Psychology, Crime & Law, Vol. 4, pp. 1–16
Reprints available directly from the publisher
Photocopying permitted by licence only

© 1998 OPA (Overseas Publishers Association)
Amsterdam B.V. Published under license
under the Harwood Academic Publishers imprint,
part of The Gordon and Breach Publishing Group.
Printed in India.

THE ROLE OF PRIOR EXPECTATIONS IN POLYGRAPH EXAMINERS DECISIONS

EITAN ELAAD[a,*], AVITAL GINTON[a] and GERSHON BEN-SHAKHAR[b]

[a] *Division of Identification and Forensic Science, Israel Police H.Q., Jerusalem 91906, Israel*

[b] *The Hebrew University of Jerusalem, Israel*

In the present experiment, prior expectations were manipulated to examine whether and to what extent polygraph examiners' judgment is affected by their expectations. Each of seven experienced police polygraph examiners examined four innocent subjects suspected of cheating on an aptitude test. The examiners were led to believe that two subjects are probably guilty and the other two are probably innocent. No relationship was found between the final judgment of the examiners and their prior expectations.

Key words: prior expectations; hypothesis-confirmation bias; polygraph; control question technique.

Psychophysiological detecting of deception (PDD) is an important area of applied psychology. It is important both from the criminal justice perspective (e.g., to find out whether a given person is deceptive regarding a specific, usually criminal, event), and for personal selection purposes (e.g., to discriminate between honest and dishonest individuals among a group of job applicants). Several methods of psychophysiological detection have been developed and used in field practice (see Reid and Inbau, 1977; Saxe *et al.*, 1985). These methods which are based on a comparison between physiological responses to relevant questions (i.e., questions that focus on the issue under investigation) and some form of "control questions" have been the focus of extensive research during the last three decades (e.g., Ben-Shakhar and Furedy, 1990; Lykken, 1981; Raskin, 1989; Reid and Inbau, 1977).

The most common PDD method in field practice is the Control Question Technique (CQT; Reid and Inbau, 1977; Raskin, 1989). Briefly, the CQT is administered in the following stages: First, the examiner becomes familiar with the facts of the case by reading the

* Corresponding author.

written report and by speaking directly to the police investigator who ordered the examination. Typically, relevant background information, such as the suspect's past criminal record, is made available to the examiner. During the next stage, the examiner conducts an extensive pre-test interview in which the examinee is given the opportunity to talk about the offense and present his or her version of the case. The series of questions, to be asked later in the actual examination stage of the polygraph test, is formulated during this pre-test interview through an interaction between the examiner and the examinee. The examiner discusses the formulation of the questions with the examinee and ensures that he or she understands them and can give a direct "yes" or "no" answer to each question. The examiner explains the testing procedure and informs the examinee that the examination is voluntary. The next stage is the actual examination stage during which the examinee is attached to the polygraph.

During the examination stage a series of questions is presented to the examinee while continuously measuring various physiologic reactions. The questions are of the following three types: (a) Relevant questions---direct crime-relevant questions of the "did you do it?" type (e.g., "Did YOU break into Mr. Jones's apartment last Friday night'?); (b) Control questions -focusing on general, non-specific misconducts, similar to the issue under investigation (e.g., "Have YOU ever take something that did not belong to You?"); (c) Irrelevant questions - focusing on completely neutral issues, (e.g., are you sitting on chair?) which are intended to absorb the initial orienting response evoked by any opening question, and to enable rest periods between the more loaded questions. Typically, the whole question series is repeated three or four times. The inference rule underlying the CQT is based on a comparison of the responses evoked by the relevant and control questions. Deceptive individuals are expected to show more pronounced responses to the relevant questions, whereas truthful individuals are expected to show the opposite pattern of responsivity (i.e., more pronounced responses to the control questions).

