

SCIENTIFIC THEORY AND SCIENTIFIC EVIDENCE: AN
ANALYSIS OF LIE-DETECTION

JEROME H. SKOLNICK†

IN an age of technology, there is bound to be increasing interest in testing and laboratory methods, so-called "scientific evidence."¹ These techniques promise a number of advantages, especially for criminal investigations and trials, not the least of which is the elimination of human bias. A finger print identifies more objectively than an eye witness. Similarly, percentage of alcohol in the blood as an indicator of whether a man is "under the influence" is less susceptible to distortion than the judgment of a policeman observing behavior in the station house. Techniques such as these can certainly be useful. Simply because a method claims to be "scientific," however, it should not be accepted uncritically.

Laboratory techniques differ in significant ways. Identification of handwriting, for instance, is a more subjective operation—depending far more on the individual making the identification—than the matching of fingerprints. As a result, analysis of handwriting is probably a less reliable² method of identification than fingerprinting since the standards of the measuring instrument are subject to greater variation. The "validity" of a technique should also be questioned—whether it actually tests what it claims to test. For example, blood-alcohol measures seem reliable enough, but have been challenged on the ground that alcohol in the blood is significantly higher than alcohol in the brain at the same moment in time;³ alcohol in the blood, therefore, may not invariably be a "valid" criterion of the legal charge of "intoxication" or "under the influence."

Finally, however accurate a method may be for proving a disputed fact, it may not be allowable in view of constitutional requirements. In *Rochin v.*

†Assistant Professor, Law and Sociology, Yale University. Donald J. Horowitz, '58 L, helped gather materials at an early stage. I would like to acknowledge the aid of my colleagues, Richard C. Donnelly and Richard D. Schwartz, and especially that of Abraham S. Goldstein, whose willingness to lend a constructively critical ear was a frequent source of encouragement.

1. See, e.g., 1 WIGMORE, EVIDENCE §§ 163-65 (3d ed. 1940).

2. For a discussion of "reliability" and "validity" see GOODE & HATT, METHODS IN SOCIAL RESEARCH 235-39 (1952).

3. Rabinowitch, *Medicolegal Aspects of Chemical Tests of Alcoholic Intoxication*, 39 J. CRIM. L., C. & P.S. 225, 238 (1948). For an opposing view, see GRAY, ATTORNEYS' TEXTBOOK OF MEDICINE 625-26 (3d ed. 1949).

California,⁴ for example, while the method used (stomach pumping) was faultless as a means for proving that the individual possessed narcotics,⁵ it was held to violate due process of law.

In this Article "science" is not conceived of as techniques and laboratory methods alone. A general goal here is to indicate how abstract scientific knowledge, especially of methodology and concepts of probability, can aid the lawyer in evaluating "scientific evidence."

Only one "scientific" technique is examined in detail, the polygraph or lie-detector. In general, it is probably the most intriguing technique in the scientific evidence array. Although the lie detector is widely known, it is also widely misunderstood. Lie detection is an extremely complicated procedure, more complicated than its proponents acknowledge. As compared with fingerprinting, it requires more personal judgment, and is less straightforward in its scientific underpinnings than breath-alcohol tests.⁶ Its very complication makes the polygraph technique exceptionally interesting to analyze from a scientific point of view. Not only does it draw upon the conclusions and methods of several disciplines dealing with human physiology and behavior, but it also presents an important probability issue which has wider application in the criminal law.

Perhaps its most intriguing quality is to be found in the curious position it holds in the field of criminal procedure and evidence. While the polygraph appears to be in wide use,⁷ its results have been excluded from trials even when sought to be introduced by the accused.⁸ The exclusionary policy of the courts has, however, been attacked by some leading commentators on evidence, who favor introducing lie-detector results in civil and criminal trials, and who even suggest that there is something unscientific about a legal system which bars such evidence.⁹ Their eagerness to endorse the technique probably arises out of the fact that of all problems associated with human testimony—accuracy of perception, ability to recall—none can be considered more destructive to the just outcome of a trial than a lying witness.

4. 342 U.S. 165 (1952).

5. *Id.* at 165.

6. See discussion in text at notes 53-54 *infra*.

7. See, e.g., UNIVERSITY OF WISCONSIN BUREAU OF GOVERNMENT, SCIENTIFIC AND LABORATORY METHODS OF JUDICIAL PROOF II, 177 (1951).

The director of the Federal Bureau of Prisons recently announced the introduction of the technique into the federal prison system. See Bennett, *A Penal Administrator Views the Polygraph*, 24 FED. PROB. 40, 43-44 (1960).

8. For a recent discussion of admissibility of lie-detector evidence, see Note, 33 TUL. L. REV. 880 (1959).

9. See Streeter & Belli, *The "Fourth Degree": The Lie Detector*, 5 VAND. L. REV. 549 (1952); Wicker, *The Polygraphic Truth Test and The Law of Evidence*, 22 TENN. L. REV. 711 (1953); McCORMICK, EVIDENCE § 174 (1954) ("it is believed that the courts wholesale exclusion of lie-detector test-results, for want of scientific acceptance and proved reliability, is not supported by the facts").

Development of The Lie-Detector

The acknowledged fallibility of the jury system encouraged thought, especially in Berkeley and Chicago during the late 1920's and early 30's, on the possibility of perfecting a mechanical means of detecting guilt or innocence.¹⁰ When hearing testimony, the ordinary juror may be inordinately influenced by such matters as personal appearance, accent, gesture, and apparent force of conviction. Machines, on the other hand, are not swayed by subjective factors. A cardiograph, for example, selects only objective information, which, when interpreted by a trained physician, provides the basis for a scientific diagnosis of heart disease. Upon this model, a mechanical instrument was developed to diagnose credibility.¹¹

Lie detection through physical change is actually a throwback to early forms of trial by ordeal. There are reports of a deception test used by Indians based on the observation that fear may inhibit the secretion of saliva. To test credibility, an accused was given rice to chew. If he could spit it out he was considered innocent, but if it stuck to his gums he was judged guilty.¹²

Until 1895, however, nobody had ever used a measuring instrument to detect deception.¹³ In that year, the Italian criminologist Cesare Lombroso—famed mostly for a physicalistic criminal theory—utilized a combination of blood pressure and pulse readings to investigate crime.¹⁴ In 1915, further experiments were conducted with blood pressure readings by Marston.¹⁵ Around the same time, Benussi,¹⁶ and shortly afterwards Burt,¹⁷ began experimenting with respiratory recordings. John A. Larson, perhaps the most scholarly of the Chicago-Berkeley group which sought to advance the "science" of lie detection, built an instrument in 1921 which he called a "polygraph"; it combined all three measures—blood pressure, pulse, and respiration.¹⁸ His junior

10. LEE, *THE INSTRUMENTAL DETECTION OF DECEPTION* 14-17 (1953). See generally, LARSON, *LYING AND ITS DETECTION* (1932).

11. See INBAU & REID, *LIE DETECTION AND CRIMINAL INTERROGATION* 4 (3d ed. 1953) [hereinafter cited as INBAU & REID].

12. Lindsley, *The Psychology of Lie Detection*, in *PSYCHOLOGY FOR LAW ENFORCEMENT OFFICERS* 90 (Dudycha ed. 1955). See generally LEA, *SUPERSTITION AND FORCE* (4th ed. 1892).

13. INBAU & REID 2.

14. *Ibid.*

15. Marston, *Systolic Blood Pressure Symptoms of Deception*, 2 *J. EXPER. PSYCHOL.* 117 (1917).

16. Benussi, *Die Atmungssymptome der Lüge*, 31 *ARCHIV FÜR DIE GESAMTE PSYCHOL.* 244 (1914).

17. Burt, *A Pneumograph for Inspiration-Expiration Ratios*, 15 *PSYCHOL. BULL.* 325 (1918); Burt, *The Inspiration-Expiration Ratio During Truth and Falsehood*, 4 *J. EXPER. PSYCHOL.* 1 (1921).

18. INBAU & REID 3.

Only the quality machines, such as the Berkeley, Keeler, Reid, and Stoelting models measure several responses. "A great number of the 'lie-detectors' in use today are of \$24.95 type or a variation thereof. These devices usually record just one phenomenon—galvanic skin reflex (G.S.R.), e.g., the Fordham pathometer . . . the basic polygraph records changes

collaborator, Leonarde Keeler, added galvanic skin response to the list.¹⁹ Keeler's is the machine currently used by such a leading firm of examiners as John Reid and Associates. There has been one improvement on it—a muscular activity recorder to obviate distortions in blood pressure readings which could be brought about by unobserved muscular flexing.²⁰

Lie Detection Procedure

Ideally, the polygraph test is conducted by an experienced examiner in an environment free of such extraneous stimuli as witnesses, stenographers, reporters, or outside noises.²¹ A properly fitted examination room is similar to that used by a cardiologist or a clinical psychologist in his work.

After a preliminary interview the subject is seated in a chair specially constructed to permit the attachment of the various measuring devices: the pneumograph tube is tied to his chest, the blood-pressure cuff is wrapped round his upper arm, and a set of electrodes is attached to his hands. The subject looks straight ahead. The examiner is seated to his side behind a desk containing a set of controls which the subject cannot see.²² These instruments begin a continuous graphic recording when the examination commences.

The questions asked are based upon the results of the preliminary interview, together with available facts and circumstances forming the basis of the accusation. They also vary according to the type of person being questioned. Nevertheless, systematically designed "model" tests are presented by Inbau & Reid, authors of the leading work on lie detection, as a means for translating wiggles on the graph, measurements of physiological activity, into final judgments of credibility.²³ As we shall see, these models are not always followed.

In criminal investigations, three arrangements of questions are used by these and presumably other reputable examiners: the "control question" test, the "card test," and the "peak of tension" test. The "control question" test, the most important, is designed to deal with a problem familiar to every scientific experimenter: how to determine whether responses are the result of the experimental stimulus—*e.g.*, of what you have said to the subject, rather than the result of some extraneous factor, such as nervousness, an extreme propensity to perspire under interrogation, or an unintentional inflection in the

in a person's pulse, blood-pressure and breathing." Harman & Arther, *The Utilization of the Reid Polygraph by Attorneys and the Courts*, 2 CRIM. L. REV. 12, 12-13 (1955).

19. INBAU & REID 4.

20. *Ibid.*

21. See LEE, *THE INSTRUMENTAL DETECTION OF DECEPTION* 77 (1953) (instruction book by a captain in the Berkeley Police Department).

22. Photographs of the test are shown in INBAU & REID 7. For a description of the test from the standpoint of the subject, see 98 CONG. REC. 258-62 (1952). See generally MacDonald, *The Lie-Detector Era*, *The Reporter*, June 8, 1954, p. 10.

23. INBAU & REID 13-15, 16-26.

examiner's voice. By a series of matched questions²⁴ the "control question" test regulates the interpretation of the graph. Matched questions reveal the norm for the subject. Asked about a robbery which never occurred, an innocent person might produce a heightened response. But his response when questioned about the robbery under investigation should be no greater.

A similar rationale lies behind the "card test," usually given immediately after the "control question" test, presumably to see how the subject reacts when he lies in a situation other than the one under investigation. But it is also used to increase the subject's confidence in the machine through a dramatic demonstration of its power to root out untruth. More will be said later regarding the necessity for subject confidence in order to carry out the examination successfully.

