Journal of Police Science and Administration
Copyright ©1985 by IACP, Inc.

Volume 13, No. 4
Printed in U.S.A.

Validity of the Psychological Stress Evaluator: A Field Study

Israel Nachshon, Eitan Elaad, and Tuvia Amsel

Israel Nachshon is Professor of Psychology at The Department of Criminology, Bar-Ilan University, Ramat-Gan, Israel. He holds a Ph.D. degree in psychology from Northeastern University. Currently he is interested in face perception and in the neuropsychology of aggressive behavior in man.

Eitan Elaad is Director of Research at the Scientific Interrogation Unit, Division of Criminal Identification, Israel Police Headquarters, Jerusalem. He holds an M.A. degree in psychology from The Hebrew University, Jerusalem. His areas of interest are psychophysiology, social psychology, and scientific interrogation.

Tuvia Amsel is the owner-director of Amsel Investigation and Interrogation, 1 Ben-Gurion Boulevard, Tel-Aviv. He holds a B.A. degree in criminology from Bar-Ilan University, Ramat-Gan, Israel. Also, he is a graduate of the Dektor School of PSE, Springfield, Virginia, and of Zonn Institute of Polygraph, Miami, Florida. He is interested in interrogation and lie detection.

The Psychological Stress Evaluator (PSE) is a device developed by Dektor Counterintelligence and Security, Incorporated, for the detection of emotional stress in the voice. The PSE is said to detect changes in the infrasonic frequency modulation, which modulates about a central frequency, and is present in both the vocal chord sounds and in the formant sounds (Dektor 1974; Inbar and Eden 1976). These modulations, which vary between 5Hz and 20Hz are supposedly "physiological tremors" or "microtremors," which are involuntary ripples superimposed upon voluntary contractions of all striated muscles, including those involved in vocalization (Lippold 1970; Papcun 1973). These modulations are controlled by the central nervous system, and hence are suppressed during stress. Their suppression may therefore serve as an index of stress (Dektor 1974).

However, since the invention of the PSE a controversy has been going on concerning the usefulness of vocal cues as stress indicators in general, and the validity of the PSE in particular (Holden 1975). Detailed reviews of the pertinent theories, modes of operation, and findings may be found in a number of recent publications (Hollien 1980; Horvath 1982; Nachshon and Feldman 1980; Vandecar *et al.* 1980). From these reviews it is evident that in contradistinction to reports made by Dektor and others affiliated with the PSE or by police officers and private practitioners, scientifically-oriented, rigorously-controlled studies have all failed to demonstrate the validity of the PSE as a detector of stress in the human voice.

To begin with, the theoretical basis of the PSE was challenged by findings showing that vocal cues, such as microtremors, may not be valid for the detection of emotional stress (see Hollien 1980; Nachshon and Feldman 1980; Suzuki *et al.* 1979). Furthermore, controlled experimental studies (Barland 1978; Brenner *et al.* 1979; Elaad 1980; Horvath 1978, 1979; Herbold 1977; Kubis 1974; Lynch and Henry 1979; Nachshon and Feldman 1980) have all failed to establish PSE validity for the detection of stress. Vandecar *et al.* (1980) reportedly found in one experiment that state anxiety was detectable by the PSE, but they failed to replicate their finding in a slightly different experiment.

However, as has been pointed out concerning both polygraph (Abrams 1972; Barland and Raskin 1973; Horvath 1976) and PSE testing (Dektor 1974), laboratory and field situations are considerably different from each other as to warrant a reexamination of PSE validity in a field situation. In particular, it was assumed (Horvath 1976) that field situations may be more stress-provoking than laboratory situations, since in the former the subject is personally involved, whereas, in the latter, he is only playing a game. It is therefore conceivable that a vocal stress evaluator,