The CQT raised a major controversy, revolving around its rationale and inference rule, as well as around the empirical question of validity (e.g., Ben-Shakhar and Furedy, 1990; Furedyand Heslegra1 1991 ;Lykken, 1974; 1978; Raskin, 1982; 1989). Some critics of the CQT have argued that the judgments of CQT examiners are contaminated because they know a great deal more than what is reveal through the physiological data gathered during the examination (e.g., Ben-Shakhar, 1991; Ben-Shakhar et al., 1986; Ben-Shakhar a Furedy, 1990). Specifically, polygraph examiners are exposed .to great deal of non-physiological information, such as information

provided to them by the investigators and impressions formed during the pre-test interview and during the test itself. It is impossible to differentiate between the impressions formed by this prior information and those gained from the purely physiological data obtained during the test phase of the polygraph examination procedure. This feature of CQT polygraph examinations, implies that judgments and conclusions derived from the physiological information are contaminated with various kinds of non-physiological information. Contamination is inherent to the CQT, because this procedure is not limited to the psychophysiological data, but rather relies on the whole examiner-examinee interaction, including the pre-test interview.

Contamination may introduce various biases, because the prior information may affect the formulation of the questions and the way they are presented to the suspects. For example, when examiners believe that a given examinee is deceptive, they may present the relevant questions in away that may affect the results in the expected direction. On the other hand, when examiners are under the impression that their suspect is truthful, the control questions might be overemphasized. In addition, prior expectations and beliefs can bias polygraph examiners when they evaluate the polygraph charts and reach a conclusion on the basis of a comparison between the responses of a given suspect to the relevant and control questions. Thus, one implication of the contamination feature of the CQT is that the weight of the strictly physiological information in the polygraph examiner's conclusion is not known.

Ben-Shakhar (1991) formulated an hypothesis based on a combination of the belief bias and the confirmation bias (Evans, 1989). This hypothesis postulates that, before the administration of the CQT, polygraph examiners typically develop a belief or hypothesis based on information gathered from previous, nonpolygraphic interrogations and from the pre-test interview. The physiological information obtained during the test phase of the CQT is then used to test this prior hypothesis. However, the hypothesis-testing process is influenced by a confirmation bias (e.g., Snyder and Swann, 1978; Wason, 1968), or by a primacy effect (e.g., Jones and Goethals, 1972). In other words, it is suggested that when the judge is supposedly considering the "objective" data neutrally for the purpose of diagnosis, or evaluation, he or she is in fact searching it for confirmation of the initial belief or the prior hypothesis which the judge entertains before looking at the data. Some data sources are sufficiently simple or well defined, that they hardly lend themselves to variable interpretation. However, in the case of CQT -polygraphy , where the rich and complex physiological data are subjectively evaluated by examiners, rather than mechanically quantified, it can readily be distorted if the search process is biased

and precommitted. Even if the judge starts out completely uncommitted, early impressions from the data might quickly suggest a hypothesis which would color the rest of the data evaluation. Moreover, under these circumstances a biased search is likely to produce supportive findings-especially if it is untempered by critical attempts to falsify the initial hypothesis, or to pit it against some competing alternative. The richness and vagueness of the information increase the likelihood of finding some confirmatory evidence. Indeed, it is possible that many polygraph charts contain some confirmatory information for almost any possible hypothesis.

Indirect supporting evidence for the confirmation bias was derived from two sources: Barland (1975) reported that in 17 out of 19 cases polygraph examiners who gave evaluations just before the administration of the polygraph test (e.g., on the basis of the prior information available to them), gave the same evaluations after the polygraph test. A second demonstration was provided by a CBS production of a "60 minutes" broadcast in 1986. Three different polygraph firms were independently called to test an alleged theft of a camera and lens from a photography magazine office employing four employees. In fact nothing was stolen from the office, but the polygraph examiners were told it could have only been done by one of the four employees. Each polygraph examiner was told: "it might have been. .." with a different employee named in each case (a decidedly weak fingering). All three polygraph examiners ultimately identified the "fingered" employee as deceptive, and cleared the other "suspects". Moreover, they expressed complete confidence in their decisions. This demonstrates not only that polygraph examiners can go wrong, but that their judgment and decision-making processes may be infected by a systematic and powerful source of bias, a bias caused by contamination.
A recent study (Elaad et al., 1994) provided the only experimental evidence for the confirmation-bias effect in the polygraph-examination context. In this study, prior expectations were manipulated by providing some examiners with the outcome knowledge (e.g., telling them that the suspect had eventually confessed). Examiners were then asked to score and assess polygraph charts from previous examinations. Results indicated that prior expectations had an impact on the examiner's judgments when the charts did not include clear indications of either guilt or innocence. However, when the physiological information contradicted prior expectations, examiners tended to ignore their expectations and base their decisions only on the physiological information. The Elaad et al. (1994) study focused on the effects of the examiner's prior expectations on the evaluation and scoring of the polygraph charts. However, prior expectations can color the entire

