Effectiveness of Techniques and Physiological Measures in the Detection of Deception

Report No. 76-5
Contract 75-NI-99-0001
National Institute of Law Enforcement and Criminal Justice
Law Enforcement Assistance Administration
U.S. Department of Justice

John A. Podlesny, M.A.
Research Assistant

David C. Raskin, Ph.D.
Project Director

Gordon H. Barland, Ph.D.
Project Co-Director

Department of Psychology
University of Utah
Salt Lake City, Utah 84112
August 20, 1976

Abstract

Control-question (CQ) and guilty-knowledge (GK) techniques for the detection of deception were studied in a mock theft context. Subjects from the local community received $5 for participation, and both guilty and innocent subjects were motivated with a $10 bonus for a truthful outcome on the polygraph exam. They were instructed to deny the theft when they were examined by experimenters who were blind with respect to their guilt or innocence. Eight physiological channels were recorded including a cardio activity monitor (CAM) and a low pressure blood pressure cuff (cardio). Numerical field evaluations of CQ tests produced 80% correct, 10% errors and 10% inconclusives. Excluding inconclusives, CQ decisions were 89% correct. Control questions were more effective than guilt-complex questions, particularly in identifying innocent subjects. There was some evidence that Backster control questions were more effective than Reid control questions. Numerical evaluations of GK tests were 90% correct with no inconclusives. Thus, there was no difference in accuracy of decisions using CQ and GK techniques. Quantitative analyses of the CQ results revealed significant discrimination between guilty and innocent subjects with measures of skin conductance response (SCR) amplitude, SCR recovery half-time, negative skin potential response (SPR) amplitude, heart rate response (HRR), finger blood volume (FBV) response amplitude and time, and finger pulse amplitude (FPA) response. Results with negative SPR amplitude, SCR recovery half-time, and SCR recovery half-time width indicated that Backster control questions identified guilty and innocent subjects significantly better than Reid control questions. The GK technique significantly discriminated between guilty and innocent subjects with SCR amplitude, SCR recovery half-time, negative and positive SPR amplitude, FBV response amplitude, and CAM systolic, diastolic, and pulse amplitude responses. The low pressure cardio identified only innocent subjects with the CQ technique and did not significantly discriminate guilty and innocent subjects with the GK technique.

Psychophysiological detection of deception (PDD) in its practical application is concerned with inferences about truthfulness or deception with regard to specific individuals and events. Its development is primarily the result of efforts on the part of field workers who have devised some ingenious and effective methods (Barland & Raskin, 1973; Reid & Inbau, 1966). Recently, the PDD has aroused the interest of scientists who have studied it in the field and in the laboratory (Barland & Raskin, 1973; Podlesny & Raskin, in press; Orne, Thackray & Paskewitz, 1972; Trovillo, 1942).

There are several advantages inherent in the laboratory study of PDD (Podlesny & Raskin, in press) most important of which is the ability to establish the guilt or innocence of the subject (ground truth). Laboratory study also allows control of extraneous variables and manipulation of variables as required. Furthermore, physiological data may be obtained with greater accuracy and in larger quantities using scientific laboratories and data processing rather than portable field equipment.

The major difficulty encountered in laboratory study of PDD concerns simulation of field deceptive contexts (Podlesny & Raskin, in press). Of the paradigms which have been used, the most valid appears to be utilization of a mock crime in which subjects perform certain acts which closely resemble a real crime while motivated with meaningful rewards for achieving a truthful outcome on a polygraph test (Raskin, 1975; Davidson, 1968). This study assesses the effectiveness of psychological techniques and physiological measures for detecting deception in the context of a mock crime.

Of the techniques currently in widespread application, the control-question (CQ) technique (Reid & Inbau, 1966) is the most sophisticated from a psychological point of view. In order to make inferences about truthfulness, the CQ technique compares the physiological responses to relevant questions which pertain to the suspect's involvement in the matter of interest with physiological responses to control questions. Control questions are formulated in a pretest interview in such a way that the suspect is very likely to be lying or not sure that he is completely truthful when he answers "no" during the test. Typical examples of control questions are given in the "Method" section. The subject's attention is focused on the control questions during the pretest interview and during review of the questions between the several charts which are run in a typical polygraph test. The importance of relevant questions is not emphasized. In that way an attempt is made to increase the salience of control questions.

The predicted relative magnitude of responses to control and relevant questions is different for innocent and guilty suspects. Innocent suspects, who know they are truthful concerning the relevant questions, are expected to produce greater responses to control questions. Guilty suspects should be less concerned about control questions and, therefore, should produce responses of greater relative magnitude to the relevant questions.

Lykken (1974) has argued that control questions can not be expected to function in that way, and that a better type of control question would concern a hypothetical crime of which the suspect has been led to believe that he is also suspected. Such questions are commonly employed in lie detection to determine the subject's level of responsivity to accusatory questions, and they are called guilt-complex questions (Barland & Raskin, 1973; Reid & Inbau, 1966). It should be emphasized that the question raised by Lykken can only be resolved empirically, and the results of research with the CQ technique (Raskin, 1975, Barland & Raskin, 1975) indicate that control questions do, in fact, function as predicted. This study is a further test of the validity of the CQ technique, and a comparison of the effectiveness of guilt-complex and control questions is also included.

The guilty-knowledge (GK) technique (Davidson, 1968; Lykken, 1960; Lykken, 1974) is based on the guilty suspect's recognition of information which is differentially available to him, since he has committed certain acts and the innocent suspect has not. A test may be constructed consisting of a series of multiple-choice questions with only one correct alternative. For example, a suspect in a bank robbery case might be asked in one question sequence to look at a series of similar pictures of people of which one is the actual teller who was robbed (the critical item). Innocent suspects are predicted to respond about equally to noncritical and critical items, while guilty suspects are expected to produce physiological responses of greater magnitude to the critical item.

The GK technique has received impressive demonstration of effectiveness in laboratory studies (Lykken, 1960; Davidson, 1968), but so far only the skin conductance response amplitude has been measured. Up to the present, the GK technique has received virtually no practical application chiefly because innocent, as well as, guilty suspects have been exposed to much of the information available to the examiner by the time a polygraph test is run (Lykken, 1974). Also, many criminal charges are of such a nature that an innocent person would know all of the relevant details of the incident. Extensive, time-consuming field work and considerable ingenuity are required to construct a valid GK test, while a CQ test may be used immediately. However, more extensive use of the GK technique may be expected, as field examiners become more sophisticated in handling cases.

This study extends knowledge of the validity of the GK technique with a variety of measures not yet studied. It also examines the GK and CQ techniques in the same deceptive context allowing comparison with respect to accuracy of decisions. Direct statistical comparison with objective measurement is not possible, since the predictions of the techniques are different with respect to responding of innocent subjects. However, it is of great interest to note which measures prove effective with the different techniques, since field application requires that knowledge.

Two types of control questions are presently being used in field PDD. Reid control questions (Reid & Inbau, 1966) place no limitation on the time period to which the question pertains, so it is possible that a control question may be construed as a relevant question by the subject. Backster (Backster, 1969) who is a prominent teacher of polygraphy in the field reasons that this could produce inconclusive results with deceptive subjects. He has introduced a form of control question which limits the time period to which control questions pertain, so that they clearly do not include the matter being investigated.

Furthermore, requiring the subject to recall events from a specific, prior period of his life may produce reactions caused by an increase in information processing (Podlesny & Raskin, in press). For examples of Reid and Backster control questions, see the "Method" section. Since it is of great practical importance to know which, if either, control question type is more effective, another purpose of this study was to make direct, objective comparison of the effectiveness of Reid and Backster control questions.

Since one purpose of this study was to examine the effectiveness of physiological measures which may be applied in PDD, a brief summary of research on each of the selected measures follows. In this study replication of earlier positive results was attempted, and promising measures which have received little or no study in the PDD were also examined. More thorough reviews may be found in Orne et al. (1972) and Podlesny and Raskin (in press).

Palmar skin resistance recordings are usually made during field polygraph examinations. Although some claims have been made that skin resistance is too labile to be useful in PDD (e.g. Reid & Inbau, 1966), research reports on skin conductance response (SCR) amplitude in mock crime contexts have been generally positive. Those reports include studies utilizing control-question techniques (Barland & Raskin, 1975; Raskin, 1975). SCR amplitude is the only measure which has been studied with the guilty-knowledge technique (Davidson, 1968; Lykken, 1960).

Recent reports have focused considerable interest on the recovery half-time of the SCR (Edelberg, 1970, 1972). Shorter recovery half-times may be related to goal-directed behavior, and longer times may indicate a defensive state. Therefore, longer times might be hypothesized to occur during deception. Kubis (1973) reported a high level of accuracy of detection in a mock crime context when SCR width was measured from the onset of the response to half recovery. In this study rise times and recovery half-times were analyzed separately.

Although palmar skin potential has been recorded in a few laboratory studies of PDD, only Raskin (1975) used a mock crime context. That study indicated that both negative and positive skin potential response (SPR) amplitude allowed significant detection of deception. However, discrimination was primarily due to larger amplitude responses of guilty subjects to relevant questions; innocent subjects did not produce larger responses to control questions.

Respiration recordings are usually made during field examinations using a bellows pneumograph. Respiration amplitude (RA) responses have been reported to discriminate deception in mock crime contexts (Barland & Raskin, 1975; Kubis, 1973; Raskin, 1975), with a decrease in RA indicative of deception. Change from prestimulus level has been the most sensitive measure. In addition changes in RA may be greater when innocent subjects answer control questions (Raskin, 1975). Respiration cycle time (RCT) was also studied in those experiments. Although Kubis (1973) reported negative results when only poststimulus measurements were made, Barland and Raskin (1975) found longer cycle times associated with deception. When RCT was measured as change from prestimulus cycle times, Raskin (1975) obtained a significant relationship between increased RCT and deception. In addition decreased RCT was associated with truthfulness.

Field polygraphs generally employ a blood pressure cuff applied to the upper arm to record relative blood pressure (Posey, Geddes, Williams

& Moore, 1969). This device is commonly called the "cardio." Until recently most field polygraphs employed mechanical recording cardios which required high pressures (about 90 mm/Hg) to operate. Barland and Raskin (1975) employed such a device in a mock crime context and used the control-question technique. They found significant detection of deception when cardio diastolic increases of greater amplitude were used as the criterion of deception, but obtained negative results with decreases in cardio pulse amplitude. Recently, electronic amplification of the cardio has been made available on field polygraphs so that they may be operated at lower pressures, eliminating much of the discomfort and danger due to occlusion of blood flow to the arm and also allowing continuous operation for longer periods. No research has yet been performed on the low pressure systems, and the validity of measures obtained with them is unknown. This study examined the validity of measures obtained using a low pressure cardio. Detailed analysis of the cardio recordings on a second-by-second basis is a novel contribution of this study.

Photoplethysmographs are presently coming into use in field PDD and are generally used to record finger pulse amplitude (FPA) and finger blood volume (FBV). Studies employing mock crimes indicate that decreases in FBV and FPA are valid indicators of deception (Kubis, 1973; Raskin, 1975). Kubis (1973) also reported that the duration or width of the FBV response is a useful measure. This study examined FBV responses using a long time constant and FPA using the same transducer coupled with a short time constant. FPA responses were analyzed on a second-by-second basis.

