A BUILT-IN VALIDITY IN POLYGRAPH FIELD EXAMINATIONS [1]

Avital Ginton
Behavioral Sciences Section, Criminal Identification Division
Israel Police Headquarters &
Criminology Department, Bar-Ilan University, Israel

An estimate of the validity of polygraph field examinations was developed using a new strategy. Forty-two pairs of real life cases (mostly unverified), each consisting of two examinees who accused each other of being deceptive about certain clear cut points, were used. The eighty four polygraph examinations were scored by experienced examiners. The examiners were completely blind as to the cases involved, nor did they identify the pairs. Assuming that in each pair there must be one deceptive and one truthful subject, It was clear that each pair which had been scored either as two deceptive or two truthful subjects must include an error. Also, It was clear that a pair which had been scored as one deceptive and one truthful subject, means either two hits or two errors. Based on these assumptions the error rate was computed. Results indicated an error rate of 19.5%. The problem of generalizing these results to polygraph field examinations as a whole is discussed.

Validity studies of polygraph examinations are basically divided into two categories. Either they are based on real life situations and examinations (field studies), or on some sort of laboratory experimental setting (analog studies).

The two categories suffer from different methodological weaknesses. The problems in field studies stem mainly from the difficulties in obtaining a reliable criterion against which the polygraph results can be validated, as well as in

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

avoiding a substantial sampling bias (Oren, 1975, Ginton et al. 1982). While these weaknesses are usually controlled in analog studies, it is the lack of realistic fear of failure on the test which questions the usefulness of this type of studies for assessing the validity of polygraph examinations in real criminal investigations (Orne,1975, Ginton et al. 1982).

To overcome these methodological handicaps is almost an impossible mission within the conventional approach (but, see Ginton et al. 1982), especially if one considers the limitations presented by ethical standards.

The present study tries to reduce the above mentioned pitfalls by adopting a completely new approach. The study is based on field examinations conducted in cases where two opposing versions regarding the occurrence of specific events were presented by two examinees in each case. Selecting only the cases in which it was practically impossible to assume that the two parties were telling their subjective truth or that the two of them were lying, it was possible to estimate the validity of the tests by mathematical computations based upon the proportions of the various results obtained by the polygraph examinations per-se. Thus there was no need to rely upon confessions, convictions or any other sort of so-called verifications, which usually result in a substantial sampling bias.

METHOD

Eighty-four records of control question polygraph examinations that had been conducted in the two main polygraph laboratories of the Israel National Police, during the years 1982-1984 were selected for this study. The 84 records refer to 42 cases with a pair of examinations per case. Each pair consisted of one examinee that put some sort of a blame on the other, who totally denied it, accusing the first examinee in fabricating the accusation. These 42 cases were the only cases that survived a selection procedure in which a careful content analysis of all the details involved in the controversies resulted in a conclusion that it is almost impossible to expect the two parties to be either both deceptive or both telling the truth regarding the relevant questions.

There were 11 cases of minor sexual offence, two cases of theft, six cases of fraud, seven cases of police brutality and 16 cases of miscellaneous issues.

The 84 records were numerically scored by six experienced examiners using a seven point scale per comparison, ranging from +3 to –3. Each record was scored by two examiners

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

and using the Latin square design, each examiner scored a total of 28 records. The scorers were completely blind as to the cases involved nor could they identify the pairs.

For detailed explanation of the numerical scoring technique the reader is referred to Barland and Raskin, 1975. Nevertheless, it is important to mention that in this technique each record receives a final numerical score. A negative score signifies that the autonomic responses to the relevant questions were on the whole stronger than to the control questions, thus indicating deception to the relevant issue. On the other hand, positive final scores indicate truthfulness. In case of 100% accuracy, the outcomes of each pair should be one positive and one negative score (opposite outcomes). However since we might expect errors, some pairs might produce, either two positive or two negative scores (identical outcomes).

MATHEMATICAL ANALYSIS

Let us define the probability of having a correct decision as P(correct)=X, and the probability of having an incorrect decision as P(incorrect)=Y. Since a decision must be made whether correct or wrong, it is clear that P(correct) +P(incorrect)=X+Y=1.

Assuming that the examinations of the two parties in each pair were conducted independently, the probability of having correct decisions for both parties equal to X^2. Following the same logic it is clear that the probability of having incorrect decisions for both parties is equal to Y^2, while 2XY expresses the probability of having one correct and one incorrect decision (see table 1).

Table 1

A model of the distribution of outcomes in paired examinations of opposing versions. X=P(correct decision**) ; Y=P(**incorrect decision**).**

		Examinees presenting one side of each case		
		Probability of Correct Decision	Probability of Incorrect Decision	Total
Examinees Presenting The counter Side of each Case	Probability of Correct Decision	X^2	XY	X
	Probability of Incorrect Decision	XY	Y^2	Y
	Total	X	Y	1.0

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

The percentage of pairs which had two identical outcomes (i.e., two positive or two negative scores per pair) is equal to the probability of having one correct and one incorrect decision (2XY). Similarly, the percentage of pairs which had two opposite outcomes (i.e., one positive and one negative score per pair) may be regarded as the joint probability of having either two correct or two incorrect decisions per pair ($x^2 + y^2$).

