

1971

## The Reliability of Polygraph Examiner Diagnosis of Truth and Deception

Frank S. Horvath

John E. Reid

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

---

### Recommended Citation

Frank S. Horvath, John E. Reid, The Reliability of Polygraph Examiner Diagnosis of Truth and Deception, 62 J. Crim. L. Criminology & Police Sci. 276 (1971)

This Criminology is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

## THE RELIABILITY OF POLYGRAPH EXAMINER DIAGNOSIS OF TRUTH AND DECEPTION

FRANK S. HORVATH AND JOHN E. REID

Frank S. Horvath is a graduate of Michigan State University with a B.S. Degree in Police Administration. In 1964 following his graduation he pursued the Study of Scientific Polygraph testing at John E. Reid and Associates and became Chief Examiner in 1970. He is licensed as a polygraph examiner in the State of Illinois and is a Charter Member of the American Polygraph Association.

John E. Reid, LLB, De Paul University, Director of John E. Reid and Associates, has made a number of noteworthy contributions to the polygraph field. He is co-author with Professor Fred E. Inbau of Northwestern University Law School of *Truth & Deception, The Polygraph (Lie-Detector) Technique* and *Criminal Interrogation and Confessions* 2nd Edition. This is his fourth article to appear in the journal. His previous ones were "Simulated Blood Pressure Responses in Lie Detector Tests and a Method for Their Detection," "A Revised Questioning Technique in Lie-Detector Tests" and "Behavior Symptoms of Lie Detector Subjects."

This study was conducted to determine if Polygraph examiners, working independently of each other, are able to successfully diagnose deception solely from an analysis of Polygraph records. Previous studies dealing with this problem have indicated that Polygraph examiners can reliably determine truth or deception from the records alone, but none of them were conducted in real-life testing situations. Davidson (1968) for example, found that by motivating students involved in an experimental crime he could correctly identify all of the innocent and 92% of the guilty subjects with the use of the Polygraph.<sup>1</sup> Lykken (1959) in a prior experiment, also using students as subjects, reached substantially the same conclusion; he identified all of the innocent and 93.9% of the guilty subjects.<sup>2</sup> Neither of these studies, however, was conducted by or with practicing Polygraph examiners, nor did they rely upon an analysis of Polygraph records obtained in actual investigations. Consequently, the studies have little value in assessing the reliability of Polygraph examiner diagnosis in real-life situations.

Kubis (1962) carried out an elaborate research program for the Air Force Systems Command of the United States Air Force. Although he used a simulated test situation for the experiments, his examiners were trained personnel. He reported

that they were able to obtain significant accuracy in identifying the thief, the lookout, and the innocent suspect. He concluded that there was sufficient validity in these experiments to warrant confidence in the lie-detecting procedure as an aid to interrogation processes.<sup>3</sup>

Ordinarily, in actual Polygraph testing, the examiner uses a complete diagnostic technique to determine deception. He takes into account detailed background information regarding the subject and the investigation; he has the benefit of actually conversing with the subject and observing the subject's attitude and behavior symptoms. In addition, he prepares and reviews the general comprehension of the questions. Since all of these auxiliary sources of information may be factors in arriving at a truth-deception diagnosis, the present study eliminated them and concentrated on Polygraph record analysis only.

In this study ten Polygraph examiners on the staff of John E. Reid and Associates agreed to analyze a number of Polygraph records independently and without the benefit of any information beyond the Polygraph records themselves. Seven of the examiners had been engaged in Polygraph testing more than one year; the remaining three were relatively inexperienced; they had been engaged in Polygraph testing from four to six months and were still participating in an internship training program.

<sup>1</sup>Davidson, P. O. *Validity of the Guilty Knowledge Technique: The Effects of Motivation*. 52 J. APPL. PSYCHOL. 62-65, (1968).

<sup>2</sup>Lykken, D. T. *The GSR in the Detection of Guilt*, 43 J. APPL. PSYCHOL. 385-388 (1959).

<sup>3</sup>Kubis, J. F. *Study in Lie Detection, Computer Feasibility Consideration*. Griffin Air Force Base, New York: Rome Air Development Center, Air Force Systems Command, United States Air Force, 1962.