The card test itself is simple in theory, readily performed, and exactly to the point. The subject is required to lie. Handed seven differently numbered cards by the examiner, he is told to select one and to remember its number. At fifteen-second intervals the examiner asks the subject if each successive card is the one whose number he is remembering. By the rules of the game the subject must answer "no" each time the question is asked, including the one time when the true answer is "yes." The examiner, who has been observing the subject's responses on the various graphs, tells the subject which card it was he had chosen to remember, informs him that he is responsive to the machine, and continues with the examination.

In theory the same as the card test, the "peak of tension" test transforms the card test from a parlor game into a life-like scrutiny. For it to be performed two conditions are necessary: first, the examiner must know of the existence of some object connected with the crime; second, no innocent subject must know of the existence of this object. Hence, the "peak" test may be useless where the details of a crime have been widely publicized.

Should circumstances be appropriate, however, the examiner will bring into play the one object known only to the culprit and to the authorities, in much the same manner as he earlier employed the selected card. For example, if the stolen object was a diamond necklace, the examiner will bring it into the examination room together with six similar objects, will exhibit each object to the subject at fifteen-second intervals, and ask, "Is this the object which was stolen?" An innocent suspect will answer "I don't know" seven times running, with no variation in graphic recording. As for the guilty suspect, his

24. An earlier form of the test interspersed key-questions (Did you rob the super-market?) with irrelevant or low anxiety provoking questions (How old are you?) The "control question" test asks presumably equal anxiety provoking questions, only one of which has to do with the crime. For instance, Did you rob the jewelry store last Saturday night? (the actual crime) as the "experimental" question; as the "control" or "guilt complex" question, Did you rob the tavern last month? (a crime never committed, or committed by someone else). Lie-detection proponents consider the change in design of questions markedly to have raised accuracy. See Reid, *A Revised Questioning Technique in Lie-Detection Tests*, 37 J. CRIM. L. & C. 542 (1947).

graph should produce an extraordinary wiggle when he is confronted with the stolen necklace. At that moment the jig, in detective novel parlance, is up.

THE SCIENTIFIC FOUNDATIONS OF LIE DETECTION

The Theory

The procedure of lie detection has been described as a proponent of lie detection might present it. Accuracy figures are available as further support for the process.²⁵ They suggest that a critic need not be persuaded solely by a description of the process; its results seem even more convincing. These figures, as compiled by Inbau & Reid, show an accuracy of around 95 per cent.²⁶ But they are unsatisfactory, for two reasons: First, only a proportion of their diagnoses have been checked.²⁷ Second, even if all cases had been "checked," the result would still be inconclusive. Validation is by confession or another inferential diagnostic process, the verdict at a jury trial. And in any given case these, and especially the verdict, may be as wrong as the lie-detector test.²⁸ There actually is no independent means of checking the phenomenon of lying, of confirming that an individual designated as a liar actually lied. In this respect, at least, the lie detector differs from the cardiograph. A diagnosis of heart disease may be checked by autopsy. There is no equivalent way of checking lying because it leaves no distinctive physical remains. Given these limitations, the best way to examine the claims of lie detection is to analyze it as a scientific theory. The theory will be stated and its assumptions examined to see how well they hold up under available evidence, and also how well they fit together.

The theory of lie detection can be summarized: the act of lying leads to conscious *conflict*; conflict induces *fear* or *anxiety*, which in turn results in

25. A discussion of accuracy is contained in Trovillo, *Scientific Proof of Credibility*, 22 TENN. L. REV. 743, 758-60 (1953).

26. INBAU & REID 110-11:

Since confirmatory or contradictory evidence is not always forthcoming after a deception diagnosis has been made in an actual case, exact figures are unavailable as to the accuracy of lie-detector test results. There is a sound basis, however, for making an estimate. The following estimate is based upon the experience of the examiners on the staff of John E. Reid and Associates during the last five year period. This estimate accords to the lie-detector technique, *when applied under the most favorable conditions*, an accuracy of 95 per cent, with a 4 per cent margin of indefinite determinations and a 1 per cent margin of possible error. In other words, in the examination of 100 subjects the examiner may make a definite and accurate diagnosis as to the guilt or innocence of 95 subjects. As to 4 of the subjects the examiner may be unable to arrive at a definite opinion as to guilt or innocence. With the 1 remaining subject the examiner may make an erroneous diagnosis of guilt or innocence.

27. Only 486 of the 1334 cases diagnosed as "guilty" were "verified." If as many as 133 of the remaining 848 were diagnosed wrongly, the error rises to .10. See Burack, *A Critical Analysis of the Theory, Method, and Limitations of the "Lie-Detector,"* 46 J. CRIM. L., C. & P.S. 414, 421 (1955).

28. See BORCHARD, *CONVICING THE INNOCENT* (1932); FRANK & FRANK, *NOT GUILTY* (1957).

clearly measurable physiological change. Lying, in short, produces intervening emotional states which reveal themselves in recordings of physiological activity. The theory contains two fundamental assumptions: first, a regular relationship between lying and certain emotional states; second, a regular relationship between these emotional states and changes in the body. Let us examine each of these assumptions in turn.

Lying and Emotion

Psychological literature suggests no regular relationship between lying and emotional states. This is hardly because psychologists regard the study of such a relationship as unimportant, but because they do not consider it consistently likely. The act of lying may evoke a variety of responses. In our society, people cannot go through life without some lying, and every individual builds up his own set of responses to the act. Lying can conceivably result in satisfaction, excitement, humor, boredom, sadness, hatred, as well as guilt, fear, or anxiety. Not uncommon are pathological individuals who, for various reasons, believe in their lies or are unconcerned about them.³⁰

Lie-detector proponents cannot have a precise concept of what emotions, if any, are produced when different people lie.³¹ The emotions may be fear, or anxiety,³² or one of these at one time and another a few moments later. With respect to the polygraph testing process, for example, a series of different emotions may be aroused at successive intervals.

Emotions and Bodily Response

The body responds involuntarily to emotions in a number of ways. Relatively benign responses include changes in any or all of the following: skin resistance (perspiration), respiration, blood pressure, heart rate, blood flow, skin temperature, muscle tension, pupillary diameter, gastric motility, and blood oxygen saturation.³³ Further, the discussion of upsetting material can, in some individuals, precipitate such painful and even dangerous somatic changes as headache,³⁴ backache,³⁵ episodes of Raynaud's disease,³⁶ and pro-

30. See Floch, *Limitations of the Lie Detector*, 40 J. CRIM. L., C. & P.S. 651, 653 (1950).

31. See Trovillo, *Scientific Proof of Credibility*, 22 TENN. L. REV. 743, 746-47 (1953).

32. Psychologists have been able to show that even such commonly related emotions as fear and anxiety are able to produce different behavioral responses. Sarnoff & Zimbardo, *Anxiety, Fear and Social Affiliation* (1960) (pre-publication report on file in Yale Law Library). It would seem, therefore, that different bodily responses might also be induced.

33. Lacey, *Psychophysiological Approaches to the Evaluation of Psychotherapeutic Process and Outcome*, in RESEARCH IN PSYCHOTHERAPY 160 (Rubinstein & Parloff ed. 1959).

34. Davis & Malmo, *Electro-myographic Recordings During Interview*, 107 AM. J. PSYCHIAT. 908 (1951).

35. Holmes & Wolff, *Life-Situation, Emotions and Backache*, 29 RES. PUBL. ASS. NERV. MENT. DIS. 750 (1950).

36. Mittelmann & Wolff, *Affective States and Skin Temperature: Experimental Study*

duction of blood, bile, and excessive hydrochloric acid in the stomach.³⁷

The recordings of blood pressure, pulse, respiration, and skin resistance produced by the polygraph may be assumed accurate if the mechanism has sufficiently rigorous specifications and is in proper order.^{37a} Such accuracy is, however, only the first measure of validity. If the polygraph is to serve its intended purpose the physiological changes must consistently coincide with subjective emotional states. This relationship, being problematic, should be examined in detail.

Three aspects of autonomic activity, or involuntary bodily reaction, may be described. These are: tension, lability, and nonspecific activity.³⁸ Tension is the amount of physiological activity occurring at a given time. Lability is a change in tension level from one time to another arising out of a change in perception or consciously held feelings. Nonspecific activity, sometimes called "spontaneous activity," has only recently been discovered to be important.³⁹ It refers to autonomic changes stimulated by sources which are neither perceptual nor conscious, although just what they are is not yet known. Bodily changes of this type can happen in the individual who is quietly resting, as well as in one who is highly aroused. Since nonspecific activity is always present, it complicates the measurement of lability.

Lie detection is based upon lability measurement. Does the subject reveal a significant heightening of blood pressure, a quickening of the pulse rate, shallowness of breath, and increased skin resistance when he is presented, for example, with the stolen object rather than an irrelevant one? Measurement of lability is complicated, however, by the fact that the amount of change—computed as a percentage or as a simple algebraic difference—is inversely related to the degree of tension existing at the moment the subject perceives the stimulus.⁴⁰ A subject with a high tension level will show changes smaller than a low tension subject; and further, a subject whose own tension levels are variable may record changes more as a result of his own tension variation than because of the effect of the stimulus.

The additional factor of nonspecific activity is a further complication. This type of activity seems to increase in relation to the subject's arousal condition, ranging from deep sleep to panic. By contrast, reactions to identifiable stimuli rise, up to a moderate state of arousal; they then drop while a subject goes

of Subjects with "Cold Hands" and Raynaud's Syndrome, 1 *PSYCHOSOM. MED.* 271, 286 (1939).

37. Mittlemann & Wolff, *Emotions and Gastrointestinal Function*, 4 *PSYCHOSOM. MED.* 5 (1942).

37a. In practice, there may be great variation in the quality of the machine used, see note 18 *supra*, and in the ability of an examiner to judge whether the machine is in proper order. See discussion in text at note 55 *infra*.

38. Lacey, *supra* note 33, at 179.

39. Lacey & Lacey, *The Relationship of Resting Autonomic Activity to Motor Impulsivity*, 36 *RES. PUBL. ASS. NERV. MENT. DIS.* 144-209 (1958).

40. See Lacey, *The Evaluation of Autonomic Responses: Toward A General Solution*, 67 *ANN. N.Y. ACAD. SCI.* 123 (1956).

from a state of moderate to a state of high arousal.⁴¹ Since lie-detector tests are apt to be given under fairly high arousal conditions, their interpretation seems especially prone to being complicated by nonspecific activity unrelated to the stimuli being presented.

Autonomic Intercorrelations

If all autonomic responses rose and fell exactly with emotional states, thereby enabling an experimenter to describe emotional states on the basis of an autonomic record, the different responses should have a precise relationship to each other. By means of an equation one should be able to describe the relationship between shallow breathing and rise in blood pressure. If this were true, a *polygraph* would be thoroughly unnecessary. A *uni*graph of any autonomic response would be sufficient to describe any other, as the heat of a chamber is sufficient to predict the pressure within. Thus, the very fact that a *polygraph* is used to detect lying indicates some irregularity.