Hence given the percentage of identical and opposite outcomes it is possible to compute the values of X and Y which indicate the accuracy rate of detection of deception by polygraph examinations.

RESULTS

As mentioned above, each record was scored blindly by two examiners. The correlation between the two scores over the 84 records which indicates the reliability was found to be 0.85. Of the 84 records there were seven records from seven different pairs for which the two scorers gave opposite final scores. These records were considered as inconclusive, and for the purpose of the present study they were eliminated from further computations. Thus, only 35 pairs (70 records) were left.

In 24 pairs the two parties received opposite scores (i.e., one positive and one negative) and in 11 pairs the scores were identical (three pairs with positive scores for both sides, and eight pairs with negative scores). The above data lead to the following equations:

$$X^2 + Y^2 = {}^{24}/_{35} = 0.686$$

$$2XY = {}^{11}/_{35} = 0.314$$

A simple computation brought about the following solution :

$$X = 0.805 \text{ and } Y = 0.195.$$

Inspection of the various types of offences in this study revealed that cases of police brutality contributed a disproportionate number of errors. When this category was eliminated, 35 pairs remained, of which six were regarded as inconclusive, leaving 29 pairs (58 polygraph records) for computation. The results for these 29 pairs were:

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

$$X^2 + Y^2 = {}^{22}/_{29} = 0.759$$

$$2XY = {}^{7}/_{29} = 0.241$$

$$X = 0.86; \quad Y = 0.14$$

Possible explanations for the higher proportion of errors in the cases of police brutality, are beyond the scope of the present article, and will be presented elsewhere.

DISCUSSION

The internal validity of the accuracy rate found in this study (80.5%) depends heavily on the assumption that the two examinations in each pair were conducted independently. Thus, one can raise the question that knowledge of the result of the first examination in each pair might influence the manner in which the second examination is conducted, leading to an artificial increase in the probability of obtaining an opposite outcome. This is especially acute when the two parties are tested by the same examiner. While it is impossible to ignore this danger, there are some indications that this is not the case. In the present study 20 pairs of examinations were originally conducted by two different examiners per pair, and only 22 pairs were conducted by one examiner per pair. Table 2 gives the distribution of the blind evaluation outcomes in these two categories. It seems that there is not a substantial difference between the pairs of examinations conducted by a single examiner and those which were conducted by two different examiners. It should be mentioned that in most cases in which two examiners were involved, the examinations were actually conducted parallel in time, eliminating any possible influence of one outcome on the other.

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

Table 2
Distribution of outcomes in paired examinations conducted by a single examiner per pair or two different examiners.

	ONE EXAMINER	TWO EXAMINERS	TOTAL
OPPOSITE OUTCOMES	13	11	24
IDENTICA L OUTCOMES	5	6	11
INCONCLUSIVE OUTCOMES	4	3	7
TOTAL	2 2	*20*	42

Another important question which stems from the present study is the generalizebility of the results to polygraph field examinations as a whole. I believe that the very fact that an innocent person is confronted with a direct accusation by another person increases his level of tension and anxiety more than in the usual situation in which a person is confronted with a circumstantial suspicion. Thus the present situations are prone to a higher rate of false positive errors than usual. On the other hand if a guilty person is confronted with such a direct accusation there is a high probability of confession before a polygraph examination takes place. Thus the guilty persons who took these examinations may be "better liars" on the average than the normal population, and therefore a higher rate of false negatives is also expected. In conclusion, I believe that the present result is an underestimation of the general accuracy of polygraph examinations as a whole.

Finally, it should be noted that among the 11 pairs with identical outcomes there were three false negatives and eight false positives, which can be used as a gross estimation of the distribution of error types.[2]

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)

REFERENCES

Barland, G.H. & Raskin, D.C. An evaluation of field techniques in detection of deception. Psychophysiology 1975, 12, 321-330.

Ginton, A., Daie, N., Elaad, E., & Ben Shakhar G. A method of evaluating the use of the polygraph in a real life situation. Journal of Applied Psychology, 1983, 67, 131-137.

Orne, M. T. Implications of laboratory research for the detection of deception. In: Ansley, N., (Ed.) Legal Admissibility of the Polygraph, 1975, Charles C. Thomas, Springfield, Illinois, U.S.A. pp. 94-119.

NOTES: [1] This article is only a condensed presentation of the study. A more detailed version will be published elsewhere in the near future.

[2] A more precise computation of the two different error rates will appear in the detailed version of this study.

Anti-Terrorism; Forensic Science; Psychology in Police Investigations. Boulder, Colo. :Westview Press; Jerusalem: Heiliger and co., **1986**, Pp. 148-154 (**Proceedings of Identa 85**. The international congress on techniques for criminal identification & counter terrorism. Jerusalem, Israel.)