such as the PSE, might be invalid for low stress situations, but rather sensitive for high stress situations (Barland 1978; Nachshon and Feldman 1980; Rockwell *et al.* 1979). Horvath's (1982) arguments with the manufacturers of the PSE about its applicability notwithstanding, it might still be worthwhile to test its validity under field conditions. Testing 66 criminal suspects, Barland (1975) found no relationship between polygraph- and PSE-based decisions concerning suspects' deception or truthfulness. More recently, the Department of Commerce in the Virginia State Police conducted a field study (cited in Horvath 1982) in which Dektor-trained raters analyzed 40 PSE records of criminal suspects. Correspondence between the polygraph- and PSE-based decisions reached 39 percent—only slightly above chance level (33 percent). Finally, testing 56 criminal suspects, Nachshon and Feldman (1980) similarly failed to find polygraph-PSE correspondence in final decisions.

Despite the negative findings, Nachshon and Feldman (1980) observed that under both laboratory and field conditions, reliability and validity of PSE chart evaluations seemed to be consistently lower for low stress (card-tests) than for high stress situations (presentation of horror pictures in the laboratory study and polygraph examinations in the field study). As suggested by Hollien (1980) and implied by Rockwell *et al.* (1979), the PSE might therefore be relatively sensitive to emotional arousal only under very stressful conditions, such as those prevalent in criminal interrogations.

Nachshon and Feldman's field study was criticized by Amsel (1979) on a number of technical grounds concerning the procedure used (questioning sequence was good for polygraph, but not for PSE examination), the mode of tape recording (Mode II, rather than Mode III was used), and the PSE evaluators' expertise.

The present field study was therefore conducted, in which those points of criticism were taken into consideration to the full satisfaction of the critic, who actually collaborated in the design and execution of the present study.

METHOD

Subjects

Forty criminal suspects (36 males and 4 females), who were undergoing polygraph examinations, served as subjects. Their ages ranged from 16 to 40 years. They were suspected of committing the following crimes: murder, rape, theft, burglary, false alarm, and cheating on an insurance company. All were interrogated by the control-question procedure; 27 by Reid and Inbau's (1966) technique, and 13 by Backster's (1969) technique.

Material and Procedure

The subjects were interrogated individually in a quiet room by five experienced polygraphers from the Section for Scientific Interrogation, Division of Criminal Identification, Israel Police Headquarters, Jerusalem. Subjects could be observed through a one-way mirror. Their galvanic skin responses (GSR) were recorded by two stainless steel electrodes attached to the right index and forefingers. Chest and abdominal respiration was measured by pneumatic rubber hoses. Finally, blood pressure and heart rate (cardiovascular responses) were measured by a cuff attached to the left arm, which was inflated up to 60 mm. mercury. Recording was done on a Lafayette polygraph (Model 76056). In addition, subjects' verbal responses were also picked up by a hidden microphone placed in front of them and delivered to a tape recorder (Uher, Model 4000, Report IC) for recording and subsequent PSE analysis. For best results, the tape was run at a speed of 19 inches per minute for recording and at 4.7 inches per minute for replay. The PSE expert (TA) participated in the suspects' interrogations, pretest interviews, and formulation of questions. He also watched the actual examination through a one-way mirror. However, he was excluded from the polygraph chart evaluation. Similarly, the PSE chart evaluation was done by him alone, without the participation of the polygraph examiners. In both cases evaluation was done globally, by inspection; no formal scoring procedure was used.

RESULTS

Scoring Criteria and Data Analysis

Final decisions based on the polygraph and PSE charts were first compared. However, these decisions are affected not only by chart evaluation, but also by other factors such as previous records, the pretest interview, and the suspect-examiner relationships (Barland and Raskin 1973).

Hence, a blind procedure was needed in order to measure the degree of correspondence between polygraph and PSE chart evaluations. Consequently, adjacent pairs of PSE response charts were presented to the PSE experts for evaluation. Each pair consisted of one

relevant and one control question. The evaluator knew neither the category of each question in a pair, nor the identity of the responder. His task was to identify the more stressful response in each pair (when the stressful response was unidentifiable, the pair was marked "inconclusive"). Following the completion of the evaluation of a given chart, frequency of stressful responses was counted separately for the relevant and for the control questions. When more stressful responses were identified for the relevant questions, the suspect was considered a deceiver; when more stressful responses were identified for the control questions, he was considered a truth-teller. Coding and decoding of the responses given to the control and relevant questions was done by the laboratory secretary, who was the only person in possession of that information.