polygraph examination, and affect the choice and the formulation of the control questions, as well as the manner by which questions are presented to the suspects. These factors may in turn affect the physiological reactions in the directions of the prior expectations. For example, examiners may put a strong emphasize on the control questions when they are convinced that the suspect is innocent, and thus produce a chart which is consistent with their prior conviction.

Darley and Gross (1983) made a distinction between a "cognitive confirmation effect" (i.e., expectancy-confirmation effects that occur in the absence of any interaction between the perceiver and the target person), and a "behavioral confirmation effect" (i.e., where expectations affect the behavior toward a target individual, such that expectancy-confirming behaviors are elicited from this individual). The study reported by Elaad et al. (1994) dealt only with the first type of confirmation bias, because the examiners were exposed only to the information included in the polygraph charts and not to the behavior of the suspects. But prior expectations of polygraph examiners may produce effects similar to the "behavioral confirmation effect" (Snyder and Swann, 1978), or the "interpersonal expectancy effect" (e.g,. Rosenthal and Rubin, 1978).

To explore the "behavioral confirmation effect" in the CQT context, it is necessary to conduct an entire CQT examination while manipulating the examiners' prior expectations. The present study was designed to meet this requirement.

METHOD

Participants

Two groups of participants took part in this experiment: (I) Twenty-eight Hebrew University undergraduate students, 10 males and 18 females, aged between 18 and 30 (mean 23.8), volunteered to participate in the study for course credit or payment. They were offered up to 55 NIS (about $22) for participating in a two-session experiment. (2) Seven polygraph examiners employed by-the Israel National Police, with at least 2.5 years of experience in operating the polygraph.

Apparatus

First, students underwent a psychological experiment in the laboratories of the Hebrew University. Second, polygraph tests were conducted in two police polygraph examination rooms located at the police

headquarters, where polygraph examinations of actual criminal cases are conducted routinely. Each room had bare walls without windows, and was sparsely furnished with a table and two chairs.
A Lafayette Factfinder field model polygraph was installed in one room, and a Lafayette Ambassador in the other. The polygraphs recorded the following physiological measures: (1) Dual respiration (thoracic and abdominal), recorded by pneumatic tubes positioned around the thoracic area and abdomen. (2) Skin resistance response (SRR), recorded by two stainless steel electrodes attached to the volar side of the index and fourth fingers of the subject's left hand. (3)Cardiovascular activity, recorded by a pneumatic blood pressure cuff positioned around the upper right arm, and inflated to a pressure between 40-50mmHg.

Procedure

To resemble field practice, the experiment took place in two sessions separated by several days (5.7 on the average). In the first, the experimenter (E) accompanied the students to an experimental room and told them to sit at a table with their back against a one-way mirror which separated the experimental room from a control room. The students were told that the aim of the experiment was to examine their capacity to perform quickly and accurately under time pressure. Before leaving the room, E gave the students a ball point pen and an aptitude test and told them that they would receive further instructions from the control room through an intercom. Three minutes after given permission to start, the students were instructed to put down the pen and turn the page over. At this point most students had answered about one third of the questions. E returned to the experimental room, removed the test sheet and placed on the table a binder which contained another paper-and-pencil test. This test was previously used by Ginton, et al., (1982) to establish a reliable criterion for deception. The students were instructed to enter their solutions to the test items on a 5 x 5 matrix according to certain written instructions. Subjects were told that the time was short and only people with superior intelligence, strong will, and emotional self control can succeed in scoring at least 15 correct answers. E added that typically only 20% of the examinees can achieve such high scores. Subjects were further informed that they would not be permitted to continue to the second phase of the study (in which they could earn up to 50 NIS) unless they receive a minimum (undefined) score.
Several sheets were connected together in the binder in the following order: An instruction sheet, which was separated after the subject read

the instructions and understood the task, was placed on the top. Next, there was a 5 x 5 matrix paper which served as an answer sheet. Beneath the answer sheet there was a scoring sheet, beneath which a chemical page was concealed. The chemical sheet received an impression of anything written on the answer sheet. The last sheet was another scoring sheet identical to the previous one.

