). However, deficiencies in quality control (D4) and procedural guarantees (D5) became systemic in both groups. Countermeasures verification is absent in 69.2% of CIT/GKT documents and 85.7% of CQT documents. This situation is understandable even though countermeasures control is methodologically important for CIT/GKT. Other authors have proven that the effectiveness of the technique depends on the lack of prior knowledge of significant stimuli by the subject (Suchotzki et al., 2021; Rothmann & Mayer, 2025).

The study shows that an acceptable level of consensus (Kendall's $W = .74$) was achieved on 11 criteria of the minimum standard after two waves of the survey. This situation indicated the presence of a common professional understanding of the basic requirements for a polygraph examination despite the differences in the functional roles of the participants. It is particularly telling that the most stable consensus was formed around procedural guarantees, as well as the procedural status of the person and consent. However, the results also show that the largest interprofessional differences were found regarding alternative explanations of reactions ($H(3) = 18.4$, $p < .001$, $\eta^2 = .17$), in which polygraph examiners gave lower ratings of necessity compared to forensic psychologists and lawyers. This discrepancy indicated a different understanding of the status of polygraph reactions (Li et al., 2025; Synnott et al., 2020). Polygraph examiners tend to consider reactions to meaningful stimuli as convincing indicators. However, forensic psychologists and representatives of the legal professions emphasized the need for alternative explanations (Patel, 2025; Kazmiruk et al., 2021).

Comparing the results of stages A and B indicated a gap between the practice of documentation and the consensus minimum standard. The largest imbalance was recorded precisely with respect to those components that are methodologically most critical for the scientific validity of a polygraph examination. Indicators of alternative explanations of reactions and countermeasures verification became important. This result has important theoretical significance, as it indicated that the existing documentation practice is technically imperfect and systematically ignores those components that determine the scientific status of polygraph examination (Funicelli et al., 2023; Haberler et al., 2024). The identified imbalance regarding informed consent (a gap of 44.2 pp) deserves special attention. Informed consent received one of the highest consensus ratings (91.8% of ratings ≥ 7). However, the actual presence of this element in the documents is only 46.7%. Thus, more than half of the

analyzed reports did not contain confirmation of voluntariness and awareness of participation. This was also emphasized by other authors and indicated the need for methodological and legal mechanisms to ensure appropriate documentation (Alqatawna, 2024; Elaad, 2018).

The results obtained have several important limitations that must be considered when interpreting them. First, the sample of documents for analysis was formed according to a purposeful strategy, which limited the possibilities of generalizing the results to the entire set of polygraph reports. Second, the assessment of the quality of documentation was implemented based on formal criteria of the codebook without verifying the actual quality of the polygraph procedures performed. However, this work made several important contributions to the scientific field. From a theoretical point of view, it was the first to empirically substantiate the gap between regulatory expectations regarding the use of the polygraph in forensic psychological examination and the actual practice of documentation. The study also operationalized this gap through a system of measurable indices (MPI, PCI) and zones of normative and methodological imbalance. From a practical point of view, the former consensus minimum standard of 11 criteria can serve as the basis for the development of regulatory documents regulating the use of the polygraph in forensic psychological examination.

D. CONCLUSIONS

The implementation of the analysis of 60 polygraph reports (stage A) and a two-wave expert survey of 110 specialists (stage B) made it possible to verify that the polygraph reports conducted at the stage of the pre-trial investigation were characterized mainly by an average level of methodological transparency (MPI $M = 44.2$) and a marginal level of procedural compliance (PCI $M = 38.4$). Systemic documentation deficiencies were identified, consisting of six codebook domains; three of them were qualified as critical: lack of alternative explanations for reactions (80.0%), verification of countermeasures (78.3%), and test invalidity (78.3%).