Heart rate (HR) has been measured in one PDD study employing a mock crime. Raskin and Hare (1976) measured sec-by-sec changes from prestimulus level for 20 sec following stimulus onset. They reported an acceleration phase reaching its peak at about the time of the subject's answer for guilty and innocent subjects responding to control and relevant questions, followed by a return to baseline for all except guilty subjects responding to relevant questions. Under the latter condition a differential deceleration phase appeared. Since HR acceleration has typically been associated with defensive responding and deceleration with attention, (Hare, 1973; Graham & Clifton, 1966; Raskin, Kotses & Bever, 1969), that finding challenges the common theory that responses of greater magnitude during deception are due to fear. The deceleratory HR response reported by Raskin could indicate an attentional response. This study also examined the HR response on a sec-by-sec basis. Another recent addition to some field polygraphs is the cardio activity monitor (CAM), a device which is claimed to measure relative pressure due to cardiovascular activity (Decker, Stein & Ansley, 1972). It is usually placed on the tip of a thumb or finger. No research employing mock crime contexts has yet been reported with the CAM, and its validity as a measure of deception is unknown. This study included a detailed examination of the form of the responses obtained from the CAM on a sec-by-sec basis.

Method

Subjects

Male subjects were recruited from the local community by a classified newspaper advertisement which offered $5 and a possible $10 bonus for participation in a psychophysiological experiment. Sixty subjects were selected from a total of 74 on the basis of no prior knowledge about the study, successful performance according to instructions, and satisfactory physiological recordings. Four of the subjects were randomly eliminated for purposes of statistical analysis. The age ranged from 14 to 71 with a median of 25 years. Education ranged from 8 to 19 years with a mean of 13.1 years. Subjects reported from 0 to 40 arrests (median = 1) for crimes other than traffic violations, and 11 subjects reported having had a prior polygraph examination.

Apparatus

The polygraph examinations were conducted in an Industrial Acoustics Company shielded chamber with the door left open. Physiological recordings were made on a Beckman Type R Dynograph located outside the chamber.

Skin conductance (SC) was recorded from Beckman 16mm Biopotential Ag-AgCl electrodes placed on thenar and hypothenar sites on the left palm which had been cleaned with 70% ethanol. A .16M NaCl electrolyte was used. Recordings were made using a Beckman 9844 skin conductance coupler with a constant applied potential of .5V. SC was recorded DC with an upper frequency cutoff of 6Hz.

Skin potential (SP) was recorded from Beckman 16mm Biopotential Ag-AgCl electrodes with a NaCl electrolyte. The active electrode was placed on the thenar surface of the left palm cleaned with 70% ethanol. The reference electrode was placed on the skin over the ulna about 6cm distal to the elbow. Electrical resistance at the reference site was reduced by repeated application and removal of Scotch tape. A Beckman 9806A AC-DC coupler was used, and recordings were made DC with an upper frequency cutoff of 30Hz.

Respiration recordings were obtained with a Beckman 7001 Respiration Waveform Belt (a strain gauge device) and a Beckman 9853A Voltage/Pulse/Pressure coupler. Respiration was recorded DC with an upper frequency cutoff of 30Hz. The belt was placed around the thorax on most subjects. With a few subjects thoracic recordings were unsatisfactory, and the belt was placed just below the ribcage.

Electrocardiograms (EKGs) were recorded from EKG lead II using a Beckman 9806A AC-DC coupler. The EKG was recorded with a time constant of .03 sec and an upper frequency cutoff of 30Hz.

The photoplethysmograph pickup was strapped with Velcro over the palmar surface of the tip of the 3rd finger of the left hand. It consisted of a Clairex CL703L CdSe photoconductive cell and a General Electric 683 miniature tungsten lamp mounted in a block of black phenolic plastic. Kodak Wratten Gelatin infrared filter No. 87C was placed over the photocell. The lamp was activated with a potential of 3V. A Beckman 9853A Voltage/Pulse/Pressure coupler contained the bridge circuitry and was used to record plethysmograms with a time constant of .1 sec and an upper frequency cutoff of 30Hz on one dynograph channel. The unfiltered output of the bridge was also connected to a Beckman 9806A AC-DC coupler which was modified for a time constant of 28 sec on a second channel. The high frequency cutoff on that channel was also 30Hz.

Cardio recordings were obtained from a pressure cuff placed on the subject's right upper arm and inflated to a pressure between 50 and 60mm/Hg during polygraph testing. The cuff was connected with latex tubing to a Stoelting Electronic Cardio Amplifier, and the output of that device was connected to the Dynograph with a 9806A AC-DC coupler. Cardio recordings were made DC with an upper frequency cutoff of 30Hz.

A Stoelting Dry Cardio Activity Monitor (CAM) was strapped firmly over the palmar tip of the second finger of the left hand. The CAM is a pressure sensitive device employing a metal diaphragm with miniature strain gauges attached to it. The diaphragm is mounted in a plastic block, and an elastic pad attached to the outside of the diaphragm makes contact with the skin. A Beckman 9853A Voltage/Pulse/Pressure coupler contained the bridge circuitry. CAM recordings were made DC with an upper frequency cutoff of 30Hz.

The outputs of the EKG, the short time-constant photoplethysmograph, cardio, and CAM channels were recorded on magnetic tape along with time marks using 2 Hewlett-Packard 3960 Instrumentation Recorders. Analog-to-digital conversion and preliminary data reduction were accomplished with a Digital Equipment Corp. PDP 12 computer. Further data reduction and analyses were performed on a Univac 1108 computer.

Procedure

Prospective subjects responded by telephone to classified newspaper ads. If they asked for information, they were told simply that the study concerned lie detection. If they wished to participate, they were told when to report to a particular room where they would find a note on the door. That note contained preliminary instructions which directed the subject to a cassette tape recorder in a desk drawer in an adjacent waiting room. If the subject asked for help, he was told that his instructions were on the note and that no further help could be given. The subject operated the tape and listened to the instructions using a headset.

Subjects were assigned in order of arrival to the next treatment combination on a randomized list containing 5 more entries than necessary to fill each cell in the experimental design. Half of the subjects were assigned to the Guilty condition and half to the Innocent condition. A secretary used the master copy of the randomized list to insert the taped instructions for the condition required. The experimenters did not know whether a given subject was guilty or innocent until all the data had been obtained from that subject and a decision had been made concerning his guilt or innocence.

Guilty subjects committed a mock theft. They searched for a specified secretary's office on a different floor of the building, approached the secretary and asked, "Where is Dr. Mitchell's office?" The secretary answered, "There is no Dr. Mitchell in this department." That interaction allowed the secretary to recognize the person as a subject. The subject then left the office and remained covertly observing until the secretary left the office. The secretary waited about 3 min and then, leaving her desk unlocked and carrying some papers, left the room. The subject then entered her office and searched the desk until he found a metal cash box. He took an envelope containing a heavy gold wedding ring from the cash box and removed the ring from the envelope. He concealed the ring on his person, destroyed the envelope, and returned to the waiting room. Guilty subjects were instructed to avoid leaving fingerprints and to prepare an alibi in case they were caught during the mock crime. They were instructed to commit the mock crime and return to the waiting room in exactly 15 min. Innocent subjects heard a brief description of the mock crime on their taped instructions, but did not enact it. They were not told any details of the mock crime. They only left the floor, waited for 15 min, and then returned to the waiting room.

All subjects were instructed that they would be given a lie detector test, and that they should deny any involvement in the mock crime or any knowledge of its details. A $10 bonus was offered to both innocent and guilty subjects for appearing truthful on the lie detector test. All subjects were told that they were suspected of taking a watch and that they should also deny having done that.

Polygraph testing was performed by three experimenters who were present during the entire examination. One experimenter acted as the polygraph examiner, a second attached the sensors and monitored polygraph recording, and the third monitored the tape recording.

The subject was advised of his rights, and the sensors were attached. A questionnaire pertaining to background information was administered, and the subject received a brief explanation of the polygraph, the autonomic nervous system, and a rationale for lie detection. He was then given a card test to allow him to become accustomed to the testing situation and to ensure adequate recordings.

Twenty guilty and 20 innocent subjects were given control-question tests (Barland & Raskin, 1975). The question sequence consisted of 10 questions with relevant questions at positions 5, 7, and 10 and control questions at 4, 6, and 9. A guilt-complex question pertaining to the watch was included at position 8. Half of those subjects received tests which included Backster (Bailey & Rothblatt, 1970) control questions which were limited with respect to the time period to which the question pertained and excluded the period of the mock crime. The other half received Reid control questions which did not have a time period limitation and included the period of the mock crime. Questions 1, 2, and 3 served primarily as buffers to habituate initial responding (Barland & Raskin, 1975). A typical Backster question sequence was a follows:

1. Is your name ________________?
2. Regarding the ring and the watch, do you intend to answer the questions about them truthfully?
3. Are you convinced that I will only ask questions on this test that you have already okayed?
4. Before the age of 18 did you ever steal any money?
5. Did you take that ring?
6. While you were in school did you ever take anything of value?
7. Did you take that ring from the desk?
8. Did you take that watch from room 702?
9. Before the age of 18 did you ever cheat anyone who trusted you?
10. Do you have that ring with you now?

Reid control questions usually began with "Have you ever" and did not refer to any age limits or time periods. For example, a Reid question

Number 4 might read "Have you ever stolen any money?" Subjects were instructed to answer each question with only "yes" or "no." Prior to the test, all questions were thoroughly reviewed with each subject, and the wording of the control questions was adjusted so that each control question elicited a "no" answer. Three or more charts were administered to each subject, but only the first three were used for objective scoring. Between charts the examiner questioned the subject about his subjective response to the control questions and made further adjustments as necessary in an attempt to increase the salience of the control questions.

Twenty subjects received guilty-knowledge tests which consisted of 5 charts each having a series of 6 alternatives to a different question concerning information related to the mock crime. The questions on those 5 charts were as follows:

Chart 1

Regarding the type of ring that may have been taken,

(1) Do you know if it was a sapphire class ring?

(2) Do you know if it was a pearl engagement ring?

(3) Do you know if it was a silver and turquoise ring?

* (4) Do you know if it was a gold wedding ring?

(5) Do you know if it was a ruby class ring?

(6) Do you know if it was a diamond engagement ring?

Chart 2

Regarding the floor of this building that the ring was hidden on,

(1) Do you know if it was the 1st floor?

(2) Do you know if it was the 12th floor?

(3) Do you know if it was the 6th floor?

(4) Do you know if it was the 4th floor?

* (5) Do you know if it was the 8th floor?

(6) Do you know if it was the 10th floor?

Chart 3

Regarding the number of the room that the ring was hidden in,

(1) Do you know if it was Room 800?

* (2) Do you know if it was Room 820?

(3) Do you know if it was Room 810?

(4) Do you know if it was Room 816?

(5) Do you know if it was Room 814?

(6) Do you know if it was Room 803?

Chart 4

Regarding the type of envelope that the ring was hidden in,

(1) Do you know if it was an inter-campus mail envelope?