A similar procedure was used for polygraph chart evaluation, for each channel separately. However, because these charts were taken from the police files, they could not be cut for paired evaluation. Consequently the response sequence was apparent, and during the evaluation of a particular response, part of the simultaneous response recorded on an adjacent channel could be seen. Nonetheless, chart evaluators knew neither the identity of the responders, nor the category of questions.

Six polygraph examiners evaluated the polygraph charts. Each chart was evaluated twice. First, two examiners evaluated one channel in each chart (that is, respiration in a third of the charts, GSR in another third, and blood pressure in yet another third), and then reevaluated those channels in half of the charts; the channels in the other half were evaluated by the other examiner. This procedure was repeated for the other two thirds of the charts and for the other four examiners.

Similarly, the PSE expert evaluated the PSE charts twice (about 3 weeks later he reevaluated the charts for the third time, but since the results of these evaluations did not differ significantly from those of the first two, they were excluded from analysis).

Charts of six subjects were excluded from analysis because of technical defects such as poor reading and the like. Consequently, data analyses were only performed on the charts of the remaining 34 suspects (for the different channels, the number of paired evaluations for the 34 suspects ranged from 134 through 189).

Reliability of Blind Evaluations

Before comparing polygraph and PSE chart evaluations, the reliability of these devices was established. For both polygraph and PSE, reliability was calculated for conclusive decisions only, in terms of correspondence between the two evaluations made. The mean interevaluation correspondence was 78.62 percent for respiration, 92.63 percent for the GSR, 80.65 percent for the cardiovascular channel, and 89.24 percent for the PSE. All correspondence scores were significantly ($p.<.001$) different from chance. These results seemed to justify further analysis of PSE validity.

Validity of PSE Chart Evaluation

Assuming that polygraph chart evaluations under field conditions are valid for detection of deception (see Horvath 1976), PSE validity was established by comparing PSE and polygraph chart evaluations.

1. Comparison of Blind Evaluations

First, decisions based on blind evaluations of polygraph and PSE charts were compared. In this comparison only unanimous polygraph decisions and consistent PSE evaluations with no more than one reservation were included. For the various comparisions, the number of responses ranged from 67 to 101.

Correspondence of chart evaluations for the various channels was calculated separately for evaluations of truthfulness (stronger responses to control than to relevant questions) and for evaluations of deception (stronger responses to relevant than to control questions). The data are summarized in table 1, which shows that, for the polygraph, interchannel correspondence ranged from about 30 percent (GSR, truthfulness vs. PSE) to about 72 percent (GSR, deception vs. respiration), with a mean percentage score of 48.44. For the PSE it ranged from about 33 percent to about 55 percent, with a mean percentage score of 43.63. Considering the various channels separately, table 1 further shows that for respiration and GSR, interchannel correspondence was always higher for evaluations for deception than for evaluations of truthfulness, whereas for the PSE it was always higher for evaluations of truthfulness than for evaluations of deception.

Regarding cardiovascular responses, correspondence was higher for evaluations of deception when compared with respiration, but higher for evaluations for truthfulness when compared with GSR and PSE. Differences between evaluations of deception and of truthfulness were smallest for cardiovascular responses (mean percentage: 5.01) and biggest for the GSR (mean percentage: 28.39). Similar differences were obtained for respiration and PSE (mean percentages: 10.75 and 8.2 for respiration and PSE, respec-

tively). However, as table 1 indicates, only a few of the results are significantly different from chance (50 percent).

2. Comparison of Blind Evaluations with Final Decisions

A second comparison involved the blind evaluations with the final polygraph and PSE decisions. Since a preliminary analysis showed that, for both polygraph and PSE evaluations, reservations on the final decisions did not affect correspondence, only "highly confident" final decisions without any reservation were analyzed.