Granted that the relationship among responses is not exact, it is apparently claimed that the several measures yield greater precision than any single one, and that "on the average" a fair degree of accuracy is obtained. This assertion would probably be true if there was a high degree of intercorrelation among the measures. A conclusion could be drawn with greater confidence if several measures having a slightly less than perfect correlation were to produce identical results. All would tend to reveal the same pattern and each would serve as a check upon the other. But in fact, psychophysicologists have been unable to find even a fairly regular relationship. A leading authority has stated,

One of the most crucial issues in psychophysiology concerns the surprisingly low intercorrelation among measures. In our work with noxious stimuli and a simple variety of autonomic variables, we have consistently found matrices of intercorrelations in which the majority of correlations approached zero.⁴²

Since the measure of any one autonomic response may not vary regularly with the reading from any other, the reliance placed by lie-detector practitioners on *several* autonomic measures seems unjustified. No one measure seems able to describe the emotional state of a subject better than any other. Hence, it is difficult to see how in a hard case the examiner can select an index to rely upon. For one subject a quickened pulse may most truly reflect his emotional state. For another, one whose pulse rate and emotional state are relatively unconnected, a soaring blood pressure level could be the autonomic activity which keeps pace with anxiety. For yet a third the blood pressure and pulse may be deceptive, while miniscule, but measurable, increases in skin resistance accurately reflect inner tension.⁴³

41. See Cohen, Silverman & Burch, *A Technique for the Assessment of Affect Change*, 124 J. NERV. MENT. DIS. 352 (1956).

42. Lacey, *supra* note 33, at 182.

43. Trovillo, *supra* note 17, at 747-48, considers this an advantage.

In sum, academic psychology and psychophysiology challenge both substantive assumptions underlying lie-detection theory: the assumption of a regular relationship between lying and emotional states, and the assumption of a regular and measurable relationship between emotional change and autonomic activity.

Methods and Techniques

Lie-detection proponents would probably not acknowledge this argument, coming as it does from academic psychology and physiology, as casting authoritative doubt on the validity of their results. They claim to be the specialists on lying—the ones upon whom judges should rely in determining whether the supporting theory has gained “general acceptance.”⁴⁴ They also imply that their collection of techniques and body of knowledge are superior to that of academic psychology.

Giving lie-detector proponents their due, it is worthwhile to examine the method more closely, to inquire into its specific techniques to determine whether academic psychology has not, in fact, overlooked advances made in the field of lie detection. Such an exercise, based on recognized standards of scientific method, should be relevant not only to an evaluation of lie detection, but also to an appraisal of other methods used to provide scientific and experimental evidence for the criminal process.

Interpretation and Reliability

Lie-detector proponents have never exaggerated their claims to coincide with the popular belief that a lie detector is a machine which rings a bell or a buzzer when a subject lies. Nor have they claimed that if any one examiner is given another examiner's test records he will arrive at the same conclusion about a subject.⁴⁵ Records alone are never enough, a fact that we might have

44. *Id.* at 762. For a discussion of the sociological meaning of factors characterizing a profession, see Goode, *Community Within a Community: The Professions*, 22 *AM. SOC. REV.* 194 (1957).

McCORMICK, *supra* note 9, at 363-64, argues that “general scientific acceptance” should not be made the standard of admissibility for scientific evidence (particularly lie detection tests); instead, he recommends that “any relevant conclusions which are supported by a qualified expert witness should be received unless there are other reasons for exclusion.” This leaves open, however, the question of whether the lie detector examiner is a “qualified expert witness.” This problem may have been one reason for forming the Academy For Scientific Interrogation, whose stated purpose is “to establish and maintain a standard of ethics and professional qualifications.” 5 *POLICE* 47 (1960).

45. INBAU & RED 116, say:

[S]ince we are dealing with nothing more than a *technique*, it is not ordinarily feasible for an experienced, qualified examiner to make a diagnosis from another examiner's test records without being at the scene of the examination itself. In making a diagnosis, an examiner must not only have before him records obtained during a carefully and properly conducted test; he must have a complete account of the subject's behavior indications and the general circumstances and conditions under

anticipated from the findings of psychophysicists about the difficulty of interpreting bodily responses. The individual judgment of the examiner, based on his own test records, is the ultimate determinant of credibility.

Lie-detector proponents do claim, however, the virtual infallibility of the lie-detecting process. As indicated, it is impossible to test the truth of this claim empirically, because there is no independent means of telling which subjects are actually lying; but within limits it is quite possible to test a part of it. *Reliability* can be tested by having several examiners test the same subject. Presumably, all should come to the same conclusion. The result which they all arrived at independently might not necessarily be correct, but similarity would establish uniformity of interpretation. On the other hand, if the conclusion varied from examiner to examiner, lie detection could be written off on the basis of empirical evidence, particularly if lie-detector examiners could be shown to be more erratic than a comparable group of lawyers or clinical psychologists judging the same subjects.⁴⁶

The fact that experiments of this kind have not been performed may be attributed in part to sheer inertia, in part to failure to comprehend the importance of such experimentation, and in part to a peculiar vocational necessity—the requirement of maintaining the myth that “lie detectors don’t lie.”

The Myth of Infallibility

This myth is essential to present methods of lie detection.⁴⁷ In a typical examination the subject is invited into a private waiting room by a receptionist who says the examiner will be ready in a few minutes. Actually, at this stage the test has already begun, for the receptionist is cast in a serious role. She offers the subject especially prepared reading matter which describes the lie detector as a virtually unerring instrument.⁴⁸ The initial hypothesis, guilty or not-guilty, is based upon the receptionist’s report of reactions to this literature. If the subject seems to be hostile, annoyed, or unsympathetic, guilt is

which the examination was conducted; and, most important of all, the diagnosing examiner must be in a position to direct some additional test variations.

46. The determination of paternity through blood tests suggests an additional laboratory for testing reliability. In some cases, the mother claims she had not had intercourse with any man other than the putative father, during the period in which conception occurred. Blood tests sometimes show the defendant could not have been the father. It would be interesting to see whether, in such cases, the findings showed the woman to be lying, and the man to be telling the truth. This would not, however, be an infallible test of reliability, for two reasons: one, the woman might have told “the truth,” as it revealed itself to her consciousness; she may have “forgotten” (unconsciously repressed) the fact she slept with other men. Two, blood tests are not always accurate. See Ross, *The Value of Blood Tests as Evidence in Paternity Cases*, 71 HARV. L. REV. 466 (1958).

47. See Reid & Arther, *Behavior Symptoms of Lie-Detector Subjects*, 44 J. CRIM. L., C. & P.S. 104 (1953).

48. Related to the writer during an interview with Richard O. Arther, a leading lie-detector theorist and examiner with John Reid and Associates, in New York City, April, 1958.

indicated. If the subject is able to show enthusiasm for the machine, he will have established himself, *prima facie*, as innocent. A comment by Reid and Arther⁴⁹ shows the importance of a subject's belief in the myth of infallibility as a test of guilt or innocence.

This belief that the innocent have in the accuracy of the lie-detector, and that they will be exonerated, is usually shown by their attitude. This attitude is one of genuine confidence in both the machine and the examiner. Because of this confidence they regard the examination as an experience they will want to relate to their family and friends.

A second purpose of the myth is to heighten the subject's bodily reactions. The routine of test administration suggested by Inbau and Reid, for example, calls for an early and emphatic communication that the machine doesn't lie.⁵⁰ In order to give physiological responses of measurable amplitude, the subject's "lies" must seem to him of some importance. If a parent told his four year old that Santa Claus would bring presents, the "lie" would hardly evoke an abnormal elevation of blood pressure or a quickened pulse, since it would have negligible emotional significance. Similarly, if a subject were tested by an instrument for which he had neither respect nor confidence the conscious lie would not necessarily be accompanied by a labile physiological response of the type needed to discriminate between the experimental and control questions. Were the machine regarded as capable of error, fear of detection would be reduced, and this lowering of fear would result in diminished physiological response. As in the trial by ordeal, the subject must believe in the efficacy of the diagnostic instrument in order for it to achieve maximum response.

The Examiner's Assessment of the Subject

Following the initial assessment by the receptionist, the examiner himself is the judge in all matters concerning the subject. He must not only be assured of the subject's belief in the machine, but must also determine whether the subject is physiologically normal, since such physiological abnormalities as cardiovascular conditions may produce aberrant recordings. He must likewise assess the subject's emotional tendencies. Some neurotics may feel guilt where no objective reason exists. By contrast, psychopathic personalities may lack feelings of guilt, even when they are "guilty" in fact. The examiner must also decide whether a subject, otherwise normal, is under extreme emotional strain or tension at the moment.⁵¹

The examiner must further decide whether a subject has rationalized his crime, thus bringing about control over his emotions. An embezzler who nurtured a grudge against his employer might feel that he has not "stolen"

49. Reid & Arther, *supra* note 47, at 106.

50. INBAU & REID 15.

51. Problems raised by physical and psychological pathology of subjects are discussed by several writers. See especially, Levitt, *Scientific Evaluation of the 'Lie-Detector,'* 40 IOWA L. REV. 440 (1955); Floch, *Limitations of the Lie Detector*, 40 J. CRIM. L., C. & P.S. 651 (1950).

anything and thus might not respond positively to evocative questions. In addition, a skillful subject, by concentrating hard on thoughts not connected with his crime, might be able to lower the height of his response. An examiner is expected to catch and correct for this defense.

The examiner is also expected to be sensitive to the behavior and attitude of the subject in the test situation.⁵² This information is used for two purposes. First, it aids the examiner in interpreting the record. Behavior symptoms and attitude, particularly belief and scepticism regarding the test, are respectively regarded as presumptive indications of "innocence" and "guilt." Secondly, tactics of the test are formulated on the basis of the examiner's appraisal of the subject. Tactics may be adjusted as the examiner sizes up the subject at different stages of the test. This is a crucial aspect of lie detecting since erroneous conclusions drawn at any point in the examination can wreck the accuracy of the result; successive questions depend to an apparently large degree on the outcome and conclusions drawn from prior questions.

Extent of Interpretation

In fitting the problem of interpretation into an evidentiary context it is instructive to compare the amount of interpretation required for lie-detector tests with the amount required in blood-alcohol tests. Measurement of the alcohol level in breath is analogous to a reading of blood-pressure or skin resistance in the sense that each is an indirect measure of what is truly sought—in the one test whether the subject is intoxicated; in the other whether he is lying. The intervening variable in the drunkenness test, per cent of alcohol in the breath, can be independently verified by drawing and analyzing a blood sample; it has been indisputably shown to bear a regular relationship to blood-alcohol content.⁵³ Furthermore, some measurable percentage of alcohol in the blood (usually .15) can be regarded as *prima facie* evidence of intoxication.⁵⁴ In lie detection, on the other hand, there is, of course, no comparable heart rate, extent of skin resistance, or blood-pressure level which may be regarded as *prima facie* proof of lying, since tension varies with the individual subject. Nor is lability—absolute or proportional—for any or all physiological measures, able to serve as *prima facie* evidence of lying.

Finally, the scientific certainty of relationships supporting breath-alcohol tests permits them to be performed properly by persons with minimal training; a patrolman may be trained for these purposes. Mistakes can be made, of course; but these deviations from a recognized standard routine can be spotted by an expert. By contrast, the polygraph is far less standardized, and its use requires a great deal of examiner interpretation.

