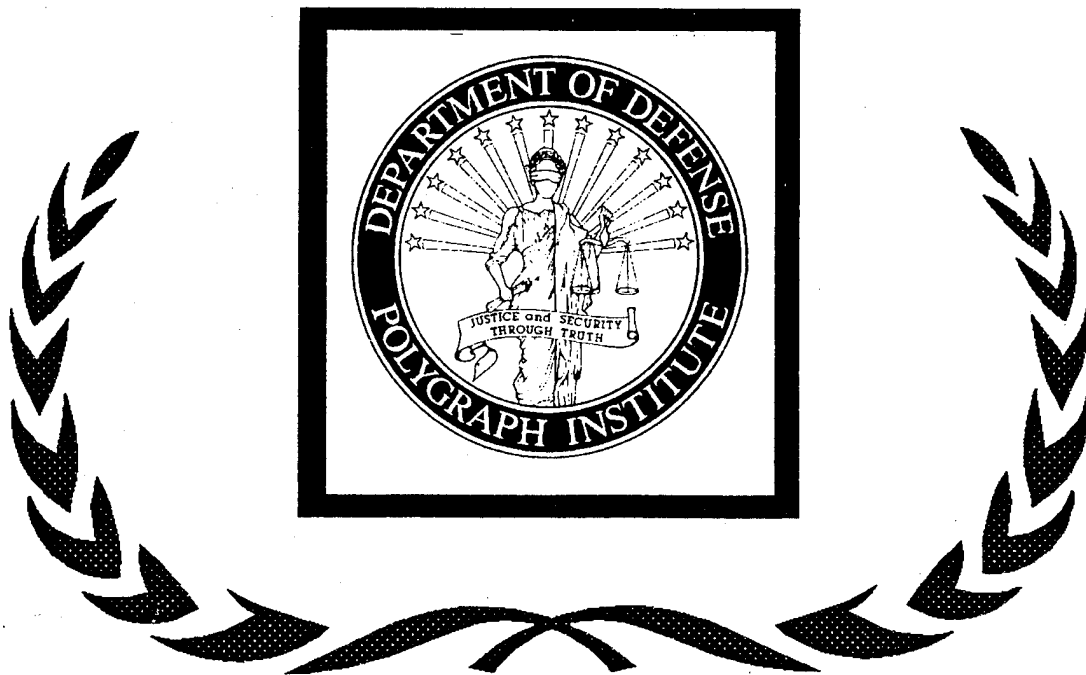




The Effects of Aural Versus Visual Presentations of Questions During a Detection of Deception Task

Barbara L. Carlton, Ph.D. and Brenda J. Smith, MS

January 1991

Department of Defense Polygraph Institute
Fort McClellan, Alabama 36205-5114
Telephone: 205-848-3803
FAX: 205-848-5332

19960221 139

DTIC QUALITY INSPECTED 1

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE January 1991	3. REPORT TYPE AND DATES COVERED Final Report (Oct 90 - Jan 91)
4. TITLE AND SUBTITLE The Effects of Aural Versus Visual Presentations of Questions During a Detection of Deception Task			5. FUNDING NUMBERS DoDPI91-P-0003
6. AUTHOR(S) Barbara L. Carlton and Brenda J. Smith			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Department of Defense Polygraph Institute Building 3165 Fort McClellan, AL 36205-5114			8. PERFORMING ORGANIZATION REPORT NUMBER DoDPI91-R-0002
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Department of Defense Polygraph Institute Building 3165 Fort McClellan, AL 36205-5114			10. SPONSORING/MONITORING AGENCY REPORT NUMBER DoDPI91-R-0002
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Public release, distribution unlimited			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) The purpose of this research was to investigate the relationship between accuracy of a detection of deception task and the stimulus mode of the question presentation. That is, will the presentation of questions on a computer screen change the accuracy rate when compared to exams conducted, more traditionally, in a verbal mode? Eighty subjects were assigned to either a guilty or innocent condition. Guilty subjects were shown a video of a mock crime scenario, while innocent subjects viewed a clip from a training video. Half of the innocent and half of the guilty groups were given the exams aurally using a tape recorder, and the other half were shown the questions on a computer terminal. Subjects were then given a guilty knowledge test by the experimenter using a Coulbourn polygraph. Overall accuracy of the decisions of the original examiner was 78%, 74% for the blind examiner, and 76% for the Lykken system. Accuracy rates for subjects in the visual condition were 83% for the original examiner, 78% for the blind evaluator, and 70% for the Lykken system. The decisions for the aural condition were 73% accurate for the blind examiner, 70% accurate for the blind evaluator, and 83% accurate for Lykken scoring system. There was no significant association between an accurate decision and the stimulus mode condition for the original examiner, the blind evaluator, or the Lykken scoring decision.			
14. SUBJECT TERMS detection of deception, stimulus mode, aural presentation, visual presentation, guilty knowledge test			15. NUMBER OF PAGES 35
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT

Report No. DoDPI91-R-0002

The Effects of Aural Versus Visual
Presentations of Questions
During a Detection of Deception Task

Barbara L. Carlton, Ph.D. and Brenda J. Smith, MS

January 1991

Department of Defense Polygraph Institute
Fort McClellan, Alabama 36205

Director's Foreword

This report describes a unique study designed to determine if the accuracy of psychophysiological detection of deception (PDD) examination decisions are affected by the manner of question presentation. The authors interpret the results as indicating that the question presentation mode has very little influence on detection accuracy rates. Examination results were scored by the original examiner, a blind examiner, and using an objective scoring system. No differences were found among the three scoring system accuracy rates.

The results of this study clearly suggest that the accuracy of PDD examination decisions is not influenced by whether examinees hear or read the questions. This conclusion should, however, be interpreted with some caution. The project was designed, implemented, and reported with above average scientific rigor. However, the non-standard procedure of having an collaborator within the subject's visual field throughout testing could have unduly influenced the results. In addition, negative results are always questionable in the absence of a statistical power analysis to indicate whether the data were sufficient to test the hypothesis. This project does, however, provide a basis for resolving the question as to the influence of visual versus aural question presentation on PDD examination decision accuracy.



Michael H. Capps
Director

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

Abstract

CARLTON, Barbara L. and SMITH, Brenda J. The effects of aural versus visual presentations of questions during a detection of deception task. January 1991, Report No. DoDPI91-R-0002. Department of Defense Polygraph Institute, Ft. McClellan, AL 36205.--The purpose of this research was to investigate the relationship between accuracy of a detection of deception task and the stimulus mode of the question presentation. That is, will the presentation of questions on a computer screen change the accuracy rate when compared to exams conducted, more traditionally, in a verbal mode? Eighty subjects were assigned to either a guilty or innocent condition. Guilty subjects were shown a video of a mock crime scenario, while innocent subjects viewed a clip from a training video. Half of the innocent and half of the guilty groups were given the exams aurally using a tape recorder, and the other half were shown the questions on a computer terminal. Subjects were then given a guilty knowledge test by the experimenter using a Coulbourn polygraph.

While the polygraph exam was being administered, a second experimenter sat across from the subject. This second experimenter was responsible for programming the subject, while the experimenter running the exam was blind to the subject's guilt/innocent status. During the exam, the subject was required to respond to the experimenter with "no" to every item. The charts were scored by the following: (1) the original examiner; (2) a blind evaluator; and (3) using a scoring system introduced by Lykken. Overall accuracy of the decisions of the original examiner was 78%, 74% for the blind examiner, and 76% for the Lykken system. Accuracy rates for subjects in the visual condition were 83% for the original examiner, 78% for the blind evaluator, and 70% for the Lykken system. The decisions for the aural condition were 73% accurate for the blind examiner, 70% accurate for the blind evaluator, and 83% accurate for Lykken scoring system. There was no significant association between an accurate decision and the stimulus mode condition for the original examiner, the blind evaluator or the Lykken scoring decision. ($x^2 = .6091$; $p < .4351$ and $x^2 = 2.0378$; $p < .1534$; $x^2 = 1.065$, $p < .3020$). There was no significant association between the type of error and the stimulus mode for the original examiner (Fisher's exact $p < .14$) or the decision rendered by the Lykken system (Fisher's exact $p < .25$) whereas the type of error was associated with stimulus mode for the blind examiner (Fisher's exact $p < .0075$). This may be due to an artifact associated with the use of the experimenter as a confronter during the exam.