(2) Do you know if it was a medium-sized manila envelope?

* (3) Do you know if it was a business-sized white envelope?

(4) Do you know if it was a small-sized manila envelope?

(5) Do you know if it was a small-sized white envelope?

(6) Do you know if it was a large-sized manila envelope?

Chart 5

Regarding the name of the doctor that the guilty person was instructed to ask for,

(1) Do you know if it was Dr. Trumbull?

(2) Do you know if it was Dr. Tolman?

(3) Do you know if it was Dr. Heisse?

(4) Do you know if it was Dr. Jordan?

(5) Do you know if it was Dr. Calvin?

* (6) Do you know if it was Dr. Mitchell?

The first alternative was included to buffer initial responding and was not scored. The remaining five alternatives consisted of a critical item and 4 noncritical items. The critical item was the correct alternative, and noncritical items were all incorrect. The critical items were positioned among the noncritical items in a pseudo-random order across charts. In the above list, critical items are identified with an asterisk. The rationale of the guilty-knowledge technique was explained to each subject in that group, and prior to each chart the question was reviewed, but no alternatives were stated until the test was administered except with Chart 4. Prior to that chart an example of each type of envelope was shown to the subject and specifically named. Subjects were instructed to answer "no" to each alternative on all the charts. With both the control-question and the guilty-knowledge tests a minimum of 15 sec was allowed to elapse between the verbal response to the question or alternative item and the beginning of the next question or item.

Following the completion of testing the charts were immediately evaluated to determine if the subject would receive the $10 bonus. Control-question charts were numerically scored using field evaluation criteria (Barland & Raskin, 1975) applied to the SC, respiration, and cardio channels which are typically used in field detection of deception. SC responses on each guilty-knowledge chart were examined to determine if the amplitude of the response to the critical item was the largest.

If the response to the critical item was the largest of the 5 responses on 3 or more of the 5 charts, the probability of that being a chance occurrence is approximately .06 and the subject was called deceptive. Otherwise, he was called truthful. All subjects who produced truthful outcomes on their polygraph test received the bonus.

Data Reduction

Numerical evaluations. The charts were scored blind by an independent rater who had not been present when the experiment was conducted. For control-question tests the criteria were similar to those described by Barland and Raskin (1975) and modified on the basis of results obtained by Raskin (1975). In addition to SC, respiration, and cardio, the plethysmograph was also scored for decreases in finger blood volume (FBV) and pulse amplitude (FPA). Each control-relevant pair was assigned a score between ±3 for each of the four channels depending on the magnitude of the difference between the responses. Negative scores were assigned when responses to relevant questions were stronger, and positive scores were assigned when responses to controls were stronger. Scores were summed across all standard charts obtained from each subject. If the total score was above +5, the subject was considered truthful; if it was below -5, the subject was judged deceptive; and scores between ±5 were considered inconclusive.

The CQ charts were also numerically evaluated to compare the effectiveness of the guilt-complex question (number 8) and the control question following it (number 9). The responses to those questions were separately compared with the responses to the relevant question at position

7. The scores were assigned in the manner stated in the preceding paragraph. Thus, the guilt-complex question was treated as a control-question, and its effectiveness as a control question was compared with the effectiveness of a standard control question. A total score was obtained for each subject on the first 3 charts.

Guilty-knowledge tests were evaluated using the criteria reported by Lykken (1959) and Davidson (1968). Each question sequence received a score of 2, 1, or 0 depending on the relative magnitude of the SCR to the critical item. If that response was the largest, a score of 2 was assigned; if it was second largest a score of 1 was assigned; otherwise, a score of 0 was assigned. The scores were totaled for the 5 question sequences for each subject. Subjects with scores of 6 or more were categorized as deceptive; those with scores below 6 were categorized as truthful. Scores were also obtained in the same way for the plethysmograph (a composite of FBV and FPA), respiration, and cardio.

Objective quantification. Objective measurements were made by persons who had no knowledge of the field evaluations or treatments administered to the subjects and also by computer. Measurements were made on the responses to each of the 3 control and 3 relevant questions on the first 3 charts for each CQ subject. For GK subjects, measurements were made on the responses to the last 5 items on each of the five charts for each GK subject. Analyses were based on mean values for control and relevant questions on each CQ chart and the value of the critical item and the mean of the last four noncritical items on each GK chart. The following measurements were obtained:

1. Skin Conductance Response (SCR) Amplitude. Increase in SCR was measured in mm of chart deflection from the onset of the first upward change in slope at least .5 sec after the beginning of the question to the highest level reached within 5 sec following the subject's answer. The responses were also converted to μmhos, but those analyses are not reported here since the responses measured in mm yielded more reliable results.

2. SCR Rise Time. The SCR rise time was measured in sec between the onset and the high point reached by the SCR as defined above.

3. SCR Recovery Half-Time. The recovery half time (Edelberg, 1970) was measured in sec for the single wave reaching the highest point between .5 sec after the onset of the question and 5 sec after the end of the answer. Where multiple peaks did not allow measurement of recovery half-time or where no measureable SCR occurred in CQ charts, the entire control-relevant pair was deleted. If the critical item could not be measured in a GK chart, the entire chart was deleted. Cell means were used to replace missing data in a small proportion of the data.

4. SCR Recovery Half-Time Width. The recovery half-time was measured for the entire SCR as defined under SCR amplitude above, and multiple peaks were disregarded. For both SCR recovery half-time and half-time width a limit of 15 sec was used as a maximum value.

5. Skin Potential Response (SPR) Amplitude. The change in mV was measured for the first wave within the period beginning .5 sec after

question onset and ending 5 sec after the end of the answer. If the first wave was negative and was followed by a positive wave, the change in mV for the positive wave was measured from the highest point reached by the negative wave. Only positive changes which reached their peak within 3 sec following their onset were defined as positive waves to eliminate scoring returns to baseline as positive waves.

6. Respiration Amplitude (RA) Response. The amplitude in mm of the last complete inspiration prior to the onset of the question was subtracted from the amplitude of the first complete inspiration following the end of the answer to yield a RA difference response. The amplitude of the poststimulus inspiration was also divided by the amplitude of the prestimulus inspiration to yield a proportional RA response.

7. Respiration Cycle Time (RCT) Response. The time in chart mm between the last 2 points of maximum inspiration prior to the onset of the question was subtracted from the time in mm between the first 2 points of maximum inspiration after the end of the answer to yield a RCT difference response. The poststimulus time was also divided by the prestimulus time to yield a proportional RCT response.

8. Heart Rate (HR) Response. Intervals between R waves in the EKG were converted to sec-by-sec HR. The value obtained for each second was the sum of the rates between the pairs of R waves overlapping the second weighted by the proportion of the second covered by each pair. Values were obtained for 2 sec prior to and 15 sec following question

onset. Deviations from prestimulus rates were obtained by subtracting the mean of the 2 prestimulus seconds from the rates for each of the 15 poststimulus seconds.

9. Finger Blood Volume (FBV) Response Amplitude. The decrease in mm of the diastolic level of the recording obtained from the photoplethysmograph channel with the 28-sec time constant was measured between the highest diastolic point within 4 sec following the onset of the question and the lowest diastolic point within 4-14 sec following question onset. The values obtained were corrected to a common gain.

10. Finger Blood Volume (FBV) Response Time. The time in sec was measured between the high and low points in the FBV response as defined above.

11. Finger Pulse Amplitude (FPA) Response. The .1-sec time-constant photoplethysmograph recordings were analyzed on a sec-by-sec basis. Each systolic and diastolic point was characterized by a relative level and a time of occurrence with respect to question onset. Sec-by-sec values for systolic and diastolic levels were obtained for the period beginning 2 sec before question onset and ending 14 sec after question onset. Those levels were the sum of the average levels for each pair of systolic or diastolic points weighted by the proportion of the sec covered by that pair. The difference between systolic and diastolic levels was obtained for each second, and the FPA response was expressed as proportion of the 2-sec prestimulus mean.

12. Cardio Pulse Amplitude (CPA) Response. The cardio recordings were measured in the same manner employed for FPA.

13. Cardio Systolic (CS) Response. The systolic sec-by-sec levels were obtained as described for FPA. CS responses were expressed as deviations from a 2-sec mean prestimulus level.

14. Cardio Diastolic (CD) Response. The CD response was obtained in the same manner as the CS response using the diastolic levels.

15. CAM Pulse Amplitude (CMPA) Response. The CMPA response was obtained from the CAM recordings in the same manner used for FPA response.

16. CAM Systolic (CMS) Response. The CMS response recordings were measured in the same manner used for the CS response.

17. CAM Diastolic (CMD) Response. The CMD responses were measured in the same manner used for CD responses.

Results[1]

Numerical Field Evaluations of Control-Question Tests.

Accuracy of decisions. All of the results of field evaluations are based on the numerical scores assigned by the independent blind evaluator who had no contact with the subjects. The outcomes of those evaluations are presented in Table 1. The combined decisions were 80% correct, 10% errors, and 10% inconclusives. Excluding inconclusives, the decisions were 89% correct. Although the results showed slightly higher accuracy of decisions with Backster control questions (94%) as compared to Reid control questions (83%), the difference was not statistically significant.

---

[1] All statistical tests employed a .05 rejection region, 2-tailed.

Table 1

Independent Rater Decisions Based on Total Numerical Field Scores for Tests Employing Backster and Reid Control Questions

	% Correct	% False Positive	% False Negative	% Inconclusive	% Correct Decisions
Backster	85	0	5	10	94
Reid	75	5	10	10	83
Combined	80	2	8	10	89

Evaluator agreement. Decisions made by the independent evaluator were compared to those made by the examiner immediately after each test was conducted. Both examiners made definite decisions in 90% of the cases, and they were in agreement on 100% of those decisions.

Effectiveness of numerical scores. Since each subject had at least three charts, the total numerical scores for the first three charts were compared for guilty and innocent subjects. The mean three-chart totals are presented in Table 2. Analysis of variance showed a significant difference between guilty and innocent subjects, $F(1/36) = 33.74$, but there was no significant interaction between Guilt and Control Question type, $F<1$. Tests on the means for the 3-chart totals showed significant identification of innocent subjects with Backster control questions, $t(9) = 6.28$, and Reid control questions, $t(9) = 3.07$. With Backster control questions there was also significant identification of guilty subjects, $t(9) = 3.16$, but there was not significant identification of guilty subjects with Reid control questions, $t(9) = 1.34$. To determine if there was a difference in the magnitude of scores for guilty and innocent subjects, an analysis of variance was performed with the sign reversed for the scores of guilty subjects. There was no significant difference in magnitude of mean scores for guilty compared to innocent subjects, $F(1/36) = 1.59$, and there was not a significant Guilt X Control Question Type interaction, $F<1$.