The type of polygraph technique used was shown to be partially related to the level of documentation quality: CIT/GKT documents had significantly higher MPI (49.1 vs. 40.3; $r = .42$) and PCI (43.2 vs. 34.6; $r = .38$) scores than CQT documents. After two waves of the survey, the expert community reached an acceptable level of consensus (Kendall's $W = .74$) on 11 criteria for a minimum standard for the use of a computer polygraph in forensic psychological expert

research. The consensus minimum standard included the following indicators: indication of the procedural status of the person, informed consent, description of the test methodology, clear scoring criteria, voluntary participation, formulation of questions and incentives, quantitative assessment indicators, alternative explanations for reactions, verification of countermeasures, availability of graphs and raw registration records, and description of calibration. Of the 11 criteria of the standard, only two had an acceptable level of compliance with the report data: the procedural status of the person (gap of 10.7 pp) and the description of the test methodology (gap of 17.2 pp). The greatest imbalance was recorded in relation to the methodologically most significant components. This indicated the orientation of the existing documentation practice to the fulfillment of minimum formal requirements.

The scientific novelty of the study lies in the fact that for the first time an empirical comparison of the actual practice of documenting polygraph examinations with a consensus minimum standard formed based on an interprofessional expert survey was carried out; a system for measuring the quality of documentation through the MPI and PCI indices was developed and validated; areas of normative and methodological imbalance were identified and operationalized as a tool for prioritizing measures to improve practice.

Prospects for further research are related to several areas. First of all, it is advisable to conduct further validation of the established minimum standard in a comparative international space. In addition, it is important to study the relationship between the level of documentation quality and court decisions on the admissibility of polygraph test results as evidence.

REFERENCES:

- Afzali, M. U., Seren-Grace, A. P., Palmer, R. W., Neumann, E., Makarious, S., Wilson, D., & Jones, R. D. (2022). Detection of concealed knowledge via the ERP -based technique *Brain Fingerprinting: Real-life and real-crime incidents*. *Psychophysiology*. <https://doi.org/10.1111/psyp.14110>
- Alqatawna, A. (2024). Utilizing Artificial Intelligence (AI) in Criminal Justice and Policing. *Comparative Law Review*, 30, 9–33. <https://doi.org/10.12775/clr.2024.001>
- Baddeley, A., Brewin, C. R., Davies, G. M., Kopelman, M. D., & MacQueen, H. L. (2023). Legal aspects of memory: a report issued by the Psychology and Law Sections of the British Academy. *Journal of the British Academy*, 11, 95–97. <https://doi.org/10.5871/jba/011.095>

- Ben-Shakhar, G., & Iacono, W. (2021). Fallacies in the estimation of the validity of the Comparison Question Polygraph Test: A reply to Ginton (2020). *Journal of Investigative Psychology and Offender Profiling*, 18, 201–207. <https://doi.org/10.1002/jip.1576>
- Blikhar, V., Zaiats, O., Pavliuk, N., & Kalka, N. (2022). Psychological and Legal Aspects of Verification and Detection of Lies during Polygraph Examination. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 13(1), 276–291. <https://doi.org/10.18662/brain/13.1/284>
- Brennen, T., & Magnussen, S. (2023). Lie Detection: What Works? *Current Directions in Psychological Science*, 32(5), 395–401. <https://doi.org/10.1177/09637214231173095>
- Elaad, E. (2018). A Comparative Study of Polygraph Tests and Other Forensic Methods. In *Interviewing and Deception* (pp. 209–231). Routledge. <https://doi.org/10.4324/9780429428517-10>
- Elton, J. (2017). The Polygraph in the English Courts. *The Journal of Criminal Law*, 81(1), 66–79. <https://doi.org/10.1177/0022018316688722>
- Funicelli, M., Salphati, S., Ungureanu, S., & Laurence, J.-R. (2023). Examining levels of processing using verbal & pictorial stimuli with the complex trial protocol in a mock theft scenario. *Biological Psychology*, 108666. <https://doi.org/10.1016/j.biopsycho.2023.108666>
- Funicelli, M., White, L., Ungureanu, S., & Laurence, J.-R. (2021). An Independent Validation of the EEG-Based Complex Trial Protocol with Autobiographical Data and Corroboration of its Resistance to a Cognitively Charged Countermeasure. *Applied Psychophysiology and Biofeedback*, 46(3), 287–299. <https://doi.org/10.1007/s10484-021-09506-2>
- Gordon, N. J., & Fleisher, W. L. (2019). The Instrumental Detection of Deception Polygraph. In *Effective Interviewing and Interrogation Techniques* (pp. 311–325). Elsevier. <https://doi.org/10.1016/b978-0-08-102610-6.00024-7>
- Goswami, D. P., & Goswami, A. (2025). Virtual Justice: The Role of Technology in Transforming Criminal Administration. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5121477>
- Haberler, J., Suchotzki, K., Meinhardt, M., & Gamer, M. (2024). Applicability and validity of the reaction time-based concealed information test in a prison sample. *Legal and Criminological Psychology*. <https://doi.org/10.1111/lcrp.12277>