To create variation in the performance of the students on the test, half of them were given five minutes to complete the test, while the others were allowed 10 minutes. After this predefined time interval (either five or ten minutes), the students were told to put the pen down, separate the answer sheet from the binder, and turn it over on the instruction sheet. E entered the experimental room, took the binder and instructed the students to take the answer sheet and follow him to an adjacent room where the task will be completed. This room had no one way mirror, and students could feel free to cheat without being observed. E returned to the control room, separated the chemical sheet from the binder, attached the correct answer keys and returned to the subject. The binder with the correct answer keys was then handed out to the students who were asked to score their own test. E left the room and returned to the control room where he assigned the actual score of the student according to the chemical sheet. Meanwhile, students had an opportunity to cheat and improve their scores by adding correct solutions to empty cells, or by changing their original wrong answers. A comparison of the scores finally assigned by the students with the scores obtained from the chemical copy made clear whether the student had cheated or not. It turned out that five of the students cheated while scoring their test. These students were thanked, debriefed and told that they would not continue to the next stage of the experiment.[1] Five noncheaters who scored 0 were also debriefed and released from the second stage. They were informed that their scores were less than that required.

The remaining 28 noncheaters were invited to the next stage. They were informed that the purpose of the study was to learn about the accuracy of the, polygraph (i.e., to find out whether the polygraph examiners would be able to determine whether the students did or did

[1]There are no formal ethical guidelines for conducting research in the Israel police. However as psychologists we are committed to the ethical guidelines of the AP A that is: Deception should be avoided unless: (a) It is scientifically essential to the study, and this is the case in the present study. (b) It is estimated that the subjects will not be distressed by the deception when debriefed. However, to minimize potential ethical problems, we decided to exclude cheaters from this study. It would be of interest to compare, in future experiments, the effects of expectations on the outcomes of polygraph examinations conducted with known cheaters and known non-cheaters.

not cheat on the test). They were promised a special bonus of 50 NIS if they would produce a truthful polygraph outcome (in the case of a deceptive result they would receive only 10 NIS). At this point they knew perfectly well that no one else is aware of the fact that they are innocent of cheating on the aptitude test. Such a state is typical for innocent suspects in actual polygraph examinations, but not for innocent participants in mock crime experiments who know very well that the experimenter is aware of their innocence. The students were further told that the polygraph examiners think they are conducting a genuine polygraph examination whose outcome would be passed on to the disciplinary committee of the Hebrew University. They were asked to collaborate by expressing concern about the possibility of being accused by the disciplinary board, and were told to disclose their score on the aptitude test. All 28 students agreed to take the polygraph test, and gave their written consent.

Finally, an appointment was made for each student to take a polygraph test at the police headquarters. There were two equal-size groups of students, those who scored less than 15 (most of them were from the group that were allowed only five minutes to complete the task) and those who scored 15 or more (most of them were from the ten-minute group). After completing the polygraph examination each student was debriefed about the real nature of the study. Most of the students expressed interest in the study's results. No one reported being distressed by the fact that they were misled in the first stage of the experiment.