The Polygraph records submitted to the examiners for analysis were obtained from twenty-five case investigations originally conducted by one of the authors. (Horvath) The cases were typical of the types usually presented to private Polygraph examiners: theft, sexual misconduct, sabotage, bribery and criminal damage to property. Subsequent to the Polygraph examination each of the selected cases had been solved by a fully corroborated confession of the guilty subject. In these twenty-five cases, seventy-five subjects had been tested originally, but the Polygraph records of only forty of them were selected for the use in this study for the following reason: the polygraph records which were dramatically indicative of truth or deception were eliminated from those submitted to the examiners because they did not require any exceptional skill to interpret. In other words, the evidence of truth and deception would be very obvious to any trained Polygraph examiner.

Twenty of the forty sets of Polygraph records chosen by the writer for this study were verified as those obtained from guilty subjects, and twenty test records were obtained from verified innocent subjects. The records contained one hundred and sixty-four (164) relevant questions which were submitted to the examiners; eighty-one (81) of these questions were verified as having been answered untruthfully during the examinations; eighty-three (83) of the questions were proven to be answered truthfully.

The recording instrument used in conducting the original Polygraph examinations was a five-channel Reid Polygraph which recorded thoracic respiration, abdominal respiration, blood pressure-pulse rate, muscular movements and pressures, and galvanic skin response (GSR). No attempt was made to determine which recording channel or channels the examiners relied upon in arriving at their decisions of truth and deception.

The subjects in each case had been given Polygraph examinations according to standard Reid Control Question Technique.<sup>4</sup> Essentially this technique consists of a pre-test interview and Polygraph testing. During the interview the examiner explains to the subject the purpose of the test and the nature of the instrument. It is at this time that the examiner seeks to condition

the subject for the test and to formulate and review with him the actual test questions. In the pre-test interview the examiner objectively notes the subject's behavior symptoms such as how he acts, looks, and talks and attempts to make an evaluation of these observations in terms of truth or deception. No attempt is made at this time to interrogate the subject with a view to obtaining a confession. At the conclusion of the interview, which lasts about twenty minutes, the examiner proceeds with the Polygraph testing.

The Polygraph testing consists of the asking of relevant, irrelevant and control questions during a number of separate tests. The questions in the 3, 5, 8, 9 and 10 positions are relevant and relate to the matter under investigation, such as, in a murder case, "Did you kill John Jones?" and "Did you shoot John Jones with a .38 caliber revolver?" The questions in the 1, 2, 4 and 7 positions are irrelevant to the issue being investigated; they deal with such matters as, "Do they call you Joe?", "Are you over 21 years of age?", etc. These irrelevant questions are asked for the purpose of establishing the subject's normal pattern of responsiveness. The remaining two questions are control questions. They are placed in the 6 and 11 positions. A control question is one which is unrelated to the matter under investigation, but is of a similar, though less serious nature and one to which the subject will, in all probability, lie; or at least his answer will give him some concern with respect to either its truthfulness or its accuracy. For instance, in a burglary investigation the control question might be, "Did you ever steal anything?" or "Except for what you have already told me, did you ever steal anything else?" The response or lack of response to the control question (in respiration, blood pressure-pulse rate, or GSR) is then compared with what appears in the tracings when the subject is asked the questions relevant to the issue under investigation. If the subject responds to a greater degree and with more consistency during the test series to the control questions than to the relevant questions, he is considered to be telling the truth regarding those relevant questions. On the other hand, if the subject responds more to the relevant questions than to the control questions, it is suggestive of lying regarding the relevant questions.<sup>5</sup>

<sup>4</sup> Reid, John E. *A Revised Questioning Technique in Lie-Detector Tests*, 37 J. CRIM. L. C & P. S. 542 (1947). TRUTH AND DECEPTION: THE POLYGRAPH (LIE-DETECTOR) TECHNIQUE, 27-32 (1966).

<sup>5</sup> The two previous paragraphs are excerpts from *The Lie Detector Technique: A Reliable and Valuable Investigative Aid*, Inbau, F. E. and Reid, J. E., 50 A.B.A.V. (5) (1964).