52. INBAU & REID 106.

53. Harger, Forney and Barnes, *Estimation of the Level of Blood Alcohol from Analysis of Breath*, 36 J. LAB. CLIN. & MED. 306 (1950).

54. DONIGAN, *CHEMICAL TESTS AND THE LAW* 10 (1957).

Training for Interpretation

The large degree of interpretation required for lie detection means that multiple skills are necessary to conduct an examination. Apart from keen personal insight, the process demands familiarity with several medical specialties, plus an understanding of clinical and social psychology. At present, most lie-detector examiners have a professional police background, and much less formal scientific training than cardiologists and psychiatrists. The comparison is appropriate because of the similarly complex judgments the evaluator must make in each case. If anything, the judgment of the lie-detector examiner is the more difficult. The psychiatrist need diagnose only the general emotional condition of the patient; the lie-detector examiner must decide whether a man is lying about a particular event, not whether a man has a *tendency* to be a liar. Under such circumstances, the several months advocated by leading examiners as sufficient training is hardly responsible, and casts grave doubt on the whole business as a serious professional enterprise.⁵⁵

Test Design

Lie-detector proponents would probably object that this description exaggerates the amount of interpretation necessary to the process. They would

55. See INBAU & REID 115.

The Bulletin of the Keeler Polygraph Institute, a leading lie-detection school which trains law enforcement personnel from many parts of the country, breaks down its curriculum as follows: 1. MECHANICAL ASPECTS, 21 hours. Thorough study of Polygraph study of Polygraph mechanical theory and machine construction; characteristics of instruments in use; Polygraph interrogation room and facilities. 2. POLYGRAPH TECHNIQUE, 30 hours. Covers test types, question types, question formulation, demonstrations, use of interpreters, applied psychology, etc. 3. INTERROGATION, 10 hours. Review of interrogation principles, general, and with the polygraph. 4. CHART INTERPRETATION, 20 hours. A study of reaction types, patterns of known psychopaths and past cases; principles of chart marking; use of analysis sheets, etc. 5. MISCELLANEOUS SCHOOL PROCEDURES, 30 hours. Class periods for examinations, reviews, practice in question formulation for specified crimes, correlation, supervised evening group study, etc. 6. MISCELLANEOUS INSTRUCTIONS, 38 hours. Study of history of technique, clerical aspects, ethics, professional organizations, test forms, self-induced physical conditions, field work, etc. 7. PRACTICE AND CASE WORK, 44 hours. Supervised student Polygraph examinations in general practice and actual case work. 8. PSYCHOLOGY, 14 hours. Fundamentals of psychology as applied to the Polygraph, abnormal behavior, etc. 9. LEGAL ASPECTS, 6 hours. The legal status of polygraph examinations, some do's and don'ts to follow, etc. 10. INTRODUCTION TO MEDICAL ASPECTS, 4 hours. Field trip to the medical section of a large museum and introductory instruction. 11. MEDICAL ASPECTS, 27 hours. Application of blood measurements and basic physiology in Polygraph examinations; includes study of circulatory, respiratory, nervous and endocrine systems; anatomy, psychiatry, etc.

Compare this with the typical training of a psychiatrist. College, 4 years. Medical school, 4 years. Interne, 1 year. Residency, 2 years. And if a psychoanalyst, an additional 4-6 years.

See, generally, *Statement of Principles Regarding Polygraph ("Lie-Detector") Examinations*, 48 J. CRIM. L., C. & P.S. 568 (1958).

probably claim that the routines they have developed—the card test, the control test, and the peak-of-tension test—obviate the need for all but a minimal level of interpretation. Although this belief may be sincere, its proponents have probably been misled by the reportedly exceptional results of the card test.

This is the test in which the subject's reactions on the polygraph are used by the examiner as a basis for telling him which card in a group of seven he has previously selected. Given prior to the control test to inspire confidence in the lie detector, the card test is a visual demonstration that "the machine doesn't lie."⁵⁶ Indeed, the card test is apt to be administered to any sceptic—from the academic, legal, or business worlds—as dramatic proof of the accuracy of the process.⁵⁷ The result of this test is actually the basis for a subtle and probably unintended deception of the subject, whoever he is—whether an accused, a professor, a judge, or a businessman interested in using lie detection for screening his personnel. There seems to be a natural tendency to generalize from the results of the card test. Indeed, the card test has probably fooled even the most competent examiners into ascribing a higher degree of accuracy to the whole enterprise than it really deserves.

If lying, albeit of a special kind, can accurately be detected by the card test, this result is a strong offer of empirical evidence to show that physiological responses vary regularly enough with lying to be probative. But even if this is true, it is incorrect to assume that the results of the card test can be duplicated by the control test. The uniformity of conditions underlying the card test is never matched by the control test. All subjects are required to lie in the card test. They must answer "no" seven times in a row to the question, "Is this the card you picked?" Always, however, one of the cards was in fact a card they picked. Therefore, the examiner is always in the advantageous position of *knowing* the subject has lied on one of seven questions. By contrast, in the real life situation of the control test there may be truthful or "innocent" sub-

56. INBAU & REID 52 say:

With the "wise guy" type of subject it is also advisable to begin the examination with a card test. We refer to the subject who, upon entering the examination room, remarks: "So this is the lie-detector! But the courts don't accept its results". The examiner's reply should be "That's the latest instrument, and with the accuracy we're now getting its admissibility as evidence isn't far off." The card test should then be given and even though the record shows and the examiner knows which card was selected the subject should be told: "I can't pick out your card. You're pretty good, so I'll adjust the sensitivity of the instrument and then give you another test." After the second card test the examiner identifies the chosen card and where feasible points out the deception response on the card. This will usually change the subject's attitude, instill in him a respect for the test procedure, and thereby better condition him for the regular tests to follow.

57. *Statement of Principles Regarding Polygraph ("Lie Detector") Examinations*, 48 J. CRIM. L., C. & P.S. 568 (1958).

jects⁵⁸ tested along with the "guilty." Therefore, the examiner cannot be at all certain that every subject tested has lied.

Real life complicates still further. Complete truthfulness and outright lying are points at the ends of a continuum. The absolutely truthful subject, one who has no knowledge whatsoever of the crime and is also emotionally tranquil, would ideally present no variation in his responses to matching questions in the control test. Such a subject is hardly likely to be encountered in real life.

A person who is really being given a lie-detector test—for example, one knows that the machine is actually being used to ascertain whether or not he is guilty of a crime—is probably more nervous than a detached sceptic to whom the card test is being demonstrated. Assume, for example, an innocent suspect with a prior criminal record who is being questioned at police headquarters about the murder of John Jones. Frightened, he may respond with racing blood-pressure and contracted respiration. The theory holds that he should be equally frightened when questioned about a murder that never took place—when asked: "Did you shoot Sam Smart in San Diego on Saturday night two weeks ago?"⁵⁹

But the validity of this assumption is doubtful. The suspect may have known Jones, or have had some connection with him, or at least have known of his murder. On the other hand, he may have never heard of Smart. Or he may be perfectly at ease about Smart's murder, knowing that a score of witnesses will testify that on Saturday night two weeks ago he was tending bar at the El Charro Club.

The real life conditions prevailing when the control test is conducted raise complications never encountered in the card test. As indicated earlier, the control test cannot assume a rigid polar classification between subjects, liars and nonliars. Furthermore, nervousness may effect the outcome. It may also be affected by differing degrees of involvement in the crime being investigated, such as having been a witness to the act or having given encouragement to the perpetrator of the crime. Innumerable shadings of "guilt" and "innocence" are possible depending on both the actual participation of the subject and on his emotional constitution.⁶⁰

58. Inbau and Reid use the words "innocent" and "guilty" interchangeably with "truthful" and "lying." This could be a dangerous assumption. A subject could be lying for a variety of reasons. For example, a subject might be covering for someone else; or he might lie to prevent himself from being inculpated in another crime; or, to prevent an embarrassing but legal fact, such as illegitimate birth, from becoming known. Moreover, "guilt" is a legal status, not simply a matter of fact. The relationship between "guilt" and "lying" may at times be close, but the two cannot always be equated.

59. INBAU & REID 16-17.

60. The latter may, in addition, be traceable to a complex of historical factors and personal attributes: interrogation by police directly before the test; lifetime experience with police and "investigators"; socio-economic class and ethnic origin.

Numerous behaviors and attitudes have been shown to vary systematically with social-class position and ethnic affiliation. See BENDIX & LIPSET, *CLASS, STATUS AND POWER* 271-81, 284-370 (1953).

Another difference is found in the "design" of the card test, which is far more simple and straightforward than the "design" of the control test. In the card test only one variable changes, all others remaining constant. The form of the key question, "Did you pick the ten of spades?" is the same for all the cards—those the subject did not pick as well as the one he actually selected. The act, picking, is constant. The place is constant. The control test, on the other hand, can make no such claims. The form of the various questions must change to account for several variables: Was it *you* did *what* to *whom* at a specific time and place. For the control test to be as sharply discriminating as the card test, only one variable at a time should be changed in any questioning period. If more than one is changed, such experimental precision as is found in the card test is lost.⁶¹

Indeed, as a result of the inconclusiveness of the control test, as compared with the card test, several control tests might have to be administered, perhaps on different days, to enable the examiner to reach a conclusion. Although additional administrations may, in some instances, serve simply as a precautionary check on earlier findings, ordinarily several repetitions are required to reach a conclusion.⁶²

The test which in its design approximates the precision of the card test is the peak-of-tension test.⁶³ Its usual use is in theft cases. In much the same way that the experimental card is matched by control cards in the card test, a stolen object is shown to the subject along with matching objects. Nevertheless, there are still several important—and some insurmountable—differences between the real life circumstances of the peak-of-tension test and the ideal laboratory conditions which prevail in the card test. Several conditions must be present in order for the peak-of-tension test to be workable.

First, some recognizable article connected with the crime must be involved (like the ten of spades in the card test). The test would not work, therefore, in an embezzlement case where particular bills would in all likelihood be indistinguishable.

Second, assuming that the article is recognizable, it must not have gone unnoticed or have been overlooked by the guilty party.⁶⁴ If part of a cache of

61. Lie-detector proponents decry laboratory tests, because their results have shown an accuracy much lower than that presumably found in practice. Trovillo, *supra* note 31, at 747 says, for example, "Simulated emotion in psychology classes, on the lecture platform, in drama, and in experimental laboratories has done more to clutter up and confuse honest polygraphic reporting than all the quackery of 50 years!" But if "simulated emotion" is not reliable, why do lie-detector examiners use it in the card test to demonstrate the infallibility of the control test?

62. INBAU & REID 25.

63. *Id.* at 23.

64. A leading lie-detector examiner encountered a case in which a thief's wife had stolen money from the loot before it was counted. The thief, misinformed as to the true amount he had stolen, was exonerated by the peak of tension test finding, but was later convicted on the basis of independent evidence. That examiner no longer uses the "peak"

stolen jewels has been recovered by the police, that part must have been examined by the guilty party, to whom the test is being administered. Moreover, if the lie detector is to provide an accurate result the subject must be able to recall with accuracy articles that he saw if and when he examined the loot.

Third, only the guilty party and the examiner must know the identity of the stolen objects. If innocent suspects learned the details of a crime from newspaper or radio reports, or inadvertently become informed as a result of police questioning, the test would be unreliable.