Effectiveness of Different Components. In order to assess the effectiveness of each of the four physiological components evaluated by numerical field scoring, the mean 3-chart total was calculated for each component

Table 2

Mean Total Numerical Field Scores of Guilty and Innocent Subjects for the First Three Charts with Backster and Reid Control Questions

	Backster	Reid	Combined
Guilty	-11.7	-6.3	-9.0
Innocent	+13.6	+14.8	+14.2

separately. Those means are shown in Table 3. Significant discrimination between guilty and innocent subjects was obtained with measures of skin conductance, $t(38) = 5.58$, and plethysmograph responses, $t(38) = 4.68$. Additional tests indicated significant identification of innocent subjects with respiration, $t(19) = 2.94$, and cardio, $t(19) = 5.02$. Respiration and cardio scores did not significantly identify guilty subjects, $t(19) = .35$ and $t(19) = 0.06$.

Effectiveness of control questions and guilt complex questions. The mean 3-chart total field scores of guilty and innocent subjects with a control question and the guilt-complex question scored against a relevant question are presented in Table 4. The Guilt X Comparison Item interaction was significant, $F(1/36) = 6.82$. When control questions were used as the comparison items, the mean scores for guilty and innocent subjects were of approximately the same magnitude, but in opposite directions. The guilt-complex comparison produced a negative mean score for guilty subjects but failed to produce a substantial positive mean score for innocent subjects, yielding only a mean of 0.0. The frequencies of scores in the wrong direction (i.e. positive values with guilty and negative values with innocent subjects) were tabulated. With the guilt-complex comparison, scores of "0" with guilty subjects were considered to be in the wrong direction since the prediction for that type of comparison is that innocent subjects would produce scores of about 0. With the control question comparison there were 3 scores in the wrong direction with guilty and 3 with innocent subjects. With the guilt-complex comparison there were 3 with guilty and 9 with innocent subjects.

Table 3

Mean Numerical Field Scores for Each Component and Total Scores for the First 3 Charts

	Plethysmograph	Respiration	Skin Conductance	Cardio	Total
Guilty	-4.2	+0.3	-5.0	.0	-9.0
Innocent	+3.1	+1.8	+5.6	+3.6	+14.2

Table 4

Mean Field Scores of Guilty and Innocent Subjects when a Control Question and the Guilt-Complex Question were Compared with a Relevant Question

	Comparison Question	
	Control	Guilt Complex
Guilty	-4.2	-3.7
Innocent	4.0	0.0

Thus, control questions were clearly more effective than guilt-complex questions in identifying innocent subjects.

Numerical Evaluation of Guilty-Knowledge Tests

Accuracy of decisions. Numerical evaluation of the guilty-knowledge tests (Lykken, 1959) using the SCR alone resulted in 90% correct and 10% false negatives. With the plethysmograph alone there were 75% correct, 20% false negatives, and 5% false positives. With respiration there were 65% correct and 35% false negatives, and with the cardio there were 60% correct, 35% false negatives, and 5% false positives.

Objective Quantitative Analyses

Skin Conductance Response (SCR) Amplitude. Mean SCR amplitude of guilty and innocent subjects to control and relevant questions is presented in Table 5. The Guilt X Question Type interaction was significant, $F(1/36) = 20.99$. Guilty subjects responded with greater amplitude to relevant questions, and innocent subjects responded with greater amplitude to control questions. There was no significant Control Question Type X Guilt X Question Type interaction, $F<1$, indicating no difference in the effectiveness of Backster and Reid control questions. Mean SCR amplitude for guilty and innocent subjects to noncritical and critical items is presented in Table 5. The Guilt X Question Type interaction was significant, $F(1/18) = 12.86$. Guilty subjects gave larger responses to critical items, and innocent subjects responded similarly to noncritical and critical items.

Skin Conductance Response (SCR) Rise Time. With the control-question technique mean SCR rise time was shorter to relevant questions (2.5 sec) than to control questions (3.3 sec), $F(1/36) = 16.73$, but there was no significant interaction of Guilt with Question Type, $F<1$. There were no

Table 5

Mean Skin Conductance Response Amplitude (mm) for Guilty and Innocent Subjects with the Control-Question and Guilty-Knowledge Techniques

	Control Question	
	Control	Relevant
Guilty	10.5	14.8
Innocent	12.6	8.9

	Guilty Knowledge	
	Noncritical	Critical
Guilty	8.7	16.1
Innocent	5.9	5.6

significant effects with the guilty-knowledge technique.

Skin Conductance Response (SCR) Recovery Half-Time. Mean SCR recovery half-time for guilty and innocent subjects to control and relevant questions with Backster and Reid control questions is presented in Table 6. The Control Question Type X Guilt X Question Type interaction was significant, $F(1/36) = 6.61$, as was the Guilt X Question Type interaction, $F(1/36) = 5.44$. Examination of Table 6 reveals that when Backster control questions were used, the mean SCR recovery half-time of guilty subjects was longer to relevant questions, and that of innocent _Ss_ was longer to control questions. However no such differentiation was found with Reid control questions. Mean SCR recovery half-times of guilty and innocent subjects to noncritical and critical items with the guilty-knowledge technique are presented in Table 6. Since most of the responses could not be scored for 1 subject in each group, the analysis used 9 subjects per cell. The Guilt X Question Type interaction was significant, $F(1/16) = 17.27$. Guilty subjects had longer SCR recovery half-times in response to critical items, and innocent subjects had similar recovery times in response to critical and noncritical items.

Skin Conductance Response (SCR) Recovery Half-Time Width. Mean SCR recovery half-time width for guilty and innocent subjects to control and relevant questions with Backster and Reid control questions is presented in Table 7. The Control Question Type X Guilt X Question Type interaction was significant, $F(1/36) = 6.40$, as was the Guilt X Question Type interaction, $F(1/36) = 10.01$. With Backster control questions mean SCR recovery half-time width was longer when guilty subjects responded to relevant

Table 6

Mean Skin Conductance Response Recovery Half-Time (sec) for Guilty and Innocent Subjects with the Control-Question Technique using Backster and Reid Control Questions and with the Guilty-Knowledge Technique

		Control Question	
		Control	Relevant
Backster	Guilty	2.2	3.2
	Innocent	2.4	1.4
Reid	Guilty	1.5	1.6
	Innocent	2.5	2.7

		Guilty Knowledge	
		Noncritical	Critical
GKT	Guilty	1.3	1.7
	Innocent	1.4	1.3

Table 7

Mean Skin Conductance Response Recovery Half-Time Width (sec) for Guilty and Innocent Subjects with Backster and Reid Control Questions

		Control	Relevant
Backster	Guilty	2.6	3.5
	Innocent	3.5	1.8
Reid	Guilty	2.1	2.4
	Innocent	3.1	3.1

questions and when innocent subjects responded to control questions. That relationship was not found with Reid control questions. With the guilty-knowledge technique there was no significant interaction of Guilt X Question Type, $F(1/18) = 3.06$.

Negative Skin Potential Response (SPR) Amplitude. Mean negative SPR amplitude for guilty and innocent subjects to control and relevant questions with Backster and Reid control questions is presented in Table 8. The Control Question Type X Guilt X Question Type interaction was significant, $F(1/36) = 4.41$, and the Guilt X Question Type interaction was also significant, $F(1/36) = 8.36$. With Backster control questions guilty subjects produced relatively larger responses to relevant questions, and innocent subjects produced relatively larger responses to control questions. Reid control questions did not produce that effect. Mean negative SPR amplitude for guilty and innocent subjects to non-critical and critical items is presented in Table 8. The interaction of Guilt X Question Type was significant, $F(1/18) = 8.64$. Guilty subjects gave larger responses to critical items, while innocent subjects gave responses of similar amplitude to noncritical and critical items.

Positive Skin Potential Response (SPR) Amplitude. With the control question technique no significant interaction of Guilt and Question Type was found with positive SPR amplitude, $F(1/36) = 3.93$. However, a significant Guilt X Question Type interaction was found with the guilty-knowledge technique, $F(1/18) = 4.69$. As shown in Table 9, guilty subjects had a significantly larger mean response to critical items, $t(18) = 2.47$, while the mean response of innocent subjects was not significantly different

Table 8

Mean Negative Skin Potential Response Amplitude (mV) for Guilty and Innocent Subjects with the Control-Question Technique using Backster and Reid Control Questions and with the Guilty-Knowledge Technique

		Control Question	
		Control	Relevant
Backster	Guilty	1.0	1.7
	Innocent	1.4	1.1
Reid	Guilty	1.1	1.2
	Innocent	.9	.8
		Guilty Knowledge	
		Noncritical	Critical
GKT	Guilty	1.3	2.5
	Innocent	1.4	1.2

Table 9

Mean Positive Skin Potential Response Amplitude (mV) for Guilty and Innocent Subjects to Noncritical and Critical Items

		Noncritical	Critical
GKT	Guilty	.5	.7
	Innocent	.7	.6

to noncritical and critical items, $t(18) = .62$.

Respiration Amplitude (RA) Response. Mean proportional and mean difference RA responses did not significantly differentiate guilty and innocent subjects with either the control-question or the guilty-knowledge technique.

Respiration Cycle Time (RCT) Response. Mean proportional and mean difference RCT response did not significantly differentiate guilty and innocent subjects with either the control-question or the guilty-knowledge technique.

Heart Rate (HR) Response. Mean sec-by-sec HR responses of guilty and innocent subjects to control and relevant questions are presented in Figure 1. The Guilt X Question Type X Seconds interaction was significant, $F(14/504) = 6.43$. Analyses of mean HR deviations from prestimulus rate were performed on Second 4, Second 11, and on the difference between Seconds 4 and 11 to determine which component of the HR response was responsible for that effect. The Guilt X Question Type interaction was significant for Second 11, $F(1/36) = 12.37$, and for Second 11 - Second 4, $F(1/36) = 10.84$, but not for Second 4, $F<1$. Thus, all subjects responded with an initial nondifferential acceleration peaking at about poststimulus Second 4. This was followed by a differential deceleratory phase with the largest mean differences around Second 11. The HR of guilty subjects decreased significantly more in response to relevant questions than to control questions between Seconds 4 and 11, $t(36) = 3.32$ and between prestimulus rate and Second 11, $t(36) = 3.98$. However, the HR of innocent subjects did not decrease significantly more to control questions than to relevant questions between Seconds 4 and 11, *t*(36) = 1.34, or between prestimulus rate and Second 11, *t*(36) = .99. Analysis of the HR responses of innocent subjects showed no significant interaction of Question Type X Seconds, *F*(14/252) = 1.70. Thus, guilty subjects responded to relevant questions with a late deceleration in HR, but innocent subjects did not have significantly different HR responses to control and relevant questions. There were no significant effects of Control Question Type X Guilt X Question Type, $F<1$, or Control Question Type X Guilt X Question Type X Seconds, $F<1$. With the guilty-knowledge technique there were no significant effects of Guilt X Question Type X Seconds, *F*(14/252) = 1.25 or Guilt X Question Type, *F*(1/18) = 3.04.

Finger Blood Volume (FBV) Response Amplitude. Mean FBV response amplitude for guilty and innocent subjects to control and relevant questions is presented in Table 10. The Guilt X Question Type interaction was significant, *F*(1/36) = 14.63. Guilty subjects produced larger responses to relevant questions, and innocent subjects produced larger responses to control questions. There was no significant interaction of Control Question Type with Guilt and Question Type, $F<1$. Mean FBV response amplitude for guilty and innocent subjects to noncritical and critical items is presented in Table 10. The Guilt X Question Type interaction was significant, *F*(1/18) = 7.82. Guilty subjects gave larger responses to critical items than to noncritical items, *t*(18) = 2.75, and innocent subjects did not have significantly different responses to noncritical and critical items, *t*(18) = 1.20.