Examiners Manipulation

Before examining the students, examiners were told that the study was designed as a replication of the Ginton et al. (1982) study, using the same method on undergraduate students. The replication was said to be necessary because of the many changes in polygraph procedures and in the population of examiners since 1982. All examiners were familiar with the Ginton et al. (1982) study which is frequently discussed at the polygraph laboratory. Nonetheless, this study was described once again to the examiners just prior to the experiment. The examiners were informed about the chemical copy which provided a reliable criterion for the classification of students as cheaters versus noncheaters, and were told that the examinees were being accused by the university authorities of cheating on a test which was part of an experiment. All examiners were perfectly aware that they were being tested for their accuracy under real life conditions. Although the examiners were told that even if the students were found deceptive, they would not be brought before the disciplinary board, several

examiners expressed discontent with the task, saying that the threat of facing the disciplinary committee was unfair. The examiners were informed about the score each student achieved on the test, and were led to believe that the score of 15 or more is most unlikely unless the student cheated on the test. To increase the credibility of our manipulation, the first examinee was an accomplice who told the examiner that her score on the aptitude test was 20, and then confessed that she had cheated on this test. After completing the examination of the 28 subjects (4 examinations for each of the 7 examiners), the examiners were debriefed as to the real purpose of the study. The responses were of interest and surprise and no one expressed discontent about the deception.

Manipulation Check

To check whether the expectancy manipulation had an impact, all the remarks made by the examiners before administering the CQT were written down. For example, remarks such as: "There is no use to examine this student, because he (or she) is going to yield a truthful outcome anyway", were made for three low scoring subjects; and remarks such as: "Is it possible to achieve such a high score without cheating?" were made about five high-scoring subjects. These remarks may suggest that the distinction between high and low scoring students and its implications were clear to the examiners. To achieve a more systematic assessment of the effectiveness of the expectancy manipulation, examiners were asked to write down their impressions of the subjects during the pre-test interview. Two independent examiners, who did not participate in the experiment, were asked to classify each examiner's notes into one of the following five categories: (1) The subject is cooperative and confident. (2) The subject is concerned with the control questions. (3) The subject is concerned, uneasy, and seems anxious, with no indication that this concern is related to the control questions. (4) The subject seemed untruthful. (5) Missing information, or irrelevant notes. The two judges were in complete agreement regarding their classifications which are displayed in Table I.

The two classification categories which indicated that the subjects are cooperative or concerned with the control questions, suggest that the examiner expected the subject to be truthful, and the categories of general stress and disbelief suggest an expectation of deceptiveness. Table 1 suggests that seven of the low scoring subjects and only a single high scorer were expected to be truthful. Similarly, seven high scorers and only a single low scoring subject were expected to be deceptive.

Table 1 Classification of the notes made by the examiners during the pre-test interview

	Cooperative	Stressed by the control questions	General stress	Disbelief	Other
High scorers	0	1	5	2	6
Low scorers	4	3	1	0	6

Record Quantification

Evaluation of the polygraph records was made according to the numerical scoring procedure which was proposed originally by Backster (1963). This procedure is a routine at the Polygraph Unit of the Israeli Police Force. According to this scoring procedure, two or three pairs of relevant-control questions are identified in each polygraph chart, and numbers (-3, -2, -1,0,1,2,3) are assigned to each pair for each physiological measure. The absolute value of the assigned number reflects the magnitude of the difference between the responses evoked by the two questions within the pair (e.g., -3 or +3 reflect a very large difference, -l or + 1 reflect a small difference and 0 reflects no difference), and the sign of the assigned number reflects the direction of the difference, such that positive numbers are associated with a pattern of larger physiological reactivity to the control question, and negative numbers reflect the opposite pattern. These numbers are then summed up across question pairs, across physiological measures and across polygraph charts to yield a total score. Thus, if for example, a polygraph examination is based on three charts and three physiological measures and if two pairs of relevant-control questions are identified for each chart, then the total score ranges between -54 and +54.

Using this scoring procedure, each examiner scored his own records. In addition, the records were given to two experienced examiners for a blind scoring. These two examiners, who were not scheduled to take part in the study, had 18 and 15 years experience in conducting polygraph examinations. While scoring the records, these scorers were unaware of the fact that all students' were innocent. They were also unaware of the score the students assigned to themselves in the aptitude test. Furthermore, they were unaware of the contents of the questions presented during the polygraph test, as well as the outcome of this test. They identified the relevant and control questions according to their corresponding serial numbers. One of these examiners scored 18 records and the other scored the remaining 10. The correlation coefficients between the total scores assigned by the original

examiners and the additional scoring were 0.92 and 0.83. This suggests that while scoring the records, the examiner who conducted the test was not strongly affected by prior expectations and by other external factors such as the subject's behavior during the test.
Typically, the following classification rule is used in a CQT: If the global score exceeds +5, an NDI (no deception indicated) classification is reached; if the global score is less than -5, the polygraph record is classified as DI (deception indicated); and if the global score ranges between -5 and +5, the record is classified as inconclusive.