Fourth, the information held by the examiner must be accurate. In a theft case, the examiner may have been misinformed by the victim, either by mistake or by deliberate intent, in order to maximize the amount to be recovered from an insurance company.

A much more serious problem of the peak test than the possibility of exonerating a guilty man is the ease with which an innocent man might be inculcated by its results. For example, an innocent man is suspected of the theft of a gold watch. A suspect who has no idea of what has been stolen is given a peak-of-tension test. He is presented first with a package of hairpins, and asked "Did you steal this?"; second, with a pair of shoelaces; third, with a stapling machine; fourth, with a gold watch; fifth, with a package of rubberbands; sixth, with a bottle of ink; seventh, with a playing card. The difference between the gold watch and the other objects is so great that the suspect is likely to respond to it, even though he is innocent of its theft. He has unintentionally been cued to respond by a foolhardy examiner. While tests as poorly conceived as this one probably do not occur in practice, there is a real possibility that a suspect will sense which among several objects that he is shown is the stolen item. Even when control objects are similar, subtle cues—raising or lowering of the voice, an unintended gesture, a change in expression, hesitation in posing a question—may coincide with the presentation of the experimental object to distort the test's result. The possibility of intentional distortion by a dishonest examiner aside, an examiner's unconscious bias may express itself either in slight modifications of behavior during testing, or in a tendency to interpret marginal or uncertain responses as proof of guilt.

The Psychology of The Examiner

Thus, in addition to difficulties of interpretation arising from the psychology and physiology of the subject—recognized, at least to some extent, by leading lie detector advocates—there is a further important source of potential error that is never taken into account at all: the psychology of the examiner. Contrast the view of the examiner held by lie-detection experts with the way

test, believing there are too many variables out of his ken. Interview with Richard O. Arther, New York, April, 1958.

psychology and other disciplines regard the tester or experimenter.⁶⁵ The proponents of lie detection seem to ignore the problem of interpretive bias. There is, it is true, much talk of "skill" and "training" as qualification for conducting tests.⁶⁶ It appears that once an examiner has been "trained," however, he is regarded as an infallible information processing machine, blind as Justice. Psychology, by contrast, views its man as a variable, wayward observer who must be protected from ever present tendencies to distort and overlook information. So axiomatic is this belief in the fallibility of human perception that an experiment which fails to take it into account is discounted on grounds of "contamination."⁶⁷ Thus, if an experimenter is trying to test the idea that babies who have spent much time in orphanages will be mentally and emotionally retarded as compared with children who have spent little or no time in orphanages, he will arrange matters so that the person who scores the tests will not know which children have had which experience.

Social-psychological studies—particularly those dealing with the effects of decisions on subsequent perceptions—demonstrate that the lie-detector technique contains elements of distortion which its proponents have apparently ignored.⁶⁸ One psychological experiment, for example, has indicated that a person tends to perceive another more favorably when the attempts of the perceiver to influence or control that other person have been successful.⁶⁹ On the basis of this study, one might predict that persons who responded positively to influence attempts—showed they believed what they were told about the lie detector—would be regarded as less culpable by the examiner regardless of whether or not they were lying while being tested. Thus, a potential distortion in the interpretation of polygraph results is introduced by the correlation examiners draw between a subject's innocence or guilt on the one hand, and his acceptance or rejection of the accuracy of the lie detector on the other.

Other studies relating to the formation of first impressions and their effect on later perceptions imply more serious criticisms of the lie-detection process.

65. Since the Second World War, a research design called the "double-blind" control experiment, has been established as the standard evaluative technique in research pharmacology. This research design not only takes into account the placebo effect, *i.e.*, the patient's psychologically based improvement; it also controls for the clinician's susceptibility to bias. Drug and placebo are packaged identically in containers labeled in a code whose key is unknown to the clinician. He is therefore, through this precaution, prevented from administering the drug in a different manner from the placebo, and also restricted from another source of bias; allowing his enthusiasm for his hypothesis to color his perception of the patient's response to the drug. Berton Rouché, *Annals of Medicine: Placebo*, *The New Yorker*, Oct. 15, 1960, p. 85, at 88.

66. See note 55 *supra*.

67. See 1 *HANDBOOK OF SOCIAL PSYCHOLOGY* 271-72 (2d ed. 1954).

68. See Bruner & Tagiuri, *The Perception of People* in 2 LINDZEY, *HANDBOOK OF SOCIAL PSYCHOLOGY* 634 (2d ed. 1954); TAGIURI & PETRULLO, *PERSON PERCEPTION AND INTERPERSONAL BEHAVIOR* (1958).

69. Thibaut & Riecken, *Authoritarianism, Status and the Communication of Aggression*, 8 *HUM. REL.* 95 (1955).

Indeed, such studies are important for understanding any system of interrogation, because first impressions seem to have a marked impact—called a “primacy effect” by psychologists—on subsequent evaluations.⁷⁰ An experiment carried out by a leading social psychologist demonstrates this effect.⁷¹ Students in three sections of a course at M.I.T. were randomly given two different descriptions of a young instructor who was to lecture the entire group, and then to be rated by the class. The lecturer was described to half the students as a 26 year old married veteran, considered by acquaintances to be rather *industrious, WARM, critical, practical, and determined*. The other half was given a description identical in all respects, except that the word *COLD* was substituted for *WARM*. Students rated the instructor after a twenty minute discussion period. The different descriptions gave rise to more participation by the *WARMS* and less by the *COLDS* in the discussion. After the performance most of the *WARMS* rated the instructor as good natured, informal, considerate, humane, sociable and humorous, while the *COLDS* formed a different impression. The hostility engendered by use of the word *COLD* carried through to the discussion period in which the *COLDS* remained relatively silent. As a result, they reinforced their initial hostility, in that way increasing the antagonism revealed in the final ratings.

Another source of error in the interpretation of polygraph results is the “halo” effect, so named by Thorndike in 1920.⁷² It refers to an individual’s tendency to make all his impressions fit together, so that a person rated high on several desirable traits would ordinarily be rated high on all. A teacher, for example, is likely to give a better grade to an examination paper turned in by a student with a “bright” reputation than to the same paper turned in by a “dull” student. Later investigators have found that the halo effect is more persistent when the traits to be judged do not manifest themselves in any particular or clear behavioral pattern, and when they have moral implications.⁷³ Traits with moral implications are certainly prominent in the lie-detector process. Further, examiners are trained to formulate questions with an implicit guilt or innocence hypothesis in mind.⁷⁴ There is, consequently, a likelihood that the examiner will bias his interpretation of an answer by his preceding hypothesis, with the chain of probable distortion running all the way back to the initial impression.

Obviously, the most direct means of eliminating interpretative pitfalls would be to base decisions upon data which required minimal interpretation. The

70. See Asch, *Forming Impressions of Personality*, 41 J. ABNORM. SOC. PSYCHOL. 258 (1946); Mensh & Wishner, *Asch on “Forming Impressions of Personality”: Further Evidence*, 16 J. PERS. 188-91 (1947).

71. Kelley, *The Warm-Cold Variable in First Impressions of Persons*, 18 J. PERS. 431 (1950).

72. Thorndike, *A Constant Error in Psychological Ratings*, 4 J. APPL. PSYCHOL. 25 (1920).

73. SYMONDS, *DIAGNOSING PERSONALITY AND CONDUCT* 113 (1931).

74. See text at notes 49, 52 *supra*.

autonomic response data yielded by the polygraph hardly achieve this degree of precision. An indirect and less satisfactory way of eliminating distortion is through high standards of training, continuing and able research, and introspective acuity. In general, the less perfect the diagnostic instrument, the greater are the professional demands to be placed upon the diagnosticians. Since the lie-detection profession falls far short of the standard required by the imperfection of its instrument, the process cannot be granted the accuracy claimed by such proponents as Inbau and Reid.

Does the technique serve any purpose at all? On the negative side, it may serve uses which are morally questionable, and which might raise legal doubts as well. The presence of a lie detector during an interrogation could induce unreliable confessions, not otherwise obtainable;⁷⁵ less dramatically, through the creation of an atmosphere of examiner omniscience key information may be extracted from the subject which he might have withheld had he not been strapped to a "scientific-looking" electronic apparatus. Unsavory though these uses may appear from the comfortable furnishings of the ivory tower, the lie-detector technique still cannot be regarded as a brutal, third-degree method of questioning a suspect. Indeed, part of the motivation among the Chicago-Berkeley group which developed the technique was the desire to create a reliable instrument in order to do away with physical coercions commonly associated with interrogation.⁷⁶ As a *scientific* instrument, however, all that can legitimately be claimed for the polygraph is that through physiological responses it may provide clues to veracity that are more detailed than those afforded by visual observation of the subject in an interview. Nevertheless, there is strong reason to doubt that these autonomic response data are any more precise in terms of permitting a systematic and reliable inference of lying.

John Larson's observation is easily as applicable to the polygraph today as it was in 1930, "[T]he technique must still remain a police tool—a very efficient police tool, to be sure, but one whose primary function is that of opening up leads to further investigation of information rather than that of being of itself *prima facie* evidence."⁷⁷

CONDITIONAL PROBABILITY AND LIE-DETECTOR ACCURACY

Even if the accuracy figures given by Inbau and Reid are valid in some situations, their validity at all times and in all places cannot be assumed.

75. LEE, *THE INSTRUMENTAL DETECTION OF DECEPTION* 160 (1953).

76. *Cf.* INBAU & REID 110.

But the lie detector itself is a commonly used method of inducing confessions. See LEE, *THE INSTRUMENTAL DETECTION OF DECEPTION* 160 (1953):

To obtain a confession where guilt is indicated is the purpose and ultimate goal of the deception [lie detector] test, though a collateral and equally important objective is elimination of the innocent. . . . The instrument and the test procedure have a very strong psychological effect upon a guilty subject in inducing him to confess.

77. LARSON, *LYING AND ITS DETECTION* 190 (1932). More recently, Larson has said, I originally hoped that instrumental lie detection would become a legitimate part of professional police science. It is little more than a racket. The lie detector, as used

Accuracy figures will vary not only with the jurisdiction, but also according to the purpose for which the polygraph is used: to ferret out security risks, to screen employees, to establish paternity, to protect insurance companies and bondsmen.⁷⁸ The proportion of liars among unmarried mothers might be much higher than among scientists working for the Atomic Energy Commission, and the probabilities of polygraph accuracy would vary accordingly. This "conditional probability" qualification applies across the board to all devices, mechanical or not (*e.g.*, the lie-detector, the cardiograph, or the Minnesota Multiphasic Personality Inventory), which try to diagnose or determine a condition or a state of being, whether syphilis, lying, heart disease, mental illness, or paternity.⁷⁹ Unless the diagnostic instrument is perfect, the probability of its accuracy in any single instance will depend upon the prevalence of the condition being diagnosed in the population to which the test is administered.⁸⁰

*Conditional Probability in Cancer Diagnosis*⁸¹

As an illustration of conditional probabilities imagine a simple diagnostic method for every form of cancer, a method with a high degree of reliability. Assume that "high reliability" means an "accuracy," or "unconditional probability," of 99 per cent. Thus, if the test were applied only to people with the disease, 99 per cent of the reactions would be positive and one per cent negative, while in application to people without the disease, 99 per cent of the reactions would be negative, and only one per cent positive. Assuming that five persons out of a thousand actually have cancer, what is the "conditional probability" of a positive reaction, the probability that a person showing a positive reaction actually has cancer? A table will help to illustrate.

in many places, is nothing more than a psychological third-degree aimed at extorting confessions as the old physical beatings were. At times I'm sorry I ever had any part in its development.