- Honts, C. R., Thurber, S., & Handler, M. (2021). A comprehensive meta-analysis of the comparison question polygraph test. *Applied Cognitive Psychology*, 35, 411–427. <https://doi.org/10.1002/acp.3779>
- Horsley, F. (2022). Clinical and forensic assessment of firesetting. In *Further Case Studies in Forensic Psychology* (pp. 136–155). Routledge. <https://doi.org/10.4324/9781003213116-9>
- Husieva, V., Avdieiev, O., Kornienko, V., & Kryvoruchko, L. (2022). Innovative technologies as alternative means of improving the efficiency of criminal justice in Ukraine. *Revista Amazonia Investiga*, 11(52), 260–268. <https://doi.org/10.34069/ai/2022.52.04.28>
- Iacono, W. G. (2024). Psychology and the Lie detector Industry: A Fifty-Year Perspective. *Biological Psychology*, 108808. <https://doi.org/10.1016/j.biopsycho.2024.108808>
- Kaplan, J., & Lundy, J. (2024). Contesting Confessions: Opinions of Expert Witnesses. *Journal of Forensic Psychology Research and Practice*, 1–26. <https://doi.org/10.1080/24732850.2024.2408388>
- Kazmiruk, S., Pampukha, I., & Blyzniuk, N. (2021). Problems of legal regulation of the use of polygraph in the security and defense sector of ukraine. *Visnyk Taras Shevchenko National University of Kyiv. Military-Special Sciences*, (2 (46)), 62–66. <https://doi.org/10.17721/1728-2217.2021.46.62-66>
- Knappe, J., Ullsperger, M., & Kirschner, H. (2025). Decoding deception with the P300: A meta-analysis of the Concealed Information Test. *International Journal of Psychophysiology*, 113236. <https://doi.org/10.1016/j.ijpsycho.2025.113236>
- Kotsoglou, K. N., & Biedermann, A. (2024). Polygraph-based deception detection and Machine Learning. Combining the Worst of Both Worlds? *Forensic Science International: Synergy*, 9, 100479. <https://doi.org/10.1016/j.fsisyn.2024.100479>
- Leo, R. A., Iacono, W. G., Denkinger, B., Harper, J., Brand, J., & Yoon, H. (2024). Brief of Amici Curiae, A Group of Concerned Scientists, Texas v. Cieliencki. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4952620>
- Li, J., Feng, G., Ling, C., Ren, X., Zhang, S., Liu, X., Wang, L., Vai, M. I., Lv, J., & Chen, R. (2025). A Novel Multi-Scale Entropy Approach for EEG-Based Lie Detection with Channel Selection. *Entropy*, 27(10), 1026. <https://doi.org/10.3390/e27101026>