In about 25 percent of Polygraph cases truth or deception may be so clearly disclosed by the nature of the responses to relevant or control questions that the examiner will be able to point them out to any non-expert and satisfy him of their significance. All records of this category were eliminated from use in this study because they do not constitute a serious test of an examiner's expertise in chart interpretation. In roughly 10 percent of the Polygraph cases the records will be uninterpretable by even the most skilled examiner. In about 65 percent of the cases, however, the responses or lack of responses, to the control questions and relevant questions are sufficiently subtle in appearance and significance so that only a highly skilled and well trained examiner will be able to interpret them for truth and deception. All of the Polygraph records given to the examiners in this study could be classified as belonging to this category.

The examiners were unfamiliar with either the cases or the Polygraph records which they were called upon to analyze. They were not allowed to discuss the project amongst themselves until all had completed it. They were not given any of the actual test questions used in the original investigations, but because of their familiarity with the technique, each examiner knew the placement of the irrelevant, relevant, and control questions by their respective numbers as recorded on the records.

The examiners were told on an individual basis that they would be allowed one full working day to analyze the forty sets of Polygraph records. They were instructed to detect the guilty subject, if any, in each investigation and also to "clear" each innocent subject. In addition to this, they were instructed to diagnose truth or deception on each relevant question asked of all forty subjects. They were admonished not to report any subject as totally inconclusive, but if they found in analyzing any particular question reaction that they could not decide truth or deception, they were allowed to report that particular question as doubtful or inconclusive. The reason for this concession was that in any given Polygraph examination some of the relevant questions may carry more "emotional weight" than others, even though they all relate to the same investigation. This is especially true in the instance where a guilty person is tested. Often he will respond to a greater degree to a question regarding whether or not he himself committed the offense than he will to a question about whether or not he knows who committed the offense, even

though he is lying to both questions asked. The more direct and more emotionally weighted question such as, "Did you shoot John Jones?" sometimes may "mask out" or otherwise "dampen" the response on the indirect or less emotionally weighted questions, such as, "Do you know who did shoot John Jones?"

Prior to being given the Polygraph records, the examiners were told that all subjects were verified as guilty or innocent, but they were not told the number of subjects in each category. More significantly, they were not told whether the Polygraph records of the actual perpetrator were included in each of the cases submitted to them for diagnosis. The examiners were given only basic factual information from each of the twenty-five cases, together with the selected Polygraph records.

The following information, chosen from one of the cases used in this study, is illustrative of the amount and the type of information presented to the project examiners:

An electric motor was sabotaged at a large midwestern rubber company. It was suspected that one of the company's employees had inserted some knife blades (which were used at the company) into the armature of the motor when it was not running. When it was turned on, the blades caused the motor to "blow up" and produced extensive damage to the surrounding area and almost seriously injured several employees.

The examiners were not told that fourteen employees were given Polygraph examinations before the guilty person was detected in the original investigation. They were supplied with only the brief factual information given above and with the Polygraph records of six of the original fourteen subjects. The six sets of records they were given were those selected from the fourteen as best fitting into the category which requires special skill to interpret. The remaining eight sets of Polygraph records were not given to the examiners. The Polygraph records of the actual perpetrator of this sabotage were not included in the six sets of records given to the examiners for analysis; this fact, however, was withheld from the examiners.

## RESULTS

*Overall Innocent-Guilty Case Judgments.* The ten examiners achieved an average 87.75 percent accuracy in solving the cases, i.e., in correctly detecting the guilty subjects and correctly identifying the innocent subjects. As can be seen from Table 1, however, there was a significant difference

TABLE 1

DISTRIBUTION OF INNOCENT-GUILTY JUDGMENTS FROM EVALUATING POLYGRAPH RECORDS BY EXAMINERS

	Actually Innocent (20)		Actually Guilty (20)		Percent Correct Judgments
	"Innocent"	"Guilty"	"Innocent"	"Guilty"	
<b>Experienced examiners</b>					
1	19	1	0	20	97.5%
2	18	2	0	20	95.0%
3	19	1	2	18	92.5%
4	19	1	2	18	92.5%
5	18	2	2	18	90.0%
6	20	0	5	15	87.5%
7	18	2	4	16	85.0%
Sub-total	131	9	15	125	91.4%
<b>Inexperienced examiners*</b>					
8	19	1	3	17	90.0%
9	16	4	8	12	70.0%
10	15	5	4	16	77.5%
Sub-total	50	10	15	45	79.19%
<b>Total</b>	<b>181</b>	<b>19</b>	<b>30</b>	<b>170</b>	<b>87.75%</b>

\* Less than six months experience.

between the experienced and the inexperienced examiners. The experienced examiners were successful in 91.4 percent of their diagnoses; the inexperienced in only 79.1 percent.