DEUTSCH, *THE TROUBLE WITH COPS* 150 (1955).

78. For uses of lie-detection, see Highleyman, *The Deceptive Certainty of the "Lie-Detector,"* 10 HASTINGS L.J. 47, 48 (1958); Note, 39 CALIF. L. REV. 439, 443 (1951); Wicker, *The Polygraphic Truth Test and the Law of Evidence,* 22 TENN. L. REV. 711, 713-14 (1953); 98 CONG. REC. 258 (1952); Arther & Reid, *Utilizing the Lie Detector Technique to Determine the Truth in Disputed Paternity Cases,* 45 J. CRIM. L., C. & P.S. 213 (1954).

79. See Ross, *The Value of Blood Tests as Evidence in Paternity Cases,* 71 HARV. L. REV. 466 (1958).

80. An excellent and comprehensive, but fairly technical, discussion of conditional probability is to be found in Meehl & Rosen, *Antecedent Probability and the Efficiency of Psychometric Signs, Patterns or Cutting Scores,* 52 PSYCHOL. BULL. 194 (1955). These authors use the term "diagnosis" to denote "the classification of any kind of pathology, behavior, or event being studied," or to denote "outcome" if a test is used for "prediction."

An introduction to conditional probability is presented in WALLIS & ROBERTS, *STATISTICS: A NEW APPROACH* 327-29 (1956).

81. See DUNN & GREENHOUSE, *CANCER DIAGNOSTIC TESTS: PRINCIPLES AND CRITERIA FOR DEVELOPMENT AND EVALUATION* 9-19 (1950).

TABLE 1: *Conditional Probabilities of Cancer Test Having Reliability of 0.99.**

	<i>Do Not Have Cancer</i>	<i>Have Cancer</i>	<i>Total</i>
<i>Diagnosed to Have Cancer</i>	(A) .01(995) = 10	(B) .99(5) = 5	(C) 15
<i>Diagnosed Not to Have Cancer</i>	(D) .99(995) = 985	(E) .01(5) = 0	(F) 985
<i>Total</i>	(G) 995	(H) 5	(I) 1000

Unconditional probability accuracy = .99

Conditional probability accuracy = .33

*Results in this and later tables are in rounded numbers.

The row consisting of cells G, H, and I reports the assumed situation accurately: out of the total 1,000, 995 do not have cancer, while five do have the disease. Being one per cent inaccurate, however, the test indicates that about 10 of the 995 noncancerous subjects have cancer (cell A). All five of those having cancer are registered as "positives" by the test. In total, the test diagnoses fifteen persons as having cancer when, in fact, only five have cancer. In conditional probability terms, then, the test is not 99 per cent accurate, as unconditional probabilities indicate, but 33 per cent accurate—only one out of every three persons shown by the diagnostic instrument to have cancer actually has the disease.

Moreover, given the same distribution of cancerous persons in the population being tested, a fairly small diminution of accuracy insofar as unconditional probabilities are concerned leads to a considerably greater drop in accuracy with respect to conditional probabilities.

TABLE 2: *Conditional Probabilities of Cancer Test Having Reliability of 0.95.*

	<i>Do Not Have Cancer</i>	<i>Have Cancer</i>	<i>Total</i>
<i>Diagnosed To Have Cancer</i>	(A) .05(995) = 50	(B) .95(5) = 5	(C) 55
<i>Diagnosed Not To Have Cancer</i>	(D) .95(995) = 945	(E) .05(5) = 0	(F) 945
<i>Total</i>	(G) 995	(H) 5	(I) 1000

Unconditional probability accuracy = .95

Conditional probability accuracy = .09

Thus, Table 2 reveals that a lowering in unconditional probability accuracy of 4 per cent, from 99 to 95 per cent, yields a drop in conditional probability accuracy from 33 per cent down to 9 per cent—55 subjects are shown to have cancer when in fact only five have. The mathematical basis for this dramatic result—an instrument billed as 95 per cent accurate is wrong in more than 90 per cent of its positive diagnoses—is really quite simple. The underlying “trick” is found in the low number of persons having cancer in proportion to number of errors the test will make.

Conditional Probability and Lie Detection

The lie-detector technique is clearly amenable to this type of analysis. The distribution of liars in the population being tested can be shown graphically to effect the result of the lie-detector technique at every assumed level of accuracy, except total inaccuracy (zero per cent) or complete accuracy (100 per cent).⁸² The illustration of conditional probabilities in cancer diagnosis is applicable to uses of the lie-detector in situations in which the distribution of liars in the population is of a similar order of magnitude (.005). For instance, lie detectors have been used to ferret out security risks applying for positions in the Defense Department.⁸³ Assuming that 25 out of every 1,000 persons

TABLE 3: *Conditional Probabilities of Lie-Detector Test In Hypothetical Security Risk Situation Where .025 of the Population are Risks*

	<i>Nonsecurity Risks</i>	<i>Security Risks</i>	<i>Total</i>
<i>Diagnosed to be Security Risks</i>	(A) .05(975) = 49	(B) .95(25) = 24	(C) 73
<i>Diagnosed Not to be Security Risks</i>	(D) .95(975) = 926	(E) .05(25) = 1	(F) 927
<i>Total</i>	(G) 975	(H) 25	(I) 1000
<i>Unconditional Probability Accuracy = .95</i>			
<i>Conditional Probability Accuracy = .33</i>			

82. Inbau and Reid claim an accuracy of 95%, and only 1% error, with 4% indefinite determinations. INBAU & REID 110-11. So long as there is some error, conditional probability is relevant, even though the percentage of error does not constitute the difference between claimed accuracy and complete accuracy. For simplicity's sake, the examples used in this paper assume the percentage of error to be the difference between unconditional accuracy and complete accuracy. Such a procedure in some of the examples makes the lie-detector appear more inaccurate than the Inbau & Reid figures suggest. However, on the basis of the preceding analysis of the lie-detector technique a one per cent error seems absurdly low. I believe even a 5% error estimate to be too low, considering difficulties with the technique discussed earlier.

83. See 98 CONG. REC. 258-62 (1952); MacDonald, *The Lie-Detector Era II*, *The Reporter*, June 22, 1954, p. 22.

applying for such positions was truly such a risk—this assumption works out more favorably to lie detection than five out of 1,000—and assuming an unconditional probability accuracy of 95 per cent, what would be the chance that any one individual tagged a liar by the machine had in fact lied?

The calculations in Table 3 yield a conditional probability of .33. For every true security risk selected by the machine, two persons are falsely designated as such. Note that conditional probability accuracy, at the assumed level of unconditional probability (.95), has increased over that found in the cancer diagnosis hypothetical case charted in Table 2. There the unconditional probability was also .95 and the conditional probability accuracy was .09. The reason for the higher conditional accuracy in Table 3 is that more true positives were postulated in Table 3—25—than in Table 2—5. An important axiom, demonstrated in Table 4, accounts for this phenomenon of the conditional accuracy approaching the unconditional when the number of true positives is raised: conditional and unconditional probabilities are equal when exactly one-half the population being tested possesses the characteristic which is being diagnosed.

TABLE 4: *Conditional Probabilities of Lie-Detector Test In Hypothetical Security Risk Situation Where .50 of the Population are Risks*

	<i>Nonsecurity Risks</i>	<i>Security Risks</i>	<i>Total</i>
<i>Diagnosed To Be Security Risks</i>	.05(500) = 25	.95(500) = 475	500
<i>Diagnosed Not To Be Security Risks</i>	.95(500) = 475	.05(500) = 25	500
<i>Total</i>	500	500	1000

Unconditional Probability Accuracy = .95
Conditional Probability Accuracy = .95

Furthermore, the conditional probability accuracy of a diagnostic instrument in selecting positives is raised when positives constitute more than half the population. In Table 3, for instance, all 927 designated as nonsecurity risks, save one, are truly not security risks. In this table, conditional probability accuracy is consequently raised from .95 to better than .99 for the selection of those who are not security risks. But, again, the machine is only 33 per cent accurate when it comes to picking those it regards as security risks.

Applying Conditional Probability

Having illustrated some probability issues inherent in calculating the accuracy of a diagnostic test, there remains the question of how to use the infor-

mation both for the lie detector and for other diagnostic instruments as well. The answer depends upon the objective to which the information is applied and, hence, varies from situation to situation.

From a policy perspective one can conceive of situations in which false designations might be permitted. For example, under the conditions specified in Table 3, two out of every three persons designated a security risk is actually not a security risk. Yet, it is possible that the danger of permitting a liar to work on a particular assignment—where failure or espionage might cost the lives of multitudes—would be so profound that the unfounded stigmatization of a few would be a necessary sacrifice. By the same token, the suffering inflicted on those designated “positive” might be so slight as not to cause concern for those who are so designated. Since polio can be prevented by an injection of vaccine, everybody in the population is automatically considered a potential victim, *i.e.*, a true positive, and no prior attempt is made to test for immunity. If cancer could be cured by a series of painful, but permanently harmless, injections, great reliance would be placed on a test which correctly diagnosed all true positives along with many false. Unfortunately, the radical forms of treatment frequently required for cancer give pause to the use of a diagnostic test which has as high an unconditional accuracy as .95.

Hence, the use of a device which designates false positives requires that two considerations be weighed: (a) the necessity that all *true* positives be found and “treated,” and (b) the severity of the “treatment.”⁸⁴

*A Conditional Probability Model of Legal Diagnoses*⁸⁵

The concept of conditional probability touches the core of any judicial system, especially a system of criminal justice. A central task of a judicial system is to decide who is “right” and who is “wrong” under the law. Unless such a system is perfect—and in criminal cases we know that it is not—the issue of false positives is important.

From the viewpoint of conditional probability, one envisions criminal adjudication as a total system, rather than as a two unit, state to individual, interaction. The state is a sorting machine, trying to identify those persons in the entire population who are responsible for acts which the state has designated criminal. Not a single individual, but a population mass, is continually being processed. In this view, every individual in the population is tested to see whether it was he who committed a particular act. An hypothesis is formed by the state regarding each individual, and for most individuals the hypothesis of guilt is rejected. For some, it is maintained at every level of the criminal process—arrest, preliminary hearing, grand jury, petit jury. Each level of ad-

84. The idea of “false-positives” is noted, but not systematically developed by Levitt, *Scientific Evaluation of the “Lie Detector,”* 40 IOWA L. REV. 440, 446 (1955).

85. Statements in this section pertaining to criminal procedure are based upon sources cited by A. Goldstein, *The State and the Accused: Balance of Advantage in Criminal Procedure*, 69 YALE L.J. 1149 (1960).

judication serves to filter out the innocent. Every previous level affects the accuracy of the subsequent level.