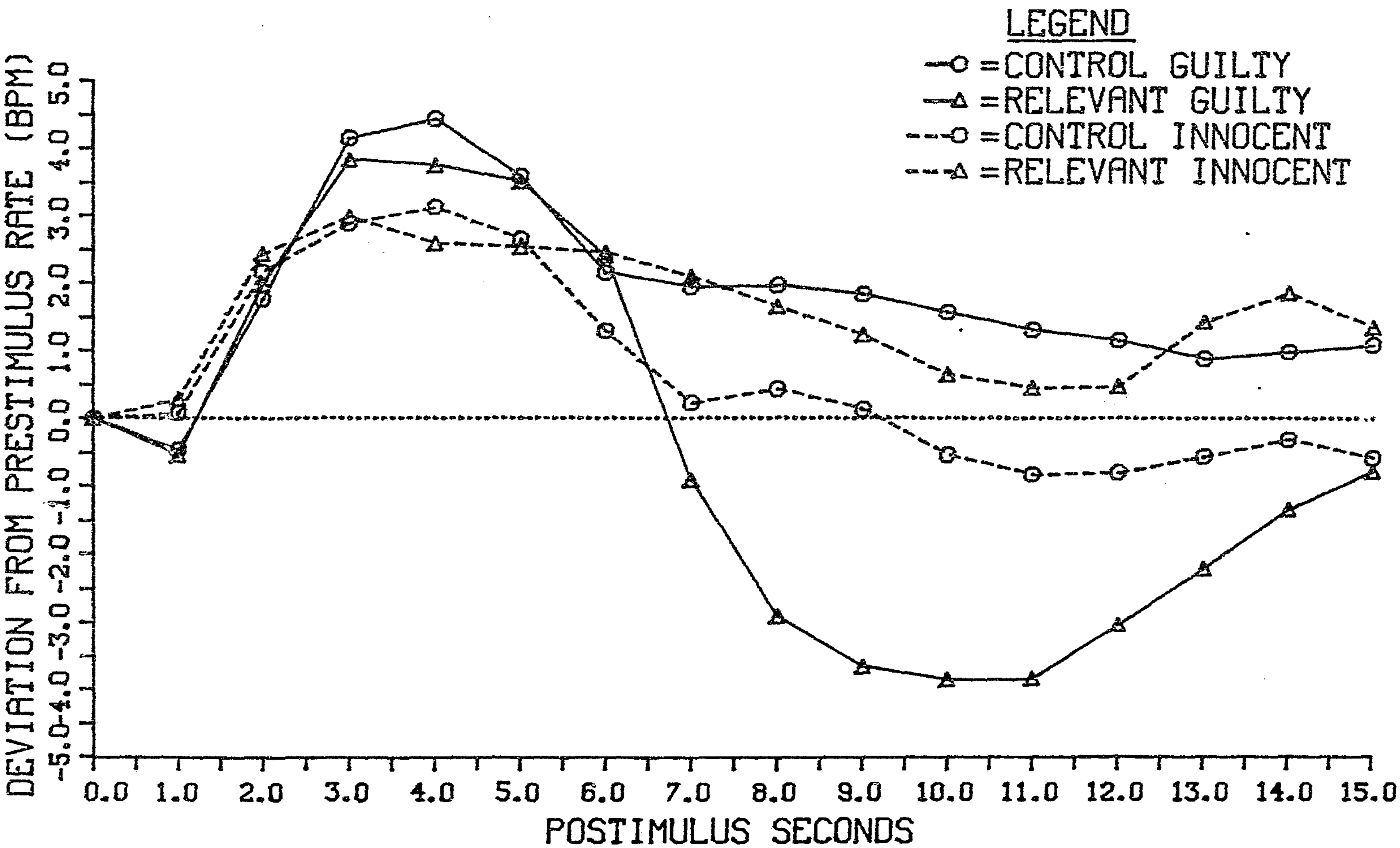

Figure 1. Mean Heart Rate Responses of Guilty and Innocent Subjects to Control and Relevant Questions

Table 10

Mean Finger Blood Volume Response Amplitude (Relative units) of Guilty and Innocent Subjects with the Control-Question and Guilty-Knowledge Techniques

	Control Question	
	Control	Relevant
Guilty	7.7	10.2
Innocent	6.8	5.9

	Guilty Knowledge	
	Noncritical	Critical
Guilty	7.6	9.3
Innocent	6.2	5.4

Finger Blood Volume (FBV) Response Time. Mean FBV response time of guilty and innocent subjects to control and relevant questions is presented in Table 11. There was a significant interaction of Guilt with Question Type, $F(1/36) = 5.33$. The responses of guilty subjects were longer to relevant questions, and innocent subjects had longer responses to control questions. There was not a significant Control Question Type X Guilt X Question Type interaction, $F<1$. With the guilty-knowledge technique there was no significant Guilt X Question Type interaction, $F<1$.

Finger Pulse Amplitude (FPA). Mean FPA responses of guilty and innocent subjects to control and relevant questions are presented in Figure 2. The Guilt X Question Type X Seconds interaction was significant, $F(13/468) = 4.32$, and the Guilt X Question Type interaction was also significant, $F(1/36) = 9.80$. Examination of Figure 2 indicates that guilty subjects produced greater decreases in FPA in response to relevant as compared to control questions. To determine if FPA reliably differentiated control from relevant questions with innocent subjects, the innocent group was analysed separately. The Question Type X Seconds interaction was significant, $F(13/234) = 3.42$. As shown in Figure 2, innocent subjects produced more prolonged decreases in FPA in response to control as compared to relevant questions. There were no significant effects of Control Question Type X Guilt X Question Type X Seconds, $F(13/468) = 1.15$, or Control Question Type X Guilt X Question Type, $F<1$. Mean FPA responses of guilty and innocent subjects to noncritical and critical items are presented in Figure 3. The interaction of Guilt X Question Type X Seconds was significant, $F(13/234) = 2.72$. Since inspection of Figure 3 suggests

Table 11

Mean Finger Blood Volume Response Time (sec) of Guilty and Innocent Subjects to Control and Relevant Questions

	Control	Relevant
Guilty	7.1	7.5
Innocent	7.0	6.5

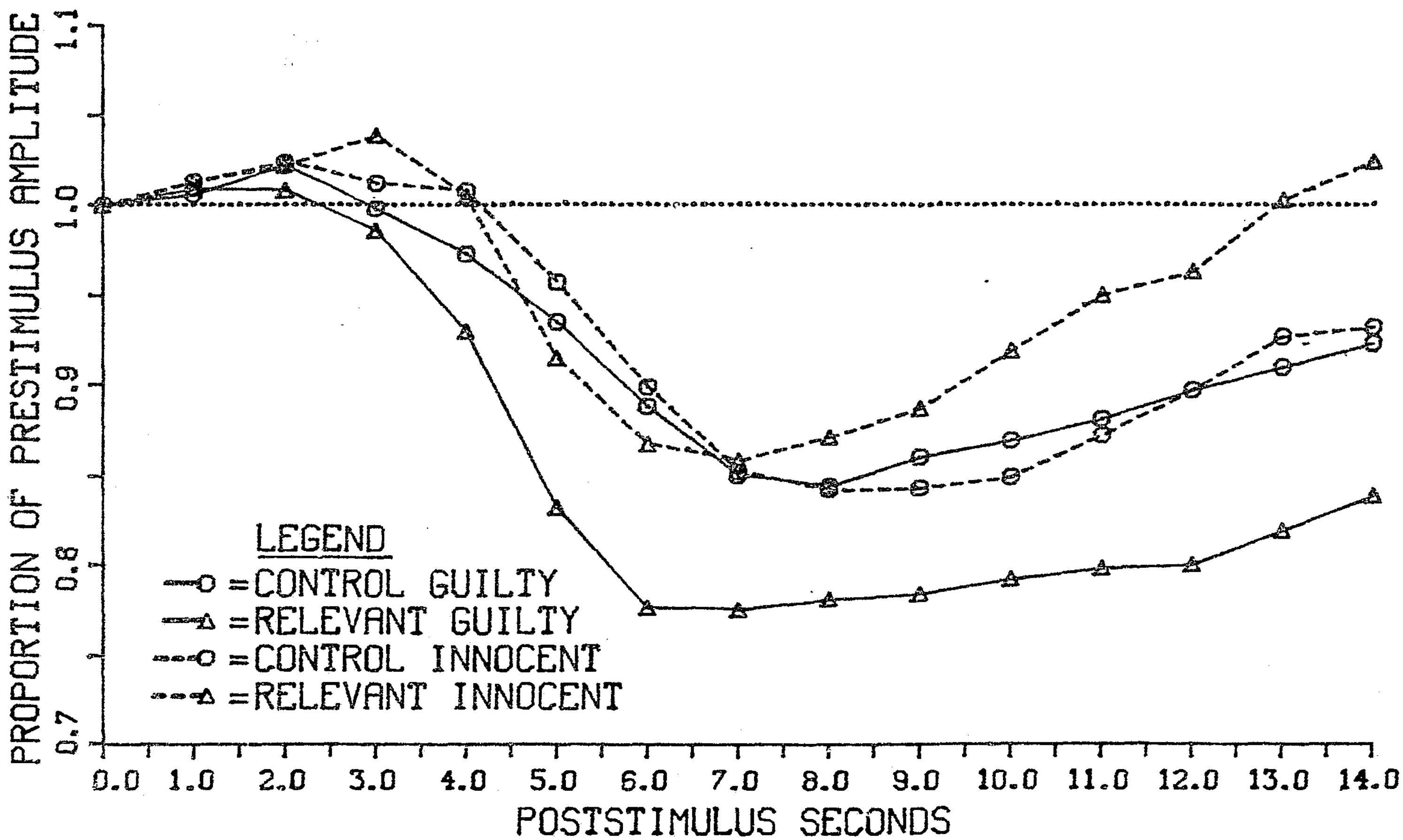

Figure 2. Mean Finger Pulse Amplitude Responses of Guilty and Innocent Subjects to Control and Relevant Questions