It should also be noted that the more experienced examiners were quite consistent with each other. Their accuracy scores ranged from a low of 85 percent to a high of 97.5 percent, with five of the seven in this group achieving a 90 percent accuracy or higher. Only one of the three inexperienced examiners achieved the 90 percent accuracy level. The remaining two achieved only a 70 percent and a 77.5 percent score, respectively.

The results also seem to support the belief of most Polygraph examiners that their errors generally favor the guilty subject, i.e., that an examiner is more inclined to report a guilty subject innocent than he is to report an innocent subject guilty.

There was a total of 400 innocent-guilty judgments to be made by the examiners; that is, each of the ten examiners was called upon to judge each of the forty subjects either guilty or innocent. One-half of the judgments were to be made on verified innocent subjects and one-half were to be made on verified guilty subjects; therefore, there were 200 judgments in each category.

Over the 200 judgments of the twenty verified innocent subjects, nineteen (9.5 percent) were erroneously judged "guilty" by the examiners; of the 200 judgments of the twenty verified guilty subjects, thirty (15 percent) were erroneously judged innocent. In analyzing this further, it should be noted that for the seven experienced examiners only nine out of 140 (6.4 percent) judgments on the twenty innocent subjects were errors, while among inexperienced examiners, 16.6 percent of their judgments on verified innocent subjects were errors. For verified guilty subjects, 10.8 percent of the experienced examiner judgments were "innocent" errors, while 25.0 percent of the inexperienced examiner judgments were "innocent" errors.

#### INDIVIDUAL RELEVANT QUESTION BY QUESTION ANALYSIS

Table 2 summarizes the data for each examiner's performance in correctly interpreting the 164 relevant questions for truth and deception.

Nine of the ten examiners achieved at least a 77.5 percent accuracy rating on the question by question analysis and six of the ten achieved better than 83 percent. Although the experienced examiners again scored significantly higher than the

TABLE 2  
DISTRIBUTION OF TRUE-LIE JUDGMENTS OF EXAMINEE RESPONSES TO 164 QUESTIONS BY EXAMINERS

Examiner Judgment	Actually True Response (83 true responses)			Actually Lie Response (81 lie responses)			Percent Correct Judgments
	"True"	"?"	"Lie"	"True"	"?"	"Lie"	
Experienced examiners							
1	79	0	4	0	1	80	96.6%
2	74	0	9	3	0	78	92.7%
3	77	0	6	14	0	67	87.6%
4	75	3	5	8	7	66	86.0%
5	64	18	1	7	11	63	77.5%
6	60	0	23	0	0	81	86.0%
7	65	11	7	14	5	62	77.5%
Sub-total	494	32	55	46	24	497	86.2%
Inexperienced examiners*							
8	71	4	8	12	13	56	77.5%
9	60	15	8	24	11	46	64.6%
10	61	13	9	19	2	60	83.8%
Sub-total	192	32	25	55	26	162	75.0%
Total	686	64	80	101	50	659	79.3%

\* Examiners with less than six months experience.

TABLE 3

EXAMINER JUDGMENTS OF THE RESPONSES OF ONE INNOCENT AND TRUTHFUL SUBJECT TO FOUR RELEVANT QUESTIONS

	Relevant Question Number				Overall Judgment
	#3	#5	#8	#9	
Experienced examiners					
1	Truthful	Truthful	Truthful	Truthful	Innocent
2	Truthful	Truthful	Truthful	Truthful	Innocent
3	Truthful	Truthful	Truthful	Truthful	Innocent
4	Truthful	Truthful	Truthful	Truthful	Innocent
5	Truthful	Truthful	Truthful	Inconclusive	Innocent, but guilty knowledge
6	Truthful	Truthful	Truthful	Truthful	Innocent
7	Not truthful	Not truthful	Not truthful	Not truthful	Guilty
Inexperienced examiners*					
8	Not truthful	Not truthful	Not truthful	Not truthful	Guilty
9	Truthful	Truthful	Truthful	Truthful	Innocent
10	Not truthful	Not truthful	Not truthful	Not truthful	Guilty