We know independently, however, that the petit jury, the final arbiter of the process, sometimes makes mistakes. But if the system is viewed in this way—abstractly, like a mathematical model⁸⁶—conditional probability is able to show that the actual number of mistakes is to some exact extent (assuming knowledge of accuracy at each stage) related to the accuracy of the previous stages.

Of course, in reality, it is impossible to determine what percentage of those coming before petit juries are in fact innocent, and what the accuracy of the petit jury would be if the distribution of guilt and innocence was 50-50. (If all who came before the jury were truly guilty, there could be no false positives.) Still, an analysis of this kind is instructive, even using hypothetical figures.

Imagine a situation similar to that in Table 2: A .95 unconditional probability accuracy for the diagnostic instrument, the jury; the population being diagnosed contains five guilty people out of every thousand. The result is mathematically the same. Ten people are falsely found guilty to every truly culpable person declared guilty.

Such a situation might appear extremely unlikely to most lawyers. They would probably argue that, although exact figures are unavailable, most of those indicted are guilty.⁸⁷ This may well be true; but an overwhelming majority of those who are indicted actually plead guilty,⁸⁸ frequently motivated by the desire for the lowered sentence that may accompany a guilty plea. Yet there are those individuals who refuse to plead guilty, who insist upon their innocence in the face of the expense of a trial, plus the likelihood of a more severe sentence. Many in this group are found guilty. Who can say what the percentage of actual guilt is among them?

If less than fifty per cent of those who protest their innocence to the end are "guilty" in fact, then the conditional probability accuracy is lower than the unconditional. If, for example, the percentage of those actually "guilty" was forty per cent, clearly a possibility, conditional probability would be lowered to .92 from an unconditional probability of .95. If the guilty proportion among those insisting on their innocence was only ten per cent, the conditional accuracy would fall to .66, still with an unconditional accuracy of .95. When calculated with an unconditional accuracy of .85, the forty per cent example and ten per cent example yield conditional probabilities of .81 and .39, respectively.

Several conclusions emerge. Both (a) unconditional accuracy must be high, and (b) the distribution in the population of the characteristic to be tested must be above fifty per cent, for there ultimately to be high accuracy. Con-

86. For a discussion of uses of mathematical models see Lazarsfeld, *Evidence and Inference in Social Research*, Daedalus, Fall 1958, pp. 99, at 124-29.

87. Goldstein, *supra* note 85, at 1162.

88. *Id.* at 1189.

centration solely upon achieving unconditional probability accuracy—for example, on how to make the jury correct more often—is not very helpful if less than half those coming to trial are guilty. This generalization becomes increasingly important the lower the percentage of those truly guilty who come to trial.

Proof at Pretrial Stage

Although this analysis would indicate that the filtering out process should be made as selective as possible, there are certain dangers in proposals to require higher standards of proof at the pretrial stages of arrest, preliminary hearing, and grand jury. As Professor Abraham S. Goldstein has cautioned, "Though a tightening of the screens might well decrease the danger that innocent men would be convicted, it might also distort other functions, such as investigation, which are perhaps equally important."⁸⁹ In short, the price of greater protection for the innocent may be a lessening in the number of guilty persons who are detected and convicted, a price that at some point may be considered too great.

But not all modifications of the screening process designed to protect the innocent will bring about fewer convictions of the guilty. If improvement takes the form of discarding inherently unreliable investigatory techniques, conviction of the guilty may increase, since an unreliable technique may give rise to two kinds of mistake—exonerating the guilty as well as condemning the innocent.

The analysis of lie-detection techniques and theories would seem to indicate that the lie detector is one such unreliable device. The uncertainty of the correlation between physiological responses and lying and the numerous possibilities for erroneous interpretation raise serious questions about the reliability of both positive and negative conclusions of guilt. For this reason, rejection of the polygraph test as a tool of police investigation would not necessarily reduce the efficacy of police investigative procedures. It might, instead, channel such procedures into the use of methods requiring less interpretation, thereby increasing the selectivity of the pretrial process with regard to both the guilty and the innocent.⁹⁰

SOME ADDITIONAL CONCLUSIONS

Police Use of Lie Detectors

The criticism of lie-detector accuracy does not necessarily demand that such tests be discarded altogether. Police investigation frequently makes use of other

89. *Id.* at 1172.

90. But the ramifications of any particular change in procedural requirements are so complex, it is difficult to predict the effects of tightening screens without a rigorously conceived series of observations to be performed after the fact. For an excellent example of what can be done to evaluate a suggested change in legal procedure, see ZEISEL, KALVEN & BUCHHOLZ, *DELAY IN THE COURT* (1959).

highly interpretative devices, such as ordinary interrogation. Harm is minimized because the interrogator does not usually regard his own conclusions about the suspect as dispositive. If the polygraph technique was generally regarded with similar skepticism, the device might do little harm in terms of inaccurately categorizing a suspect, and it might conceivably be used simply as an adjunct to ordinary procedures.

It may turn out, however, that the label "lie detector" attached to the polygraph, has already fixed the "image" of the device with a presumed certainty long beyond the point of no return in the general population as well as in police circles. Lie-detector proponents may be finding themselves on the horns of an advertising dilemma. If they revise their estimates concerning the certainty of the polygraph with the vigor necessary to destroy the myth of "lie detection," they may also destroy the desire of police to make limited use of the technique.

The accuracy of the lie-detector technique, however, is only one of the values relevant to its use by police. The question of whether the polygraph should be used at the early stages of the criminal process is extremely complex. One may take the position, on moral grounds, that the test should not be used at all because of the posture it permits the state to take in confronting the accused. Its chief function appears to be to induce confessions by deception, convincing the suspect that "the machine doesn't lie." As such, it is probably used against persons least capable of judging its scientific authenticity.⁹¹ The "tricks" used by police are probably most effective against inexperienced persons who "waive" their rights unthinkingly.

Criminal law takes a curious attitude toward waiver of rights. It seems to assume that the innocent will "assert their rights," that is, will not testify, will demand counsel, will refuse to cooperate with police. Fact may be quite the opposite. The innocent and unsophisticated suspect may be precisely the one who will not demand procedural protections. If questioned, he will talk freely, submit to tests, such as the lie-detector, and may feel disinclined to bring an attorney into the picture. Therefore, an argument against strengthening procedural protections on grounds of protecting society's interest in apprehending the guilty, may be misleading. There may be enough procedural protections for the hardened criminal—since he is sufficiently knowledgeable to demand his rights—and too few for the inexperienced suspect.

Use of the lie detector might also lead to a lower standard of arrest if police were to rely upon it as a "screening device."⁹² Suspects would be brought into

91. See "Letter From an English Policeman On Use of Judges' Rules," in FRYER, *SELECTED WRITINGS ON THE LAW OF EVIDENCE AND TRIAL* 845 (1957). The lie detector is probably often used against such persons, since its chief function appears to be to induce confessions, Lee, *supra* note 76.

See "Letter From an English Policeman On Use of Judges' Rules," in FRYER, *SELECTED WRITINGS ON THE LAW OF EVIDENCE AND TRIAL* 845 (1957).

92. See McInerney, *Routine Screening of Criminal Suspects by the Polygraph (Lie-Detector) Technique*, 45 J. CRIM. L., C. & P.S. 736 (1955).

the station house less discriminately, with a feeling of assurance by police that the polygraph will separate the innocent from the guilty. Since the number of innocent persons tested would be increased in such a situation, the probability of selecting false positives would be raised.

Finally, with increasing acceptance at the police level, there is a greater likelihood that pressures will arise from those involved directly or indirectly with its use, from defense attorneys as well as prosecutors, to accept lie-detector evidence as court-appointed testimony at the trial.

Whatever the drawbacks to lie detection, its proponents might nevertheless argue however, that lie detection is preferable to some investigatory methods presently employed by police. Even if lie detection is based on deception of the suspect, deception appears to be a commonly accepted police technique. Textbooks on standard methods of criminal interrogation advise examining officers to deceive suspects in order to induce confessions.⁹³ On the theory that a little deception is better than brutality, lie detection has been supported as an alternative to "third degree" methods.⁹⁴ Instances of police coercion are widely reported,⁹⁵ although there is presently no systematic knowledge of how

93. See, e.g., KIDD, POLICE INTERROGATION 124-25, 133-86 (1940). Deceptive interrogation methods are fully explained and encouraged in O'HARA, FUNDAMENTALS OF CRIMINAL INVESTIGATION 95-114 (1956).

Some notable examples of deception suggested by him are: *Pretense of Physical Evidence*. "The interrogator . . . pretends that certain physical evidence, appropriate to the case, has been found by laboratory experts. The average person has mystical notions of the power of scientific crime detection and will accept practically any claims that science may make. Thus, the detective can mix pseudoscience in his statements. . . . In a homicide, the interrogator can refer to hair found at the scene of the crime, which can be shown, under the microscope to be the suspect's hair. For added realism, the suspect can be invited to look into the microscope."

Knowledge Bluff. "The interrogator reveals a number of pertinent items of evidence which are definitely known. He is thus able to convince the subject that it is futile to resist since the interrogator obviously has sources of knowledge. The interrogator should prepare himself for this approach by learning a great number of facts about the crime in question and about the subject's background. He must create the impression that he possesses an unlimited store of knowledge. This is not too difficult if the subject is confused and is normally not too bright."

Reverse Line-Up. "This technique is applicable in crimes which ordinarily run in series, such as forgeries and muggings. The accused is placed in a line-up, but this time he is identified by several fictitious witnesses or victims who associated him with different offenses. It is expected that the subject will become desperate and confess to the offense under investigation in order to escape from the false accusations." (Italicized titles are O'Hara's).

Not all textbooks on criminal investigation endorse deceptive techniques, however. See MULBAR, INTERROGATION 21-22 (1951) ("Never lie nor deceive the subject. It is dangerous because once he catches you, he will never again cooperate. Never make a promise that cannot be fulfilled. You will succeed only in losing his respect for you along with your own self-respect."). See also his chapter in SNYDER, HOMICIDE INVESTIGATION 77 (1950).

94. INBAU & REID 110; see also MacDonald, *supra* note 83, at 29.

95. For illustrations of police coercion, see HOPKINS, OUR LAWLESS POLICE (1931); Westley, *Violence and the Police*, 59 AM. J. OF SOCIOLOGY 34 (1953).

common such practices are.⁹⁶ Systematic observations might in fact disclose an inverse relationship between the use of lie detection by police organizations and reliance upon "third degree" methods. Increased police interest in so-called "scientific" methods might be followed by an increase in self-respect, and accompanied by a heightened feeling of "professional" and social status. Elevating the occupational status of police work may result in an increased reluctance to resort to violence, since violence is typically identified as a characteristic of lower social classes.⁹⁷ Thus, however poorly it performs its manifest function of separating liars from truth-tellers, the lie detector conceivably could serve desirable hidden purposes at the station house level.⁹⁸

Whatever the relative merits of these arguments, the fact is that lie-detection techniques are widely used.⁹⁹ Indeed, the leading police periodical regularly maintains a column devoted to polygraph testing.¹⁰⁰ In view of this apparently widespread acceptance and the establishment of a professional organization designed, at least in part, to overcome judicial resistance,¹⁰¹ the issue of whether lie-detection test results should be introduced at trial becomes increasingly important.