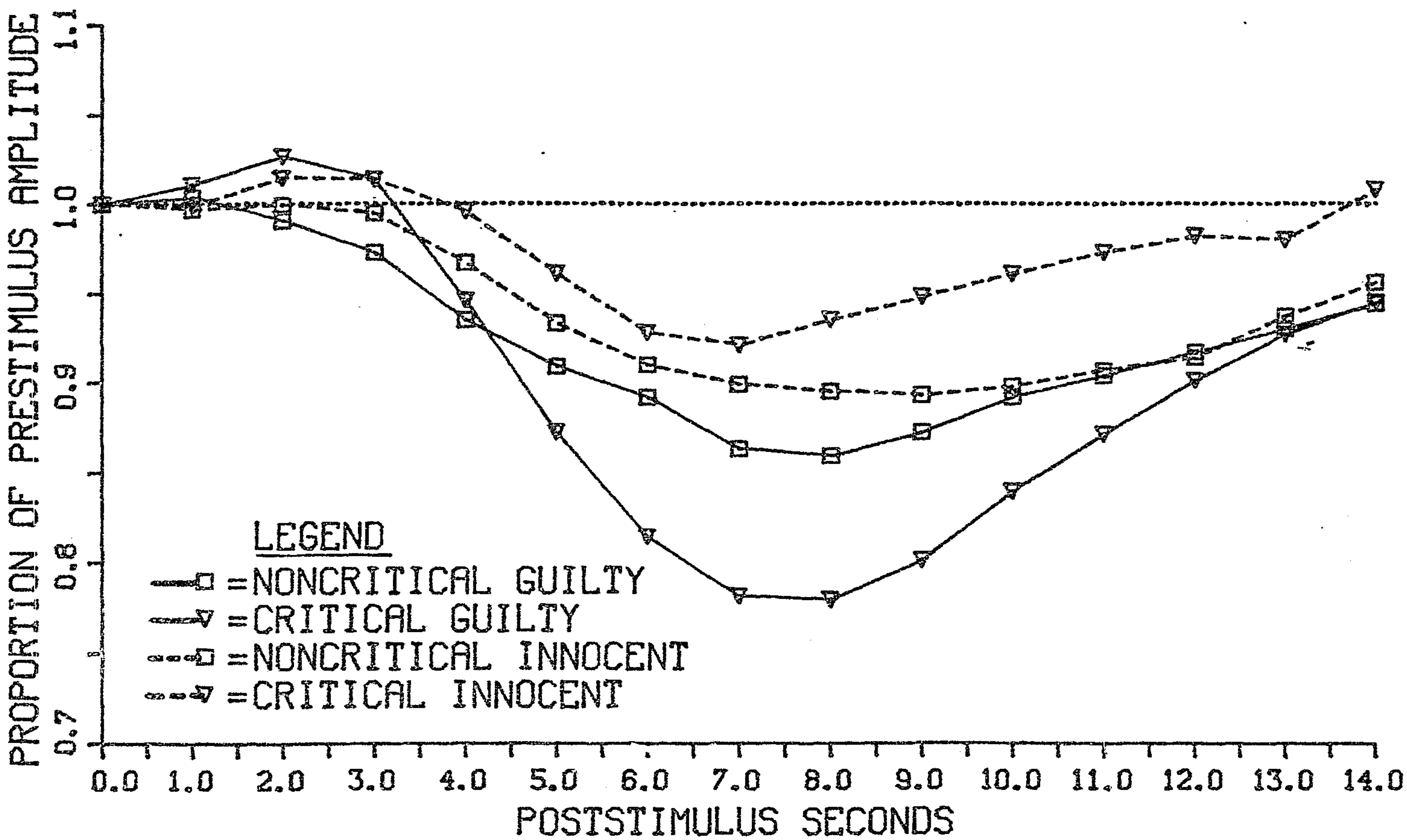

Figure 3. Mean Finger Pulse Amplitude Responses of Guilty and Innocent Subjects to Noncritical and Critical Items

that innocent subjects may have produced greater FPA responses to noncritical items, accounting for part of the interaction, analyses were performed on the innocent and guilty groups separately. Innocent subjects did not produce a significant Question Type X Seconds interaction, $F(13/117) = 1.05$, while guilty subjects did, $F(13/117) = 3.12$. As predicted, guilty subjects produced FPA responses which differentiated critical from noncritical items, while innocent subjects did not.

CAM Systolic (CMS) and Diastolic (CMD) Responses. All CAM control-question analyses were made using 6 subjects per cell rather than 10 because of inadequate recordings from some subjects. Mean CMS responses of guilty and innocent subjects to control and relevant questions are shown in Figure 4. The Guilt X Question Type X Seconds interaction was significant, $F(13/260) = 2.24$. However the responses of innocent subjects to control and relevant questions are clearly not different, and the interaction appears to be mostly due to greater decreases in systolic level of guilty subjects in response to relevant questions. Control-question subjects did not produce a Guilt X Question Type X Seconds interaction on the CMD measure, $F(13/260) = 1.35$. There were no significant effects of Control Question Type X Guilt X Question Type, $F(1/20) = 1.04$, or Control Question Type X Guilt X Question Type X Seconds, $F(13/260) = 1.39$. Mean CMS responses of guilty and innocent subjects to noncritical and critical items are shown in Figure 5. Both the Guilt X Question Type and Guilt X Question Type X Seconds interactions were significant, $F(1/18) = 5.11$, and $F(13/234) = 6.51$. Guilty-knowledge subjects also produced similar CMD responses with a significant Guilt X

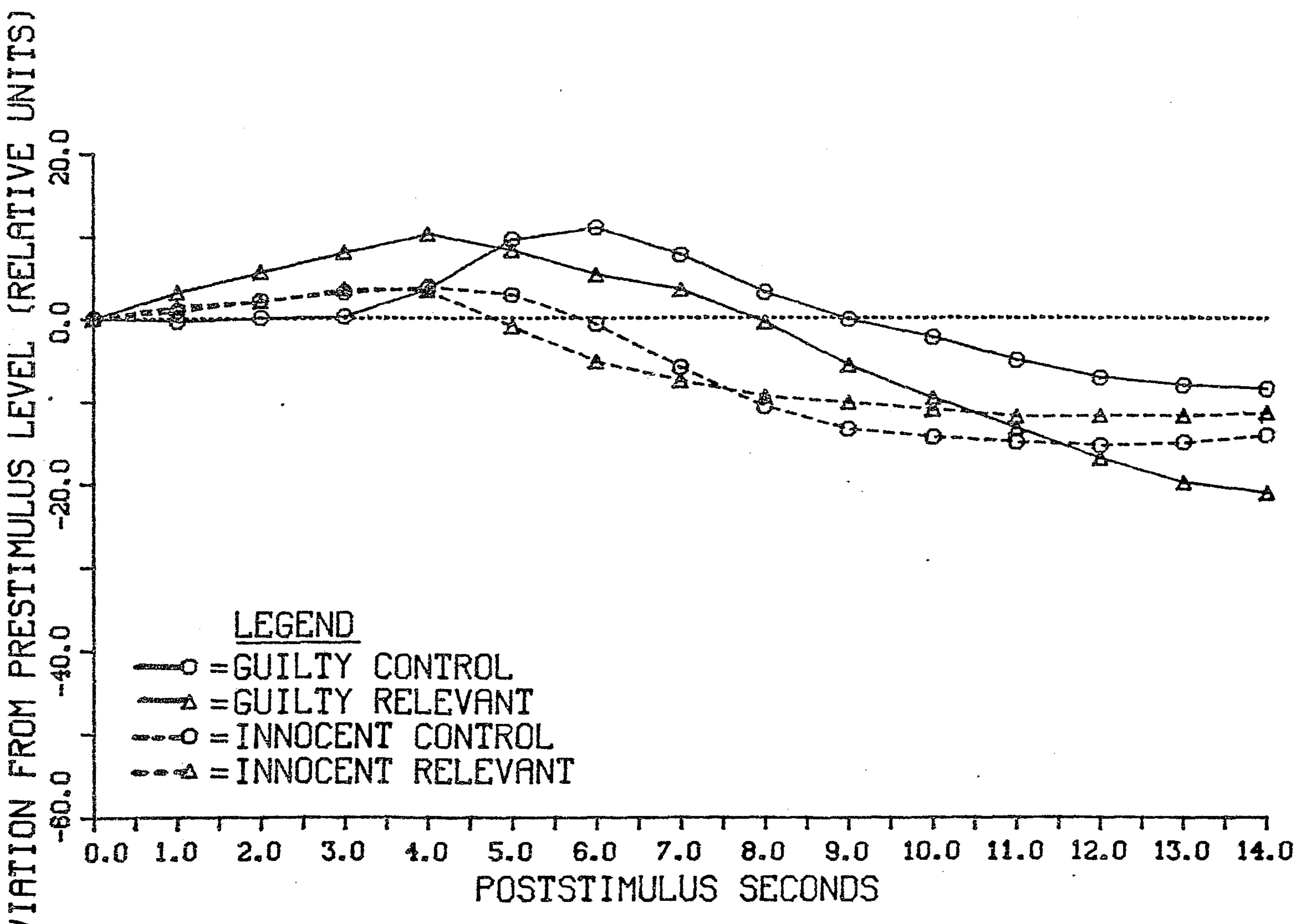

Figure 4. Mean CAM Systolic Responses of Guilty and Innocent Subjects to Control and Relevant Questions

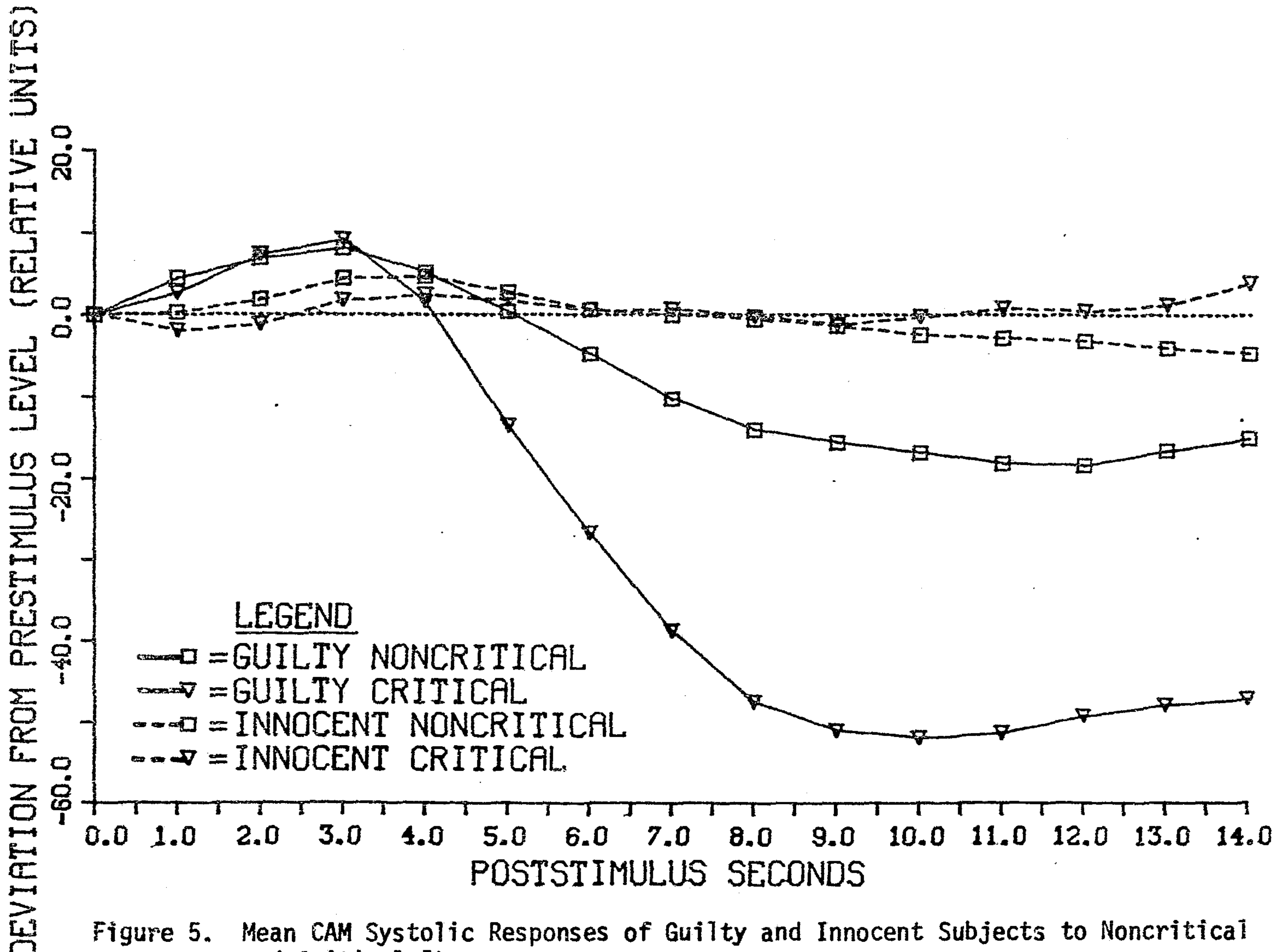

Figure 5. Mean CAM Systolic Responses of Guilty and Innocent Subjects to Noncritical and Critical Items

Question Type X Seconds interaction, $F(13/234) = 5.43$. Guilty subjects had greater mean decreases in CMS and CMD level in response to critical items. Innocent subjects produced similar, small responses to noncritical and critical items.