The data are summarized in table 2, which shows that correspondence between blind evaluation and final decisions was similar for the polygraph and the PSE. For the polygraph, mean percentages were 55.41 and 58.08 for decisions of truthfulness and deceptions, respectively; for the PSE, they were 50.87 and 49.49 for decisions of truthfulness and deception, respectively. For the polygraph, correspondence ranged from about 37 percent (respiration, deception) to about 81 percent (GSR, deception). For the PSE, it ranged from about 38 percent (respiration, deception) to about 64 percent (PSE, truthfulness).

3. Comparison of Final Decisions

In the last analysis, final polygraph and PSE decisions were compared. In 17 cases, both polygraph and PSE decisions were conclusive without any reservation. Among those, in 16 cases (94.11 percent) the

TABLE 1

CORRESPONDENCE IN BLIND EVALUATIONS

Blind evaluation		Criterion							
		Respiration		GSR		Cardiovascular		PSE	
		Truthful	Deceitful	Truthful	Deceitful	Truthful	Deceitful	Truthful	Deceitful
POLYGRAPH									
Respiration	Total number of cases	—	—	29	43	35	32	41	63
	Number of agreements	—	—	15	31**	13	14	16	21
	Percent correspondence	—	—	51.72	72.09	37.14	43.75	39.02	33.33
GSR	Total number of cases	27	45	—	—	29	49	29	67
	Number of agreements	15	31*	—	—	15	22	16	29
	Percent correspondence	55.56	68.89	—	—	51.72	44.89	55.17	43.28
Cardiovascular	Total number of cases	35	32	42	36	—	—	51	50
	Number of agreements	13	14	15	22	—	—	25	21
	Percent correspondence	37.14	43.75	35.71	61.11	—	—	49.02	42.00
PSE	Total number of cases	48	46	54	42	54	47	—	—
	Number of agreements	16	21	16	29*	25	21	—	—
	Percent correspondence	33.33	45.65	29.62	69.04	46.29	44.68	—	—

*$p<.05$
**$p<.01$

TABLE 2

CORRESPONDENCE BETWEEN BLIND EVALUATIONS AND FINAL DECISIONS

Blind Evaluation		Final Decision: Truthful	Final Decision: Deceitful
		Polygraph	
POLYGRAPH			
Respiration	Total number of cases	43	19
	Number of agreements	20	7
	Percent correspondence	46.51	36.84
GSR	Total number of cases	40	21
	Number of agreements	24	17*
	Percent correspondence	60.00	80.95
Cardiovascular	Total number of cases	35	19
	Number of agreements	22	10
	Percent correspondence	62.86	52.63
PSE	Total number of cases	44	22
	Number of agreements	23	13
	Percent correspondence	52.27	61.90
		Psychological Stress Evaluator	
POLYGRAPH			
Respiration	Total number of cases	81	13
	Number of agreements	35	5
	Percent correspondence	43.21	38.46
GSR	Total number of cases	72	14
	Number of agreements	29	8
	Percent correspondence	40.28	57.14
Cardiovascular	Total number of cases	72	16
	Number of agreements	40	8
	Percent correspondence	55.56	50.00
PSE	Total number of cases	96	21
	Number of agreements	62*	11
	Percent correspondence	64.46	52.38

*$p < .005$

TABLE 3

FINAL DECISIONS OF POLYGRAPH AND PSE CHART EVALUATIONS

Polygraph Decisions	PSE Decisions: Truth N	Truth %	Deception N	Deception %	Total N	Total %
Truth	22	64.71	1	5.88	23	67.65
Deception	6	17.65	3	8.82	9	26.47
Inconclusive	1	2.94	1	2.94	2	5.88
Total	29	85.29	5	14.71	34	100.00

decisions corresponded to each other. The data of all 34 cases are presented in table 3, which shows that, overall, polygraph and PSE decisions corresponded in 25 (73.53 percent) cases, and did not correspond in 7 (20.59 percent) cases. In two cases, polygraph decisions were inconclusive. If these inconclusive cases are excluded, correspondence will reach 78.13 percent.