Key-words: detection of deception, stimulus mode, aural presentation, visual presentation, guilty knowledge test (GKT)

Table of Contents

Title Page	i
Director's Foreword	ii
Abstract	iii
List of Figures	v
List of Tables	vi
Method	2
Subjects	2
Equipment/Apparatus	2
Procedure/Method	2
The Confronter	5
Scoring	6
Original Examiner	6
Blind Evaluator	7
The Lykken Scoring System	7
Results	7
Questionnaire Results	7
Polygraph Examination Results	9
Overall Accuracy	9
Role	10
Stimulus Mode	12
Distribution of correct calls as a function of stimulus mode	13
Distribution of error-type as a function of stimulus mode	14
Discussion	16
References	20
Appendix A: Description of Research/Volunteer Affidavit	A-1
Appendix B: Examination Question and Alternatives	B-1
Appendix C: Research Questionnaire	C-1

List of Figures

1. Percentage of overall accuracy of all three scoring systems; i.e., the original examiner, the blind evaluator, and the Lykken scoring system . . . 10
2. Percentage of accurate decisions for both innocent and guilty subjects by the original examiner, blind evaluator, and those made using the Lykken scoring system 10
3. Percentage of correct decisions made in the two stimulus mode conditions by the original examiner, the blind evaluator, and those made using the Lykken system 12

List of Tables

1.	Probability, frequency, and expected frequency distributions of innocent subject currently identifying critical items	8
2.	Frequency of correctly identified critical items for each question	9
3.	Inter-scoring system/evaluator matrix	9
4.	Contingency table for role versus decision of the original examiner	11
5.	Contingency table for role versus decision of the blind evaluator	11
6.	Contingency table for role versus decision of the Lykken scoring system	11
7.	Distribution of the correct original examiner decisions as a function of stimulus mode	13
8.	Distribution of the correct blind evaluator decisions as a function of stimulus mode	13
9.	Distribution of the correct Lykken scoring decisions as a function of stimulus mode	14
10.	Distribution of the errors made by the original examiner as a function of stimulus mode	14
11.	Distribution of errors made by the blind examiner as a function of stimulus mode	15
12.	Distribution of the errors made using the Lykken scoring system as a function of stimulus mode	15

The method of presenting questions in field polygraph exams has remained relatively unchanged since 1917. Examiners are taught to ask questions in an unemotional tone of voice to be sure it is the content of the question and not the delivery that is associated with any physiological reaction.

The advent of television and personal computers has made presentation of written material on a video screen rather common. There is, however, a dearth of research on the application of this common technology to polygraph testing. Application of visual technology in physiological detection of deception (PDD) has both certain advantages and disadvantages.

No doubt it would increase the cost of apparatus in the field and, until perfected, might be more awkward to use than verbal presentations. However, using a computer to deliver the questions might be a good way of insuring that physiological responses are associated with the content of the questions, and not any intentional or unintentional verbal or nonverbal behavior on the part of the examiner. If this is true, the use of visually presented techniques would take the field of PDD a long way toward standardization. Also, there is little research that examines the accuracy of a polygraph test given to someone with impaired hearing, where visual presentation of the questions may be a necessity.

Lacking conclusive research support, there has been no temptation to adopt visual presentation methods. To-date, only one investigation can be found in current literature which compared the effects of the type of stimulus mode in which the questions are presented.

An investigation by Beijk (1980) attempted to evaluate potential differences found in skin resistance responses as a function of mode of stimulus presentation on a numbers test. A prior experiment found a significant 'hit' rate on a numbers test. A follow-up experiment was conducted to examine different modes of presentation (auditory versus visual) and found no significant difference between visual and auditory presentation of the stimuli. The authors "conclude that a small difference in experimental procedure, be it an attempt to change motivation of the mode of stimulus presentation, did not significantly change the results found in Experiment 1" (p 276).

