EUROPEAN
POLYGRAPH

Volume 3 • 2009 • Number 3–4 (9–10)

Tuvia Shurany*
[uczelnia]
James Allan Matte**
[uczelnia]
Einat Stein***
[uczelnia]

Influence of case facts on blind scorers of polygraph tests

The first author (Shurany) was asked by a client if the Quality Control Reviewer should be provided with the facts of the case before blind scoring the physiological data recorded on the polygraph charts related to that polygraph examination, and whether it would have an influence on the reviewer's judgment and evaluation. The question generated this research study, which was preceded by a field study by Dror & Rosenthal (2008) that employed meta-analytic procedures to determine the degree of reliability and bias ability of forensic experts. The sample sizes in Dror et al's study involving fingerprint identification were

* e-mail
** e-mail
*** e-mail

quite small, some with only five and six experts. However, its findings revealed a biasing effect that increased the proportion of incorrect judgments from 30% to 70%. Because fingerprint analysis is so well established, it is expected that the results of Dror et al's study would apply equally to other forensic

Method

Participants and procedure

Three groups of polygraph examiners were selected for this study comprising a total of 82 polygraphists; from Mexico (n=35), United States (n=34) and Bulgaria (n=13). All Polygraphists were skilled and experienced examiners. They were instructed to evaluate and score the polygraph charts presented to them via PowerPoint slides in the manner in which they were trained, using the 3-position scale.

Commencing in the early 1970's, the Quality Control Review Section of the Department of the Army Criminal Investigation Division Command (CID) which reviewed all polygraph examinations conducted by the CID world-wide, used a 3-position scale to score the polygraph charts and whenever the reviewer's score disagreed with the score of the original polygraphist, the 7-position scoring scale was used to arrive at a final determination. (Brisentine, 1974, 1995). Several research studies have been conducted and reported regarding the effectiveness of the 3-position versus the 7-position scoring scales in polygraph chart analysis. Blackwell's (1999) field study found "the PDD examiners mean level of accuracy was 75.7% and 66.3% for the 7- and 3-position scoring scales, respectively." Blackwell stated that "without exception, the overall level of accuracy generated by the examiners when using the 7-position scoring scale was higher than when using the 3-position scoring scale. The same was true when looking at the overall percentages for either the innocent examinations or the guilty examinations." Krapohl (1998) found that the 3-position scale with a cutoff (fixed threshold) of +/–4 was statistically equivalent to the widely accepted 7-position scale with the +/–6 cutoff score (fixed threshold). However, Krapohl also found that "the highly experienced raters in this study rarely used the full range of available values in the 7-position scale, employing the narrower range of the 3-position scale for about 90% of the question comparisons." Capps and Ansley (1992) and Van Herk (1991), like Krapohl, found that the accuracy of the 7- and 3-position scales depended on the threshold used. The Backster and Matte Comparison Techniques (Matte, Backster 2007) use

an *increasing* threshold, whereas other Zone Comparison Technique modifications (DoDPI, Utah) employ a fixed threshold. (Matte, Backster 2000; Matte 2002).

The charts were selected from cases that used the Backster Zone Comparison Technique. All the Polygraphists were presented with 8 separate examinations each containing 4 charts. The first and last polygraph examination consisted of the same confirmed No Deception Indicated (NDI) charts. The second through the seventh examination consisted of unconfirmed NDI and DI (Deception Indicated) results which were used as a buffer to prevent recognition of examination #8 as being the same as examination #1.

Before the presentation of examination #1, no case facts were given to the polygraphists. Conversely, immediately prior to the presentation of examination #2 through #7, they were given fictitious facts of the case. Thus, they knew these facts prior to their evaluation and scoring of the charts. At examination #8, the polygraphists were presented with fictitious case facts and information supposedly provided by the police that fingerprints of the examinee found at the scene of the crime matched with 90% probability.