At the Trial

One school of thought, whose most articulate spokesman in the United States is Helen Silving, would do away with polygraph evidence at trial, irrespective of its probative value, on both due process and moral grounds.¹⁰² Professor Silving would not even permit the lie detector to be used by the accused in his own defense.¹⁰³

Professor Silving's argument, however, is based upon an erroneous assumption, that lie detection "tests the unconscious." In fact, lie detection does not bring out *repressed* materials in the individual's life history. Miss Silving

96. The behavior of police has drawn the attention of several legal scholars, but none has yet conducted an empirical study of either police arrest or interrogation practices. Commentaries which do exist have necessarily been based upon fragmentary and secondary materials such as police manuals, yearbooks, textbooks, and isolated reports.

See J. Goldstein, *Police Discretion Not to Invoke the Criminal Process*, 69 YALE L.J. 543, 554-62 (1960); Foote, *Law and Police Practice: Safeguards in the Law of Arrest*, 52 NW. U.L. REV. 16, 29 (1957).

97. Schneider & Lysgaard, *The Deferred Gratification Pattern: A Preliminary Study*, 18 AM. SOC. REV. 142, 143, 145 (1953).

98. See generally Merton, *Manifest and Latent Functions*, in SOCIAL THEORY AND SOCIAL STRUCTURE 19 (1957).

99. See DEUTSCH, *THE TROUBLE WITH COPS* 150-51 (1955). FBI Director J. Edgar Hoover, however, has expressed strong dislike of lie detection and skepticism about its value. *Id.* at 151.

100. See, e.g., 5 POLICE 47 (1960).

101. See note 44 *supra*.

102. Silving, *Testing of the Unconscious in Criminal Cases*, 69 HARV. L. REV. 683, 687, 688-89, 702 (1956).

103. *Id.* at 693.

wrongly joins the polygraph with so-called "truth serums" under the heading "objective tests."¹⁰⁴ What the polygraph actually tests is "conscious" conflict between the answer given to the interrogator and the facts as believed by the accused.¹⁰⁵ In theory, the polygraph, by measuring autonomic responses, discloses whether the subject believes what he is saying. He is at all times conscious of the content of his speech. If he lies, the conscious conflict produces an involuntary physiological response, significantly different from that which would have been evoked had he not lied. Presumably, the physiological responses recorded by the polygraph arise out of emotions felt by the individual while lying. It is important to emphasize that such feelings are always *consciously* felt. Freud makes this emphatically clear: "It is surely of the essence of an emotion that we should feel it, *i.e.*, that it should enter consciousness. So for emotions, feelings, and affects to be unconscious would be quite out of the question."¹⁰⁶ There is, thus, no interference by lie detection with "freedom of the will" in the sense of unconscious probing. Of course, the subject's answers may be unconsciously motivated, but this is true of any testimony. In that sense, there is never any "freedom of mind and will."¹⁰⁷

Although some authorities claim otherwise, it appears that lie detection is a form of "testimony," and that the privilege against self-incrimination therefore applies to the technique.¹⁰⁸ But our law surely *permits* an accused to testify or to introduce "bodily evidence" on his own behalf if he wishes; indeed, it sometimes allows the state to introduce evidence from the *body* of the accused, regardless of his wishes.¹⁰⁹

In deciding whether to admit such evidence three standards are apparently applied by the courts: consent, dignity, and probative value. Practically, the

104. Silving, *supra* note 102, at 683.

For a discussion of testing with drugs, see Dession, Freedman, Donnelly & Redlich, *Drug-Induced Revelation and Criminal Investigation*, 62 *YALE L.J.* 315 (1953).

105. As discussed earlier, see text accompanying note 66 *supra*, this is a fundamental limitation of lie detection which prevents it from ever achieving general applicability. Assuming it had no other difficulties associated with it, the polygraph is applicable only in instances where there is correspondence between the event as it happened, and beliefs about it held consciously by the subject of the test.

106. Freud, *On the Unconscious*, in 4 *COLLECTED PAPERS* 109-10 (American ed. 1959).

107. It is also true that the polygraph will record physiological activity arising out of unconscious processes. To this extent, it does "test the unconscious." But the polygraph test does not look for this information. Indeed, to the extent it occurs, it tends only to confuse the interpretation of consciously motivated responses.

108. But see McCORMICK, *EVIDENCE* 266 (1954); INBAU, *SELF INCRIMINATION* 67 (1950). These authorities assert that lie detection results are not "testimony" in the conventional sense. They assert that the physiological response, not the verbalized answer to the question put by the examiner is the crucial part of the test. But as a matter of fact, the physiological response is directly associated with the testimony. Without testimony, there would be no differential physiological response to observe. This is not true of other "uses of the body." The percentage of alcohol in the blood, for instance, does not vary with the testimony of the accused. With the lie-detector, however, physiological responses are a function of testimony.

109. See McCORMICK, *EVIDENCE* 264 & n.13 (1954) (collecting cases); A. Goldstein, *supra* note 85, at 1189 & n.134 (collecting cases).

consent standard would present no barrier, since polygraph tests cannot be administered unless the suspect is willing to cooperate.¹¹⁰ Furthermore, it is difficult to conclude that the measurement side of lie detection—the reading of blood pressure, pulse, and respiration—is less “dignified” than other methods of obtaining bodily evidence authorized by the Supreme Court.¹¹¹ The issue of “dignity” is raised only if the technique were to have so little accuracy that the subject is deceived when the examiner tells him, “The machine doesn’t lie.” Given this background, the question shifts to the third, pragmatic, standard. If the physiological measurements associated with an individual’s speech, *willingly given*, were regularly as probative of lying as the amount of alcohol in the blood is of intoxication, or as pumping the stomach can be of possession of narcotics, there would be no reason why an accused should not be permitted to introduce the results of a lie-detector test into evidence if he chose to do so.

Unfortunately, no such test exists. What exists instead is, at best, the opinion of a skilled interrogator, guided to some extent by the systematic measurement of physiological responses. Moreover, because of basic theoretical limitations, there is little reason for supposing that a test with very high unconditional accuracy will ever be developed.

If, despite the objections of dignity and accuracy, polygraph evidence were ever to be admitted into the courtroom, it should never be introduced as court-appointed testimony. The technique does not warrant even the degree of confidence accorded to the interpretive testimony of a medical doctor or psychiatrist. The introduction of evidence by opposing lie-detector experts would, of course, be as highly confusing and time wasting as any other battle of experts. Proponents of lie detection take a dim view of the prospect of conflicting lie-detection testimony.¹¹² It would threaten the status of their so-called “profession”; it would also destroy the myth of infallibility upon which they depend in part for their results.¹¹³

110. Consent is not always a barrier to securing evidence having to do with the body of the accused. In his concurring opinion in *Rochin*, Mr. Justice Douglas says: “Of course an accused can be compelled to be present at the trial, to stand, to sit, to turn this way or that, and to try on a cap or a coat. . . . But I think that words taken from his lips, capsules taken from his stomach, blood taken from his veins are all inadmissible *provided they are taken from him without his consent*. They are inadmissible because of the command of the Fifth Amendment.” *Rochin v. California*, 342 U.S. 165, 179 (1951). (Italics added.)

111. *Breithaupt v. Abram*, 352 U.S. 432 (1956), permits blood to be extracted with a hypodermic needle from the body of an unconscious suspect, by an attending physician. In the dissent Mr. Chief Justice Warren writes: “Of course, one may consent to having his blood extracted or his stomach pumped and thereby waive any due process objection.” *Id.* at 441.

112. See *INBAU & REID* 133.

113. See text at notes 47-50 *supra*. To protect themselves, lie detector examiners would have to make sure that only one examiner testified at trial. A minority of jurisdictions have made an exception to the general exclusionary rule when parties have stipulated to this effect. See, e.g., *People v. Hauser*, 85 Cal. App. 2d 682, 193 P.2d 937 (Dist. Ct. App. 1948).

SUMMARY

The scientific basis for lie detection is questionable. There seems to be little evidence that upholds the claim to a regular relationship between lying and emotion; there is even less to support the conclusion that precise inferences can be drawn from the relationship between emotional change and physiological response.

The degree of interpretation required by the lie-detection process is high, certainly higher than is admitted; and the techniques of the process enhance the possibility of error from interpretation. Lie-detector tests are not comparable to chemical tests; the latter are based upon firmly established relationships and are, hence, more routinized in procedure. Chemical tests can be carried out with precision by a technician—mechanically, even though the moving parts of the mechanism are human.

Lie detection requires at least as much interpretation as tests performed by clinical psychologists or various medical specialists. It differs from those tests in three ways: first, it claims that it can diagnose a *particular* item of behavior. No reputable clinical psychologist or psychiatrist would assert a *general* ability to diagnose, from tests or interviews, whether an individual had sometime in the past performed a specific act. He would judge only tendencies to behave. Second, lie-detector examiners who are called upon to make this most difficult judgment usually have a police background; rarely, if ever, do they have the education and training required of psychological and medical specialists. Lie detection requires the making of physiological, psychological, and sociological judgments which even practitioners in those fields would draw only with caution. Third, the chance of distortion is magnified by the uncontrolled psychological response of the individual examiner. Considerable naiveté is shown regarding this factor; and the minimal precautions that are actually taken to offset examiner bias would be appropriate only for tests as precisely routinized as chemical tests.

Whatever the unconditional accuracy of the lie detector, the number of false positives it diagnoses is going to be related to the number of true positives in the population being tested. This fact would make the use of lie detectors, even if they had high unconditional accuracy, questionable in those situations—such as personnel screening—in which there are few true positives in the population.¹¹⁴ This point is true of any diagnostic procedure, not only of the lie-detector technique. Indeed, the adjudicatory process itself may be viewed as a series of diagnostic steps, with the conditional accuracy of each dependent upon the ability of earlier stages to filter out false positives.

The polygraph may be of some use in this respect, as a relatively benign method of procuring confessions. The additional question of whether it should actually be used by police in this manner, is answerable partly in relation to one's position on how far the state may go to induce confessions. For those who weigh the pragmatic heavily—who fear the guilty might escape out of tenderness to the accused, or even that harsher methods might be resorted to

114. See McInerney, *supra* note 92.

as an alternative to the polygraph—police use seems advisable. But the lesser evil argument is open to question. As a matter of state morality, traditional civil protections afforded the accused ought to be encouraged and carried out. A resort to this type of expediency puts the state in a curious moral position; admitting that it cannot control its sanctioning powers it permits police to lie,¹¹⁵ on the grounds that if they were not allowed this freedom they would revert to brutality. Such circumstances are more appropriate to a “Garrison State”¹¹⁶—in which those skilled in violence are most powerful—than to a democracy.

115. *But see* Spano v. New York, 360 U.S. 315, 322-24 (1959). In this case, the Supreme Court appears to rely heavily upon police deception of the accused as a ground for holding his confession inadmissible, and his conviction improper. In Leyra v. Denno, 347 U.S. 556, 559-61 (1954), the Supreme Court held that a confession was coerced where a psychiatrist was falsely introduced to the accused as a person intended to provide medical relief, when in fact his “techniques . . . were used to break petitioners will in order to get him to say he had murdered his parents.” *Id.* at 559.

116. See Lasswell, *The Garrison State and Specialists on Violence*, in *THE ANALYSIS OF POLITICAL BEHAVIOUR* 146 (1951).