RESULTS

The frequencies of the three classification categories (DI, NDI, and Inconclusive) under each expectation condition, as well as the classifications made by the two blind scorers. are displayed in Table 2. Inspection of Table 2 suggests that there are no differences between the distributions of the three classifications under the two experimental conditions (12 out of 14 subjects were classified as innocents under the guilt-expectation condition, and II out of 14 subjects were classified as innocents under the innocence-expectation condition). Similar results were obtained when the classifications made by the blind scorers are considered. The table also indicates that across experimental condi - tions, the rate of true negative classifications was 82.1 %, with a false positive error rate of only 7.1% and an inconclusive rate of 10.7.% However, the classifications' distributions are extremely skewed, and therefore it is difficult to test whether the differences between the two conditions are statistically significant (the vast majority of the cases were classified as innocents, and the frequencies of the other categories were much too small).
Therefore, we conducted statistical tests on the continuous scale of the total scores, which may serve as a more sensitive measure of the

Table 2 Frequencies of CQT classifications of innocent subjects obtained under high, low, and no expectancy of cheating.

Polygraph Classificatiom	*Original examiners*			*Blind scorers*		
	NDI	*Inconclusive*	*DI*	*NDI*	*Inconclusive*	*DI*
Expectancy of cheating						
High	12	1	1	12	1	1
Low	11	2	1	10	3	1
Marginal sum	23	3	2	22	4	2

expectancy manipulation than the final classifications made by the examiners. The mean total scores under each experimental condition was computed for each examiner. These means as well as the overall means across examiners, and the mean scores obtained by the 2 blind scorers are displayed in Table 3. Inspection of Table 3 reveals that the small differences between the two experimental conditions are not in the expected direction (the mean global score for the cheating-expectation condition is larger than the mean score obtained for the innocent expectation condition).

To examine whether the polygraph examiners might be affected only by strong expectations, we analyzed the extreme cases (students who scored above 19 on the test versus students who scored below 9). The mean polygraph scores for these two sub-groups, were 11.1 for the high scorers (expected cheaters), and 7.17 for the low scorers (expected non-cheaters).

DISCUSSION

The results of the present study indicate that manipulating prior expectations had no effect on the examiners' judgment when they conducted the CQT from its outset, and were allowed to interact with the suspects, rather than just analyze the charts. The relatively high correlation between the scores assigned by examiners exposed to the experimental manipulation and those that were not exposed to it further suggests that the examiners expectations did not affect their scoring. Furthermore, even when only extreme cases were considered,

Table 3 Mean CQT scores computed for each polygraph examiner under the two experimental conditions, and the corresponding scores computed for the blind scorers

Examiners	*Original examiners* *Expectancy of cheating*		*Blind scorers* *Expectancy of cheating*	
	High	*Low*	*High*	*Low*
1	19.5	12.0	13.0	11.0
2	8.5	11.5	11.0	10.5
3	0.0	6.0	2.5	10.0
4	6.0	11.5	6.5	7.5
5	14.0	9.0	10.5	10.0
6	17.5	17.5	19.0	16.5
7	17.5	3.5	18.5	3.0
Marginal means	11.86	10.14	11.57	9.79

the expectations had no effect on the outcomes of the polygraph examinations. These results are consistent with Smith and Kida (1991) who argued that experts performing familiar tasks are not subject to the same kinds of biases found when students perform abstract, unfamiliar tasks.