The first group of 35 polygraphists from Mexico evaluated examinations #1 through #8 on the same day. One of them asked the first author (Shurany) whether they had seen the charts of examination #8 before. He was instructed to simply evaluate the charts without providing him an answer to his question, and all of the polygraphists were able to hear the question and answer.

The above mentioned procedure was applied with another group of 34 polygraphists from the USA and 13 polygraphists from Bulgaria, with the exception that examinations #1 through #4 were presented on one day and examinations #5 through #8 were presented 48 hours later. This change in procedure was necessitated by the experience of the first group where one of the polygraphists recognized examination #8 as being identical to examination #1.

The score sheets from examinations #1 and #8 from all participants were entered into an Excel sheet. The names of the polygraphists were omitted from the score sheets which were identified only by a random number. The score sheets of examination #1 were compared with the score sheets of examination #8.

Results

In order to examine whether there are differences between the Charts' scores with prior knowledge vs. without prior knowledge, and in order to explore whether these differences are similar among the Polygraphists from the different countries, a 2 (knowledge) x 3 (country) MANOVA with repeated measures on knowledge (without vs. with) was performed.

The Manova revealed significant differences between the Polygraphists' scores to the same chart with vs. without prior knowledge ($F_{(2, 78)} = 17.9$, $p<.001$, $Eta^2=.32$).

Univariate ANOVAs revealed that the scores following previous knowledge were significantly lower as compared to the scores of the same chart without previous knowledge(see Table 1), for each compared question (Q5, Q7).

For Q5 the scores decreased by 37% on average, with a maximum decrease of 8 points, and for Q7 the average decrease was 26%, with a decrease of as much as 10 points in one case.

Table 1. Means and Standard Deviations of Charts' scores with vs. without prior knowledge

	Prior Knowledge					
	Without		With			
Measures	M	SD	M	SD	$F_{(1,79)}$	Eta^2
Q5	3.44	2.73	2.17	2.60	27.24***	.26
Q7	5.04	2.56	3.70	3.15	19.90***	.20

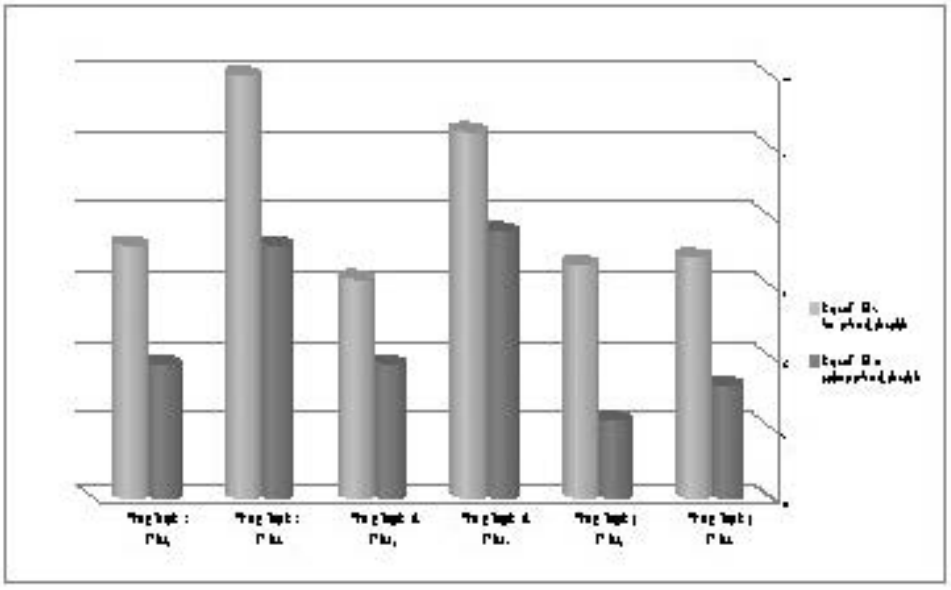
****p<.001*

No interaction of knowledge X country was detected ($F_{(4,156)} = 0.62$, $p>.05$), i.e. the decrease in scores with vs. without prior knowledge was found to be similar among the polygraphist from the different countries (see Figurer 1)