It was assumed that correspondence may be reached for about 67 percent of the truthful suspects and for about 33 percent of the deceivers (Elaad and Shachar 1978). The differences between the expected and the obtained distributions of the suspects was analyzed by the normal approximation to the binomial distribution. As the analysis showed, the 78.13 percent correspondence between the two distributions is significant ($z2.48$; $p<.05$, against an apriori probability of 67 percent).

Considering the truth-tellers and the deceivers separately, it appears that among the nine deceivers detected by the polygraph only three (33.33 percent) were also found to be deceivers by the PSE (as expected by an apriori probability). On the other hand, among the 23 truth-tellers detected by the polygraph, 22 (95.65 percent) were also found truthful by the PSE.

DISCUSSION

The present study examined PSE validity under field conditions. Concurrent validity was estimated by simultaneously recording autonomic and vocal responses for comparison of the decisions based on the PSE chart evaluations with those based on the polygraph chart evaluations, which served as criteria (no information concerning ground truth was available). However, discussion of validity scores ought to be preceded by a consideration of the reliability of the measures used.

For the polygraph, interevaluation correspondence of blind chart evaluations ranged from about 78 percent for respiration to about 93 percent for the GSR. These figures resemble those reported elsewhere. Evaluating 102 charts of criminal cases, Raskin, Barland and Podlesny (1978) found about 84 percent correspondence between two evaluators when inconclusive cases were included, and 100 percent correspondence between the two when those cases were excluded. In another field experiment, Ginton *et al.* (1982) found 81 to 94 percent correspondence among eight evaluators.

PSE interevaluation correspondence of about 89 percent similarly resembles earlier findings obtained under field conditions for intraevaluator (Kradz 1974) and interevaluator (Heisse 1975) reliability. However, severe methodological criticism was leveled against these studies (see Horvath 1982; Nachshon and Feldman 1980). When compared with a better controlled field study (Nachshon and Feldman 1980), the reliability obtained in the present study seems considerably higher.

Considering the validity of PSE blind evaluations when they are compared with the blind decisions made for the various polygraph channels, correspondence seems to be rather low; only in one case (with GSR, deceivers) did correspondence exceed the 50 percent level. However, this figure very much resembles correspondence scores obtained for the various polygraph channels. Out of 18 interchannel comparisons, only in seven cases did correspondence exceed the 50 percent mark (in two cases, by about 1 percent only), and in only three of them was it significantly better than chance.

Apparently, decisions based upon blind evaluations of single polygraph channels are not valid, and consequently may not serve as criteria for PSE validity. That this is indeed the case is evident by the data shown in table 2, which reveals that except for one case (GSR, deceivers) correspondence between single channel blind evaluations and final polygraph decisions are quite low—below 63 percent.

Blind evaluations, it appears, are invalid not only for the polygraph, but for the PSE as well, since none of the polygraph and PSE blind evaluations corresponded significantly better than chance with PSE final decisions. Validity of PSE chart evaluations can therefore be determined only by comparing final polygraph- and PSE-based decisions.

Although the validity of the polygraph examination techniques which were used in the present study are disputable (see Lykken 1979, 1981; Podlesny and Raskin 1977; Raskin and Podlesny 1979), some studies show that when properly administered by trained polygraphers these techniques may be valid for the detection of deception under field conditions. In a recent controlled field study, Ginton *et al.* (1982) checked polygraph evaluations against ground truth information concerning the subjects' deception and truth-telling. Under these conditions, polygraph-based decisions of conclusive cases were correct in about 86 percent of the cases. Polygraph-based decisions could therefore be used as criteria for evaluating PSE validity.