CAM Pulse Amplitude (CMPA). Control-question subjects did not produce a significant effect of Guilt X Question Type, $F(1/20) = 2.87$, or Guilt X Question Type X Seconds, $F(13/260) = 1.59$, with the CMPA measure, nor were there any effects of Control Question Type X Guilt X Question Type, $F(1/20) = 1.46$, or Control Question Type X Guilt X Question Type X Seconds, $F<1$. Mean CMPA responses of guilty and innocent subjects to noncritical and critical items are presented in Figure 6. Guilt X Question Type and Guilt X Question Type X Seconds interactions were both significant, $F(1/18) = 4.83$, and $F(13/234) = 3.29$. Guilty subjects produced greater mean CMPA decreases in response to critical items while innocent subjects produced small, similar responses to noncritical and critical items.

Cardio Systolic (CS) and Diastolic (CD) Responses. All CQ cardio analyses were performed with 8 subjects per cell due to inadequate recordings from some subjects. Since the CS results were essentially the same as the CD results, only the latter are reported here. Mean CD responses of guilty and innocent subjects to control and relevant questions are presented in Figure 7. The Guilt X Question Type X Seconds interaction was significant, $F(13/364) = 2.30$. The Guilt X Question Type interaction was also significant, $F(1/28) = 4.52$. Those effects were due to the differential responses of innocent subjects. Guilty subjects did not respond

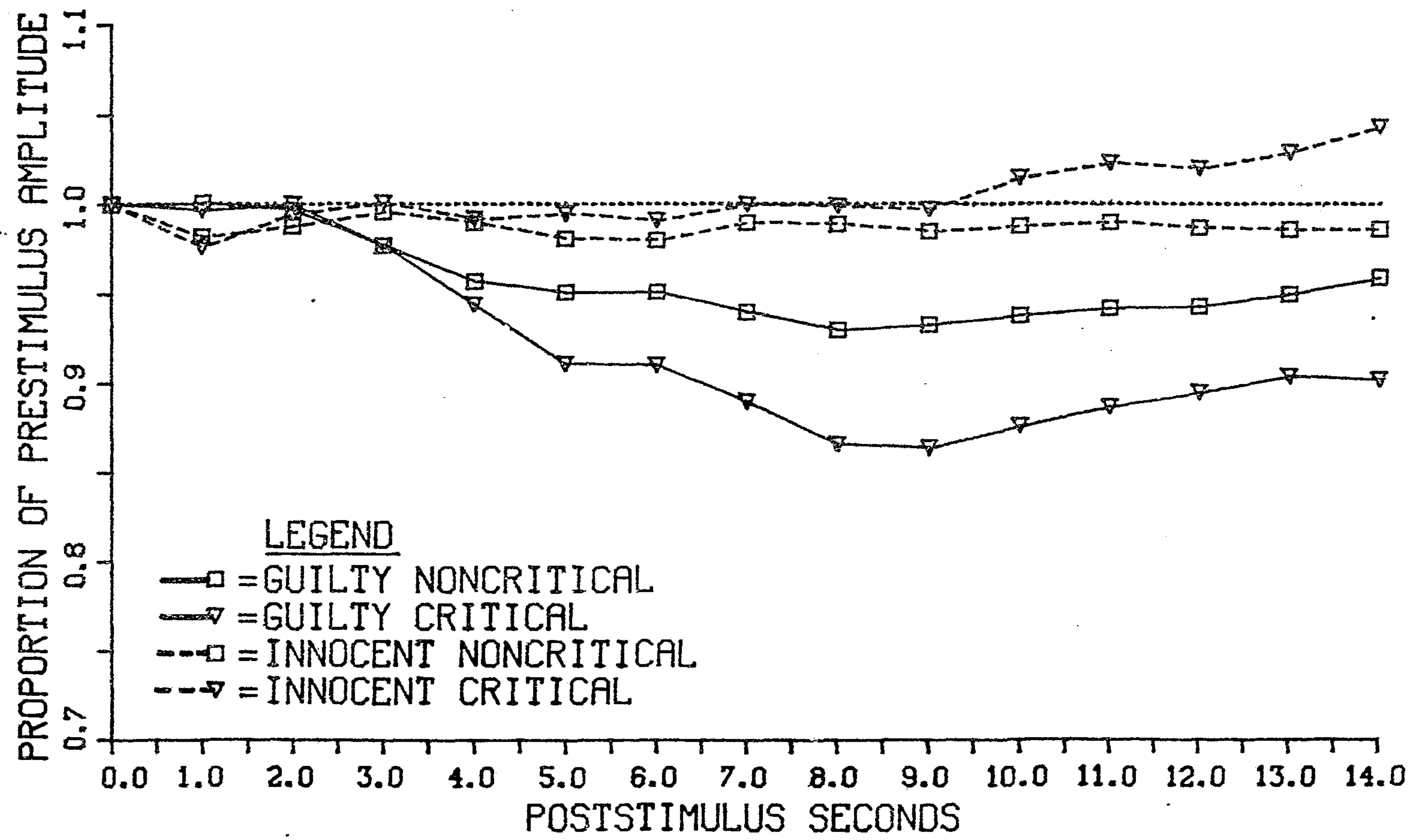

Figure 6. Mean CAM Pulse Amplitude Responses of Guilty and Innocent Subjects to Noncritical and Critical Items

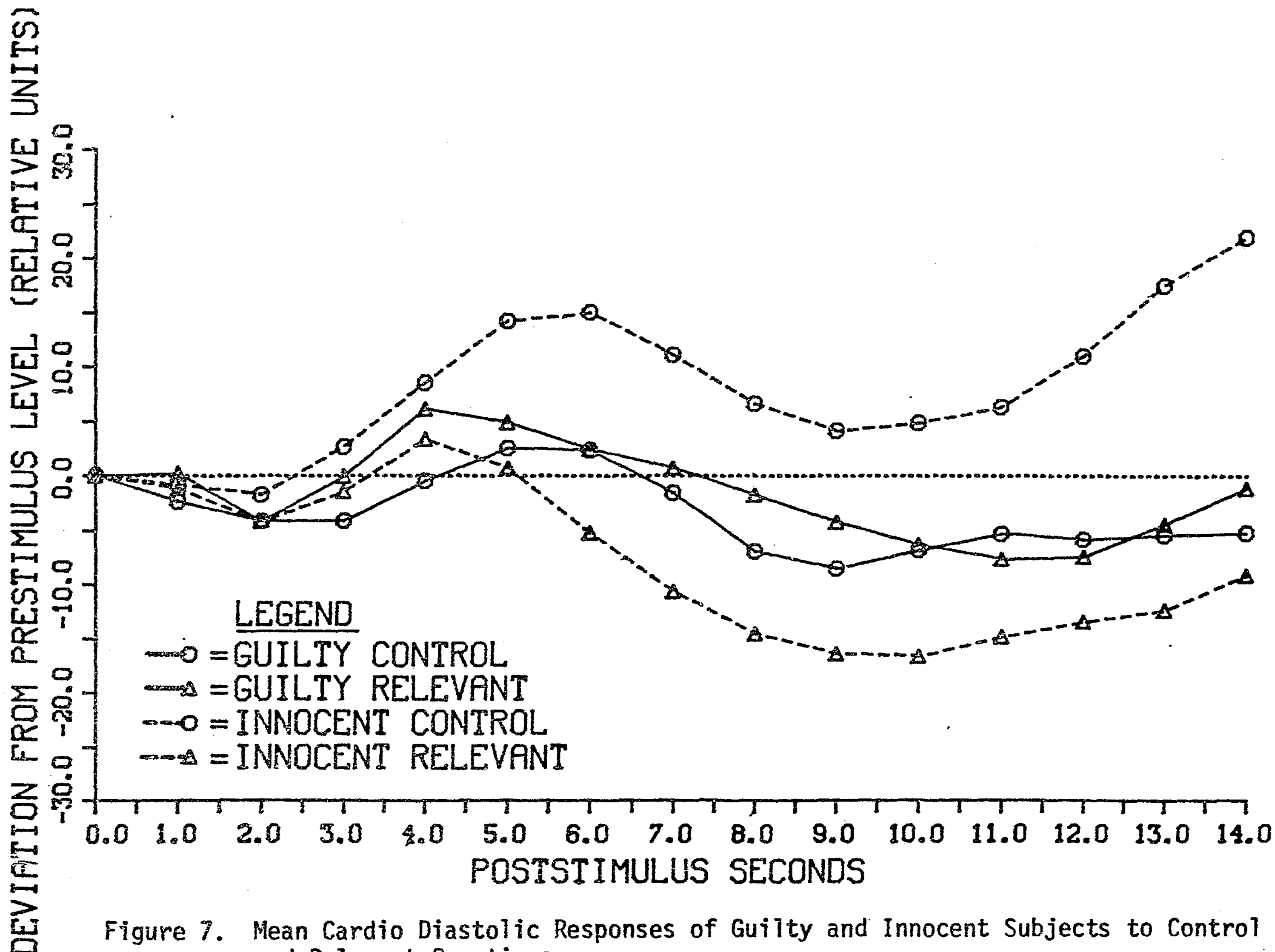

Figure 7. Mean Cardio Diastolic Responses of Guilty and Innocent Subjects to Control and Relevant Questions

differentially to control and relevant questions; however, innocent subjects responded to control questions with increases in mean CD level and to relevant questions with decreases. There were no significant effects of Control Question Type X Guilt X Question Type X Seconds or Control Question Type X Guilt X Question Type, *Fs*<1. All GK cardio analyses were performed with 9 subjects per cell due to inadequate recordings from two subjects. There were no significant effects of Guilt X Question Type X Seconds, F<1, or Guilt X Question Type, *F*<1, with the CD measure.