However, it is possible that the manipulation of the expectancy factor was not sufficiently strong, or that the examiners were skeptical about the true nature of the experiment. To examine this possibility we conducted a manipulation check, and our analysis of the examiners' impressions of the subjects during the pre-test interview, displayed in Table 1, suggests that the examiners' impressions were affected by the expectancy manipulation. Another indication of the impact of the expectancy manipulation was obtained at the end of the polygraph examination when examiners were asked to write down their impressions of the students. Among the 14 students who achieved high scores on the aptitude test, nine made good impressions on the examiners. All of them were classified as truthful on the basis of the CQT, although in five of these nine cases the examiners expressed concern about the accuracy of their classification. Two high scoring students made bad impressions on the examiners, and were classified as deceptive and inconclusive. In the three remaining cases the impressions were either missing or unclear. Among the 14 low-scoring students, nine made good impressions on the examiners. In eight of these cases the examiners expressed confidence in the truthful outcomes of the polygraph test, while the ninth was classified as inconclusive. Two low scoring students made bad impressions, and were classified as "inconclusive", and as "deceptive". In three additional cases who were classified as truthful, the impressions were either missing or unclear.

This analysis indicates that the classifications made by the examiners at the end of the examination correspond to their impressions of the charts and their impression of the examinee, but not to their prior expectations. It seems that the formation of polygraph examiner's impressions regarding the examinee's truthfulness is a complex process, and that impressions and beliefs may change during the course of the test. The examiner integrates pre-test expectations derived from information gathered before the test with verbal and nonverbal behavior of the examinee during the pretest interview and with physiological and non-physiological information gathered during the administration of the test. It is possible that when both behavioral and physiological clues contradict the prior expectations, the impact of these expectation is minimized or lost.

Thus, the present results are inconsistent with the hypothesis that judgments of CQT -polygraph examiners reflect primarily the prior

expectations and beliefs held by the examiners before conducting the CQT. Because the CQT is based, to a large extent, on the examiner-examinee interaction, it is possible that prior expectations formed by prior information interact with the impressions formed by the examiners during the course of the examination. This complex interaction may affect the interpretation of the physiological information and the final conclusion reached by the examiner. This may account for the differences between the present results and those reported by Elaad et al. (1994). While the latter study indicated that prior expectation have an impact on the process of chart interpretation and scoring when the charts do not contain information that clearly contradict these expectations, the results of the present study suggest that prior expectations have no impact when the examiner conducts the CQT from its outset, and interpret the polygraph charts. If it is assumed that prior expectations interact with the behavior of the suspect during the examination, and that expectations and beliefs may change during this process, then it is not surprising that prior expectation affect examiners' judgment when these examiners have access only to the polygraph charts but are not exposed to the suspects' behavior. In other words, it is suggested that the polygraph examiners may rely on their impressions from the subjects' behavior, and when these impressions contradicted the prior expectations based on the alleged results on the aptitude test, they relied on their impressions and disregarded the information about the test score. This possibility gains some support from the post-experimental descriptions of the examiner' impressions of the students.

Several alternative accounts may be suggested to explain the presen findings and the differences between them and those obtained by Elaad et al. (1994). First, it is possible that prior expectations have no impact on CQT polygraph tests conducted on innocent subjects, but only on chart interpretation. However, it is difficult to draw such a conclusion from a single experiment, and it is clear that further experimentation will be required before this issue is resolved. Further studies should use different methods for manipulating expectations and beliefs, and different groups of polygraph examiners. If this account is true, it is still not clear.why the polygraph examiners were not affected by their expectations when they scored the polygraph charts in the final phase of the examination. It is possible that most of the records included clear indications of innocence, and when these indications contradicted the prior expectations, the examiners tended to rely on the specific physiological information. This account is consistent with the findings of Elaad et al. (1994).

Finally, it should be remembered that the present experiment differs in many ways from the real life context of polygraph examination. In

particular, examiners in this experiment knew perfectly well that their ability as professional polygraphers is being tested, because the truth was known to the experimenter. It is possible that in such a situation examiners tend to be more cautious than in the typical CQT examination context where their performance cannot be readily evaluated.

ACKNOWLEDGMENTS

This research was supported by a grant from the Sturman Center for Human Development at the Department of Psychology, The Hebrew University of Jerusalem. We wish to thank the polygraph examiners of the Israeli police for their assistance in the execution of this study, and Ygaal Atali, Sharon Liberty and Michal Tur-Kaspa for their help in the data collection.