In the present study correspondence between the polygraph-based and PSE-based conclusive decisions reached 78.12 percent (94.11 percent where no reservations were made). However, as a detailed data analysis further showed, correspondence was in fact very high (95.65 percent) only for the polygraph-based truth-tellers, but not for the deceivers (33.33 percent). This finding contradicts Nachshon and Feldman's (1980) data showing no significant correspondence between polygraph- and PSE-based decisions under field conditions. A closer look at the two studies shows, however, that in Nachshon and Feldman's (1980) study, PSE chart evaluations were conducted by three evaluators who were specially instructed for their task, but who by no means could be considered PSE experts. By contrast, in the present study the PSE evaluator was a licensed, experienced PSE expert. It is hence possible that the contradictory results obtained in both Nachshon and Feldman's 1980 and present study reflect the differential levels of expertise of the evaluators involved.

The plausibility of this explanation may be enhanced by comparing the correspondence in the two respective studies separately for truth-telling and deceiving subjects. In the present study, correspondence was very high for the truth-tellers, but very low for the deceivers. However, in Nachshon and Feldman's (1980) study two out of the three evaluators showed the reverse trend: about 75 percent correspondence for deceivers, and about 35 percent correspondence for truth-tellers (for the third evaluator, similar figures, between 50 percent to 60 percent correspondence, were obtained for the two groups of suspects). Presumably as a safeguard against false positive errors (decisions that a given person is a deceiver whereas in fact he is a truth-teller), when in doubt, the PSE expert, who participated in the present study, tended to decide that the suspect was a truth-teller. This acquired tendency clearly did not affect decisions made by the nonexpert PSE evaluators who had taken part in Nachshon and Feldman's (1980) study. Consequently, Nachshon and Feldman's suspects had an apriori 50 percent chance to be regarded as truth-tellers or as deceivers. By contrast, in the present study some suspects had a better chance to be considered truth-tellers than deceivers. Since most suspects agreeing to undergo criminal polygraph examinations eventually turn out to be truth-tellers (Elaad and Schachar 1978), the chances of correspondence between the decisions of the PSE expert and those of the polygraph experts were considerably higher than 50 percent for the truth-tellers and considerably lower than 50 percent for the deceivers. This differential probability is reflected in the results of the present study.

In conclusion, the present study showed that when conducted expertly and when evaluated globally, field PSE examinations may be valid for the detection of truth-tellers who undergo criminal interrogations. However, the fact that the PSE expert actually witnessed the polygraph examination makes it virtually impossible to single out the effect of the PSE alone on the decision made by the PSE evaluator. Therefore, more research is needed before a final conclusion concerning PSE validity, its scope of applicability, and the differential contributions of the factors affecting final decisions (PSE charts, examiner's expertise) can be reached.

ACKNOWLEDGEMENT

The authors acknowledge the help given by Deputy Commander Meir Kaplan, Head, Division of Criminal Identification, and Chief Superintendant Dr. Avital Ginton, Head, Section for Scientific Interrogation, Israel Police Headquarters, Jerusalem.

REFERENCES

Abrams, S. 1972. The polygraph: Laboratory vs. field research. *Polygraph* 1:145-150.

Amsel, T. 1979. Personal communication.

Barland, G. 1975. Detection of deception in criminal suspects: A field validation study. Unpublished Ph. D. dissertation, University of Utah, Salt Lake City, Utah.

Barland, G. H. 1978. Use of voice changes in the detection of deception. *Polygraph* 17:129-140.

Barland, G. H., and Raskin, D. C. 1973. Detection of deception. In *Electrodermal activity in psychological research*, edited by W. F. Prokasy and D. C. Raskin. New York: Academic Press.

Backster, C. 1969. *Technique fundamentals of the tri-zone polygraph test*. New York: Backster Research Foundation.

Brenner, M.; Branscomb, H. H.; and Schwartz, G. E. 1979. Psychological Stress Evaluator: Two tests of vocal measure. *Psychophysiology* 16:351-357.

Dektor Counter Intelligence and Security, Inc. 1974. Psychological Stress Evaluator. In Committee on Government Operations, House of Representatives, The use of polygraph and similar devices by federal agencies, pp. 238-241.