- Lukács, G., & Ansorge, U. (2023). Response time concealed information test using fillers in cybercrime and concealed identity scenarios. *Memory*, 1–17. <https://doi.org/10.1080/09658211.2023.2195179>
- Lukács, G., Kleinberg, B., Kunzi, M., & Ansorge, U. (2020). Response Time Concealed Information Test on Smartphones. *Collabra: Psychology*, 6(1). <https://doi.org/10.1525/collabra.255>
- Malewski, H., & Kraujalis, L. (2024). Psychophysiological Research and its Place in the Lithuanian Scientific Discourse. *Theory and Practice of Forensic Science and Criminalistics*, 34(1), 58–75. <https://doi.org/10.32353/khrife.1.2024.04>
- Mergaerts, L., & Van Daele, D. (2023). Polygraph Testing in Criminal Proceedings: Safe and Sound or Swimming Against the Tide? *European Journal of Crime, Criminal Law and Criminal Justice*, 31(3-4), 292–324. <https://doi.org/10.1163/15718174-bja10048>
- Mykityshyna, K. (2024). Kilka uwag na temat dopuszczalności dowodu z ekspertyzy poligraficznej w procesie karnym Ukrainy. *Problemy Współczesnej Kryminalistyki*, 27, 211–246. <https://doi.org/10.52097/pwk.8465>
- Nagale, T., & Khandare, A. (2025). CIT-LieDetect: A Robust Deep Learning Framework for EEG-Based Deception Detection Using Concealed Information Test. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 8(1), 152–167. <https://doi.org/10.35882/jeeemi.v8i1.1300>
- Oswald, M. (2020). Technologies in the twilight zone: early lie detectors, machine learning and reformist legal realism. *International Review of Law, Computers & Technology*, 34(2), 214–231. <https://doi.org/10.1080/13600869.2020.1733758>
- Patel, N. (2025). Polygraphy as a tool for forensic psychological evaluation in cases of workplace harassment: Advances and ethical concerns. *Journal of Digital Security and Forensics*, 2(2), 50–54. <https://doi.org/10.29121/digisecforensics.v2.i2.2025.66>
- Quattrocchio, S. (2025). Criminal law enforcement through AI. In *Research Handbook on the Law of Artificial Intelligence* (pp. 349–372). Edward Elgar Publishing. <https://doi.org/10.4337/9781035316496.00025>
- Rothmann, R., & Mayer, E. (2025). AI-based behaviour analysis in prisons: A

- multidisciplinary assessment of an Austrian case study. *European Journal of Criminology*. <https://doi.org/10.1177/14773708251349702>
- Shalash, O., Métwalli, A., Sallam, M., & Khatab, E. (2025). A Multimodal Polygraph Framework with Optimized Machine Learning for Robust Deception Detection. *Inventions*, 10(6), 96. <https://doi.org/10.3390/inventions10060096>
- Suchotzki, K., Verschuere, B., & Gamer, M. (2021). How Vulnerable is the Reaction Time Concealed Information Test to Faking? *Journal of Applied Research in Memory and Cognition*, 10(2), 268–277. <https://doi.org/10.1016/j.jarmac.2020.10.003>
- Synnott, J., Dietzel, D., & Ioannou, M. (2020). Open Access: A review of the polygraph: history, methodology and current status. In *Reviewing Crime Psychology* (pp. 50–74). Routledge. <https://doi.org/10.4324/9780429346927-5>
- Ukhno, O. (2021). Genesis and Issues of Using Latest Technologies and Artificial Intelligence in Criminalistics, Forensic Expert Activity and Pre-Trial Investigation. *Theory and Practice of Forensic Science and Criminalistics*, 25(3), 40–59. <https://doi.org/10.32353/khrife.3.2021.04>
- Wolsink, L. N., Meijer, E. H., Smulders, F. T. Y., & Orthey, R. (2024). The concealed information test with a continuously moving stimulus. *Psychophysiology*, 62, 14714. <https://doi.org/10.1111/psyp.14714>
- Yefimenko, I., Slipchenko, V., & Vaško, A. (2023). Critical infrastructure as an object of criminal encroachment: General characteristics and features of the investigation organisation. *Naukovij visnik Natsional'noi akademii vnutrishnih sprav*, 28(2), 41–51. <https://doi.org/10.56215/naia-herald/2.2023.41>
- Yukhno, O., & Turenko, D. (2025). Problematic issues of using a polygraph in the detection, disclosure and pre-trial investigation of criminal offenses. *Archive of Criminology and Forensic Sciences*, 11(1), 83–99. <https://doi.org/10.32353/acfs.11.2025.05>