\* Examiners with less than six months experience.

inexperienced, both groups combined had only an overall error of 20.7 percent. This figure, however, is somewhat misleading, because it includes as errors those relevant questions which the exami-

ners reported as inconclusive or on which they were unable to make any diagnosis. This error was usually made by examiners when they analyzed the Polygraph records of a guilty subject and cor-

rectly interpreted the more direct relevant questions, but were unable to interpret an indirect relevant question due to the "masking out" effect described above. If these inconclusive question errors are eliminated, examiners actually made only an 11 percent error; that is, they judged only 11 percent of the relevant questions opposite their verification, thus achieving an overall 89 percent accuracy rating.

To further illustrate the results of the question analysis, Table 3 indicates how accurately each examiner interpreted the Polygraph records of one of the six subjects in the previously described sabotage case.

The relevant questions asked of all subjects in this case were as follows: Question #3, "Did you insert two mill knife blades into the armature of of that motor?"; Question #5, "Did you put those blades into that motor?"; Question #8, "Did you cause that damage to the mill motor?"; Question #9, "Do you know who put those knife blades in the mill motor?" There was no question asked in #10 position. The irrelevant and control questions were placed according to the format previously explained.

The subject (used as an example in the table) was asked the four relevant questions. Since it had been verified that his answers were truthful to all questions, his records should have been analyzed by the examiners as being those of an "innocent" subject and as consisting of four truthful relevant question responses. Only examiners 1, 2, 3, 4, 6 and 9 judged the subject in this manner. Examiners 7, 8 and 10 judged this subject as "guilty" and the four relevant question reactions as "lies". Examiner 5 judged this subject as innocent by finding him telling the truth to Questions #3, #5 and #8, but recorded him as inconclusive on Question #9. (knowledge question) This was recorded as an error.

#### DISCUSSION

These data clearly support the claim of Polygraph examiners that they can reliably diagnose truth and deception or detect the guilty and identify the innocent solely from an analysis of Polygraph records. In actual practice, of course, a Polygraph examiner has the benefit of all the detailed factual information in the case beforehand, as well as the behavior symptoms of the subject at the time of the test and moreover in many case

situations he has the full complement of Polygraph records of all the subjects in the case before he issues an opinion as to whether the subject is truthful or not. In actual testing situations the examiner places the utmost reliance upon responses or lack of responses on Polygraph records, but he is afforded the additional opportunity to evaluate the attitude of the subject and to make allowances for a resentful or angry attitude, a condition which could cause an error in interpretation of Polygraph records. An opportunity to observe the subject and evaluate his attitude toward the test would allow an examiner to diagnose truth and deception more reliably than the examiners in this study.

If the examiners had been given all of the Polygraph records in each case and were aware of the fact that one of the subjects must be guilty, the accuracy ratings for both experienced and inexperienced examiners would have been greatly improved. This would have allowed the examiners to compare the Polygraph records of one subject with those of another subject in the same investigation.

Although the results of the present study attest to the reliability of Polygraph examiner's ability to diagnose truth and deception, they also attest to the value of practical experience in qualifying examiners as experts. The accuracy of the experienced examiners was significantly better than that of the inexperienced examiners. This was probably due to the fact that the experienced examiners had more practical knowledge of the limitations of the Polygraph technique in that both groups of examiners had been taught the "theory" of the technique in the same manner. The examiners with the most experience were more able to apply consistently the fine points of the theory, which assisted them in diagnosing truth and deception with greater accuracy.

#### ACKNOWLEDGEMENTS

The writers are indebted to Fred E. Inbau, Professor of Law, Northwestern University for his assistance and suggestions in the preparation of this paper and to Philip Ash, Ph.D. Professor of Psychology, University of Illinois, Chicago Circle Campus for his critical review and expert arrangement of the statistical data. The writers also gratefully acknowledge the conscientious and unstinting efforts of the staff of John E. Reid and Associates without whom this paper could not have been written.