Elaad, E. 1980. Detection of stress in the human voice: A pilot study. Unpublished manuscript.

Elaad, E., and Schachar, E. 1978. Polygraph and field validity. *Crime & Soc. Deviance* 6:16-24 (Hebrew).

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. *J. Appl. Psych.* 67:131-137.

Heisse, Jr., J. W. 1975. The 14 question Modified Zone of Comparisons Test. *Steering Comments* 3(June):1-4.

Herbold, H. 1977. The influence of delayed response upon voice stress as measured by the PSE. *Polygraph* 6:125-131.

Holden, C. 1975. Lie detectors: PSE gains audience despite critics' doubts. *Science* 190:359-362.

Hollien, H. 1980. Vocal indicators of psychological stress. In *Forensic Psychology and Psychiatry*, edited by F. Wright, C. Bahn, and R. W. Rieber. (Annals of the New York Academy of Sciences, Vol. 347). New York: The New York Academy of Sciences.

Horvath, F. 1976. Detection of deception: A Review of field and laboratory procedures and research. *Polygraph* 5:107-145.

———. 1978. An experimental comparison of the Psychological Stress Evaluator and the Galvanic Skin Response in detection of deception. *J. Appl. Psych.* 63:338-344.

———. 1979. Effect of different motivational instructions on detection of deception with the Psychological Stress Evaluator and the Galvanic Skin response. *J. Appl. Psych.* 64:323-330.

———. 1982. Detecting deception: The promise and the reality of the voice stress analysis. *J. Forensic Sci.* 27:340-351.

Inbar, G., and Eden, G. 1976. Psychological Stress Evaluators: EMG correlation with voice tremor. *Biological Cybernetics* 24:165-167.

Kradz, M. 1974. The Psychological Stress Evaluator. In *U.S. Government, The use of polygraphs and similar devices by federal agencies*, pp. 244-254. Washington, D.C.: U.S. GPO.

Kubis, J. F. 1974. Comparison of voice analysis and polygraph as lie detection procedures. *Polygraph* 3:1-47.

Lippold, O. C. J. 1970. Physiological tremor. *Sci. Amer.* 224 (March):65-73.

Lykken, D. 1979. The detection of deception. *Psych. Bul.* 86:47-53.

Lykken, D. T. 1981. *A tremor in the blood*. New York: McGraw-Hill.

Lynch, B. E., and Henry, D. R. 1979. A validity study of the Psychological Stress Evaluator. *Canadian J. Beh. Sci.* 11:89-94

Nachshon, I., and Feldman, B. 1980. Vocal indices of psychological stress: A validation study of the Psychological Stress Evaluator. *J. Pol. Sci. & Adm.* 8:40-53.

Papcun, G. 1973. The effects of psychological stress on speech: Literature survey and background. Paper presented at the 86th meeting of the Acoustical Society of America, Los Angeles.

Podlesny, J. A. 1977. Physiological measures and the detection of deception. *Psych. Bul.* 84:782-799.

Raskin, D. C.; Barland, G. H.; and Podlesny, J. A. 1978. Validity and reliability of detection of deception. Project report, National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice.

Raskin, D. C., and Podlesny, J. A. 1979. Truth and deception: A reply to Lykken. *Psych. Bul.* 86:54-59.

Reid, J. E., and Inbau, F. E. 1966. *Truth and deception*. Baltimore: Williams & Wilkins.

Rockwell, D.; Hodgson, M.; and Cook, D. 1979. Psychological Stress Evaluator: An attempt at validation. Unpublished manuscript.

Suzuk, A.; Watanabe, S.; Takeno, Y.; Kosuji, T.; and Kasuya, T. 1979. Possibility of detecting deception by voice analysis. *Polygraph* 8:318-324.

Vandercar, D. H.; Greener, J.; Hibler, N. S.; Spielberger, C. D.; and Bloch, S. 1980. A description and analysis of the operation and validity of the Psychological Stress Evaluator. *J. Forensic Sci.* 25:174-188.