Furthermore the significant decrease in chart's score following prior knowledge produced 9 (11%) false positives as compared to one false positive (1.2%) with no prior knowledge. Generally, the rate of inaccurate decisions (i.e. false Positives or inconclusive, to a confirmed no deception indicated charts) was doubled as a result of prior knowledge (65.84% as compared to 32.9% respec-

tively). McNemar test for comparing frequencies in repeated measures with dichotomic variables (accurate/inaccurate decision) revealed that this difference in inaccuracy rate is significant ($c^2 = 25.04$, $p<.001$).

Figure 1. Average Score Change with vs. without prior knowledge



Conclusion

The results of this study indicate that knowledge of case facts does have an influence on the polygraphist's evaluation and scoring of the physiological data recorded on polygraph charts. Therefore a Quality Control Review should be conducted in three stages:

The authors wish to express their gratitude to Dr. Frank Horvath for his contribution of research information and statistical data analysis and his invaluable advice on the conduct of this study. The authors also wish to dedicate this study to the memory of Dr. Ronald M. Reese and William J. Schove, Jr. disciplines including forensic psychophysiology. However, another study by Elaine Elaidi, et al (1998), that used blind scoring of polygraph records, reflected a different result. Prior expectations affected the polygraphists' judgments when the polygraph charts reflected inconclusive results which offered no indications of guilt or innocence.
However when the objective physiological evidence reflected strong indications of guilt or innocence which clearly contradicted the polygraphist's expectations, judgments were not affected by these expectations. Hence these results indicate that the effect of prior expectations on the interpretation and scoring of polygraph charts is limited to inconclusive polygraph records.

1. Blind scoring should be conducted without any knowledge of the case facts, examiner's impression, and results of the polygraph examination.
2. After blind scoring has been accomplished, the polygraphist should then review the case facts for adequacy of test question formulation.
3. Review the audio/video recording for assurance that no procedural violations were committed by the original polygraphist.

References

Blackwell N.J. (1999), *PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examinations from actual criminal investigations*, Polygraph, 28 (2), 149–175.

Brisentine R.A. (1995, Nov. 13), Telephone conversation with J.A. Matte.

Brisentine R.A. (1974, March), *Polygraph Research in the U. A. Army*, Polygraph, 3 (1), 66–80.

Capps M.H., Ansley N. (1992), *Anomalies. The contributions of the cardio, pneumo, and electrodermal measures towards a valid conclusion*, Polygraph, 21 (4), 321–340.

Dror I., Rosenthal R. (Jul 2008), *Meta-analytically Quantifying the Reliability and Biasability of Forensic Experts*, Journal of Forensic Science, 53 (4), 900–903.

Elaad E., Ginton A., Ben-Shakhar B. (1998), *The Effects of Prior Expectations and Outcome Knowledge on Polygraph Examiners' Decisions*, Polygraph, 27 (4), 288–301.

Krapohl D.J. (1998), *A Comparison of 3- and 7-position Scoring Scales with Laboratory Data*, Polygraph, 27 (3), 210–218.

Matte J.A. (2002), *2002 Supplement to Forensic Psychophysiology Using The Polygraph*, J.A.M. Publications: Williamsville, New York.

Matte J.A., Backster C. (2000), *A Critical Analysis of Amsel's Comparative Study of the Exclusive v. Nonexclusive Comparison Question*, Polygraph, 29 (3), 261–266.

Matte J.A., Backster C. (2007), *Psychological Structure and Theoretical Concept of the Backster Zone Comparison Technique and the Quadri-Track Zone Comparison Technique*, Polygraph, 36 (2), 84–90.

Van Herk M. (1991), *Numerical evaluation: Seven point scale +/–6 and possible alternatives; A discussion*, Polygraph, 20 (2), 70–79.