References

Backster, C. (1963). Polygraph professionalization through technique standardization. Law and Order **11**, 63-64.

Barland, G. H. (1975). Detection of deception in criminal suspects: A field validation study. Ph.D. dissertation, University of Utah.

Ben-Shakhar, G. (1991). Clinical judgment and decision making in CQT polygraphy: A comparison with other pseudoscientific applications in Psychology. Integrative Physiological and Behavioral Science, **26**, 232-240.

Ben-Shakhar, G., Bar-Hillel, M. and Lieblich, I. (1986). Trial by polygraph: Scientific and juridicial issues in lie detection. Behavioral Science and the Law, **4**, 459-479.

Ben-Shakhar, G. and Furedy, J. J. (1990). Theories and Applications in the Detection of Deception: A Psychophysiological and International Perspective. New York: Springer-Verlag.

Ben-Shakhar, G., Lieblich, I. and Bar-Hillel, M. (1982). An evaluation of polygrapher's judgments: A review from a decision theoretic perspective. Journal of Applied Psychology,**67**, 701-713.

Darley, J. M. and Gross, P. H. (1993). A hypothesis-confirmation bias in labeling effects. Journal of Personality and Social Psychology, **44**, 20-33.

Elaad, E., Ginton, A. and Ben-Shakhar, G. (1994). The effects of prior expectations and outcome knowledge on polygraph examiners' decisions. Journal of Behavioral Decision Making, **7**, 279-292.

Evans, J. St. B. T. (1989). Bias in Human Reasoning: Causes and Consequences. London: Lawrence Erlbaum.

Furedy, J. J. and Heslegrave, R. J. (1991). The forensic use of the polygraph: A psychophysiological analysis of current trends and future prospects. In: J. R. Jennings, P. K. Ackles and M. G. H. Coles (Eds.), Advances in Psychephysiology, **4**, Greenwich, CT: JAI Press.

Ginton, A., Daie, N., Elaad, E. and Ben-Shakhar, G. (1982). A method for evaluating the use of the polygraph in a real life situation. Journal of Applied Psychology, **67**, 131-137.

Jones, E. E. and Goethals, G. (1972). Order effects in impression formation: Attribution context and the nature of entity. In: E. E. Jones et al. (Eds.), Attribution: Processing the causes of behavior. Morristown, NJ: General Learning Press.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. Journal of Applied Psychology, **44**, 258-262.

Lykken, D. T. (1974). Psychology and the lie detection industry. American Psychologist, **29**, 725-739.

Lykken, D. T. (1978). Uses and abuses of the polygraph. In: H. L. Pick (Ed.), Psychology: From Research to Practice. New York: Plenum Press.
Lykken, D. T. (1981). A Tremor in the Blood: Uses and Abuses of the Lie Detector. New York: McGraw-Hill.
Raskin, D. C. (1982). The Scientific basis of polygraph technique and their uses in the judicial process. In: A. Trankell (Ed.), Reconstructing the Past: The Role of Psychologists in the Criminal Trial. Stockholm, Sweeden: Norsted & Soners.
Raskin, D. C. (1989). Polygraph techniques for the detection of deception. In: D. C. Raskin (Ed.), Psychological Methods in Criminal Investigation and Evidence. New York: Springer Publishing Company.
Reid, J. E. and Inbau, F. E. (1977). Truth and Deception.. The Polygraph ("Lie Detection") Technique. Baltimore: Williams & Wilkins.
Rosenthal, R. R. and Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. The Behavioral and Brain Sciences, **3**, 377--415.
Saxe, L., Dougherty, D. and Cross, T. P. (1985). The validity of polygraph testing: Scientific analysis and public controversy. American Psychologist, **40**, 355-366.
Smith, J. F. and Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. Psychological Bulletin, **109**, 472--489.
Snyder, M. and Swann, W. J. Jr. (1978). Behavioral confirmation in social interaction: From social perception to social reality. Journal of Experimental Psychology, **14**, 148-162.
Wason, P. C. (1968). On the failure to eliminate hypotheses--a second look. In: P. C. Wason and P. N. Johnson-Laird (Eds.), Thinking and Reasoning, Penguin, Harmondworth, Middlesex.