Cardio Pulse Amplitude (CPA) Response. CPA responses did not differentiate guilty and innocent subjects with either the CQ or GK technique.

Discussion

In general, the results of this study confirm the validity of the CQ and GK techniques for the detection of deception. They also suggest that Backster control questions may be preferable to Reid control questions. With respect to specific measures, the results agree substantially with previous reports. In addition, several novel measures were identified. Numerical field scoring of CQ tests resulted in significant identification of guilty and innocent subjects and a high level of accuracy of individual decisions (89%). Numerical evaluation of GK tests yielded similarly high accuracy (90%) without any inconclusives.

Numerical scoring of the different components produced a number of positive results with the CQ technique. Total scores showed significant discrimination between guilty and innocent subjects, and similar results were obtained for skin conductance and plethysmograph measures. In addition, respiration and cardio measures produced significant identification of innocent subjects.

Recent theoretical attacks on the efficacy of control question deception tests (Lykken, 1974) were not substantiated. The results of this study clearly demonstrate that control questions do function as predicted with a variety of measures. That prediction states that guilty subjects will produce responses of greater magnitude to relevant questions and innocent subjects will produce greater responses to control questions. The results of this study support that prediction in several ways. The size of numerical field scores for innocent and guilty subjects did not differ significantly, indicating that the scores of innocent subjects were about the same magnitude as those of guilty subjects and in the opposite direction. Objective quantification provided substantial evidence that control questions functioned as predicted with measures of SCR amplitude, SCR recovery half-time, SCR recovery half-time width, negative SPR, FBV amplitude, FBV response time, and FPA.

The assertion that guilt-complex type questions would be more effective than control questions (Lykken, 1974) is also not supported by the results of this experiment. When control questions and guilt-complex questions were compared using field numerical scores, control questions were significantly more effective in identifying innocent subjects, and control questions produced much lower frequencies of scores in the deceptive direction with innocent subjects.

When Backster control questions, which excluded the time period of the mock crime, were compared with Reid control questions, which did not temporally exclude the mock crime; the Backster type of control question was superior in some respects. Although numerical field scores produced significant identification of innocent subjects with both types of control questions, only the Backster control questions produced significant identification of guilty subjects using numerical scores. With the objective quantitative measures of SCR recovery half-time, SCR recovery half-time width, and negative SPR Backster control questions were significantly more effective than Reid control questions. In no instance were Reid control questions found to be superior to Backster controls. These results especially the prolongation of SCR, support an information-processing explanation of the greater effectiveness of Backster control questions (Podlesny & Raskin, in press).

The results also strongly support the validity of the GK technique. Numerical field scoring produced highly accurate decisions with skin conductance and moderately accurate decisions with the plethysmograph. However, accuracy with respiration and the cardio were only slightly better than chance. Noncritical and critical items operated as predicted with a variety of quantitative measures. With several measures guilty subjects produced significantly greater responses to critical items, and innocent subjects produced responses to noncritical and critical items which did not differ significantly. Those measures included SCR amplitude, SCR recovery half-time, negative SPR, positive SPR, FBV, FPA, CMS, CMD, and CMPA. With the exception of SCR amplitude (Davidson, 1968; Lykken, 1959) none of those measures had previously been studied with the GK technique.

The results with specific measures are generally in agreement with previous reports. The positive results with SCR amplitude replicate the results of other studies using the CQ technique (Barland & Raskin, 1975; Raskin, 1975) and GK technique (Davidson, 1968; Lykken, 1959). The results with negative SPR extend those of Raskin (1975) by showing that innocent subjects may produce greater mean responses to control questions. The finding that a differentially larger, long-latency HR deceleration occurred to relevant questions for guilty subjects but not for innocent subjects closely replicates the findings of Raskin and Hare (1976) and suggests that attentional processes may be important in the detection of deception with guilty persons. The results with FBV and FPA confirm those of Kubis (1973) and Raskin (1975) and also indicate that FPA can reliably differentiate innocent subjects. The latter finding may have been made possible by the detailed sec-by-sec analysis used in this study. Finally, the negative result with CPA obtained with a low pressure cuff agrees with the results reported by Barland and Raskin (1975) which were obtained with a high pressure cuff. However, the negative results with positive SPR using the CQ technique do not agree with the finding by Raskin (1975) that guilty subjects produced reliably greater positive SPRs in response to relevant questions.

The finding of positive results with numerical field scores of respiration only for innocent subjects and the lack of positive results with the objective quantitative measures of RA and RCT are in contrast to previous findings (Barland & Raskin, 1975; Kubis, 1973; Raskin, 1975).

Those studies, however, used pneumatic bellows transducers, while this study employed a device which consisted of a bent strip of spring steel strapped over the chest and having strain gauges attached to it. Comparison of the recordings obtained with that device to those from bellows or mercury strain gauge transducers indicates a relatively high level of noise (mainly from heart pulsations) and generally poor signal quality. Furthermore, the transducer used in this study is relatively obtrusive. Subjects may have attended more to their breathing than in previous studies, thus obliterating RA and RCT effects which might otherwise have been present. In addition, the partially positive findings with field numerical evaluations seem to indicate that the quantification of respiration recordings was not sensitive to information which was obtained by field evaluation procedures. Thus, there is reason to conclude that the weak respiration effects obtained in this study should not be interpreted as strongly contradicting previous findings.

The novel results of this study have important practical and theoretical implications. SCR rise time was shorter in response to relevant questions for both innocent and guilty subjects. Thus, while SCR amplitude and recovery time are valid indicators of deception, the rise time of the SCR is not useful. SCR recovery half-time and the more practical measure of SCR recovery half-time width were shown to be associated with deception using the CQ technique, and SCR recovery half-time was valid with the GK technique. Those results support the hypothesis that defensive responding may partly explain the efficacy of PDD, since Edelberg (1972) found that longer SCR recovery half-times were related to defensive states. However, longer recovery times may also be caused by prolonged information processing in response to questions which are more salient or require greater efforts to retrieve information. The SCR recovery half-time results with the GK technique, which depends largely on recognition of items, strongly support an information processing explanation.

The results obtained with the low pressure cardio cuff indicate that mean CD and CS responses significantly identified innocent but not guilty subjects with the CQ technique, and the cardio was not effective with the GK technique. Similar results were obtained with the field numerical evaluations of cardio responses with CQ and GK techniques. Thus, the effectiveness of the low pressure cardio may be limited. However, higher pressure devices may be useful, since Barland and Raskin (1975) reported discrimination of both guilty and innocent subjects. Furthermore, recent results obtained with criminal suspects in our laboratory using a pressure of 70 mm/Hg in the cuff have produced very encouraging identification of both deceptive and truthful subjects. The limited results obtained in the present study may have been due to the use of an inflation pressure which was too low and provided a poor coupling between the arterial system and the cuff as might be suggested by the report of Posey et al. (1969).

The CAM produced clear and significant differentiation with the GK technique, but only very marginal effects with the CQ technique. Those results indicated that the CAM may be of practical value when used with the GK technique but not with the CQ technique. Since the CAM is rigidly attached to the finger, it probably responds to changes in both blood pressure and vasomotor tone. If so, with the relatively emotion-provoking CQ technique, increases in BP and vasoconstriction may tend to cancel each other. The CAM may be more effective with the GK technique because differential blood pressure changes may not be present, as indicated by the results with the cardio. However, the useful results obtained with the GK technique were similar in response form to those obtained with the photoelectric plethysmograph. Thus, the CAM may have less usefulness than a finger plethysmograph.

In summary, the results of this study indicated that in a mock-crime context both the CQ and GK techniques were equally effective in terms of accuracy of decisions and with a variety of measures, and Backster control questions were somewhat more effective than Reid control questions. It should be noted that the use of the CQ technique results in a small percentage of inconclusive outcomes, whereas the GK technique permits a decision in every case.

The findings of this study are of practical importance to the application of PDD and support its validity. However, field research with various techniques and measures remains extremely important due to the difficulty inherent in making inferences from the laboratory to field contexts. Furthermore, the limited applicability of the GK technique in field settings and its lack of demonstrated superiority over CQ techniques in accuracy of decisions leave the CQ technique as the current method of choice in criminal investigation.

References

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

Bailey, F. L., & Rothblatt, H. B. Investigation and preparation of criminal cases, federal and state. Rochester, N. Y.: The Lawyers Cooperative Publishing Company, 1970, 296-299.

Barland, G. H., & Raskin, D. C. Detection of deception. In W. F. Prokasy & D. C. Raskin (Eds.), Electrodermal activity in psychological research. New York: Academic Press, 1973.

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

Davidson, P. O. Validity of the guilty knowledge technique: The effects of motivation. Journal of Applied Psychology, 1968, 52, 62-65.

Decker, R. E., Stein, A. E., & Ansley, N. A cardio activity monitor. Polygraph, 1972, 1, 108-124.

Edelberg, R. The information content of the recovery limib of the electrodermal response. Psychophysiology, 1970, 6, 527-539.

Edelberg, R. Electrodermal recovery rate, goal orientation, and aversion. Psychophysiology, 1972, 9, 512-520.

Graham, F. K., & Clifton, R. C. Heart rate change as a component of the orienting response. Psychological Bulletin, 1966, 65, 305-320.

Hare, R. D. Orienting and defensive responses to visual stimuli. Psychophysiology, 1973, 10, 453-464.

Kubis, J. F. Analysis of polygraphic data, part 2. Polygraph, 1973, 2, 89-107.

Lykken, D. T. The GSR in the detection of guilt. Journal of Applied Psychology, 1959, 43, 385-388.

Lykken, D. T. Psychology and the lie detector industry. American Psychologist, 1974, 29, 724-739.

Orne, M. T., Thackray, R. I., & Paskewitz, D. A. On the detection of deception. In N. S. Greenfield & R. A. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart and Winston, 1972.

Podlesny, J. A., & Raskin, D. C. Physiological measures and the detection of deception. Psychological Bulletin, in press.

Posey, J. A., Geddes, L. A., Williams, H., & Moore, A. G. The meaning of the point of maximum oscillations in cuff pressure in the indirect measurement of blood pressure. Part I. Cardiovascular Research Center Bulletin, 1969, 8, 15-25.

Raskin, D. C. Psychopathy and detection of deception in a prison population. (Report No. 75-1, Contract 75-NI-00-0001, U. S. Department of Justice). Salt Lake City, Utah: University of Utah, Department of Psychology, June, 1975.

Raskin, D. C., & Hare, R. D. Psychopathy and detection of deception in a prison population. Manuscript submitted for publication, 1976.

Raskin, D. C., Kotses, H., & Bever, J. Cephalic vasomotor and heart rate measures of orienting and defensive reflexes. Psychophysiology, 1969, 6, 149-159.

Reid, J. E., & Inbau, F. E. Truth and deception: The polygraph ("lie detector") technique. Baltimore: Williams & Wilkins, 1966.

Trovillo, P. V. A history of lie detection. Journal of Criminal Law and Criminology, 1939, 29, 848-881.

END

7 [illegible]