Beijk used a type of information test (Podlesny and Raskin, 1977). There are several types of information tests. One type of information test that might prove to be useful in the field is the guilty knowledge test or GKT. An information test presumes that a guilty person possesses knowledge or information that an innocent person would not. It is the exposure of this knowledge or information that is associated with the response made during the polygraph exam.

According to Andreassi, the GKT is superior to the more typically used control question technique (CQT), because it is standardized, error rates can be specified with GKT, and researchers believe that it is less vulnerable to faking or the use of countermeasures (Andreassi, 1989).

The purpose of this research is to compare the distributions of decisions obtained when the questions are presented verbally to those rendered when the questions are presented visually on a GKT. Does one mode of presentation result in more accurate decision concerning deception?

Method

Subjects

Twenty-two female and 60 male basic trainees at Fort McClellan, Alabama participated as subjects in this investigation. Due to excessive movement, the data for two of the male subjects were not included in the final analyses. Subjects were, for the most part, in average to excellent health. The age of the subjects ranged from 17 and 33.

Equipment/Apparatus

Subjects' physiological data was recorded using a Coulbourn Skin Conductance Coupler and preamplifier (S71-22). The coupler was set on AC coupling, sensitivity on 1000 mV/micromho, using silver-silver chloride electrodes attached to the palmar side of the index and middle fingers of the subject right hand. The data was collected on a PC Brand 286 with an NEC Multisync monitor using CODAS Software by DATACQ. CODAS is a data acquisition program which digitizes analog information and stores it in a file in the computer, no hard copy is made. After the data has been digitized and stored, the data was printed out on hard copy using a HP LaserJet Series II printer.

The questions presented in the visual condition were presented on a Zenith IBM PC Compatible using Harvard Graphics Software. The questions presented in the aural condition were delivered via a Marantz PMD 221 Portable 3-head Cassette Recorder.

Procedure/Method

Upon arrival at the Institute, subjects were met and briefed on the purpose of this investigation. The purpose and procedure of the study was fully explained to all subjects. Subjects were also given a copy of a justification and explanation sheet. At this time, subjects were asked to read and sign a volunteer affidavit or participation consent form. Copies of the justification/explanation sheet and the volunteer affidavit can be found in Appendix A. The volunteer affidavit informed the subject that his/her participation is solely voluntary. The form specifies that if the subject wishes to discontinue their participation, she/he may do so at any time and no penalty will

be assessed. Due to the specific nature of the exam, no personal or biographical information was required, therefore, the subjects were not asked questions of a personal nature.

All subjects were given a guilty knowledge test. There were five questions and each question had six alternatives or possible answers. The specific questions and alternatives, with the critical item identified, can be found in Appendix B. The questions were presented in the same sequence, as were the alternatives, for all subjects. Subjects were informed that one of the six alternatives was the correct alternative, however, only a guilty person would know which alternative was correct. Prior to each question, the experimenter told the subject what the question would be, but did not go over the alternatives. Subjects were then told that if they were innocent, none of the alternatives would be any more meaningful than the rest, however, if they were guilty then they would know exactly what the correct alternative was and they could expect it to be presented at some point during the recording of the question.

Subjects were instructed not to respond to the question itself, but to wait until they were presented with an alternative. The required response to each alternative was "NO". Since the question began with "Do you know . . . ," an innocent person would never be forced to lie since they would not know which of the alternatives was true.

Subjects were randomly assigned to one of the following four conditions: (1) Aural-Guilty, (2) Aural-Innocent, (3) Visual-Guilty, and (4) Visual-Innocent.

Subjects were randomly programmed innocent or guilty individually. All subjects viewed a short video. Subjects who were programmed guilty viewed a video of a mock crime. The video depicted the theft of a gun and some money. The video was shot from the criminal's perspective, meaning as if the camera person was committing the crime. The criminal's face was never shown. however, the arms were visible at times during the crime and they were easily identified as a man's arms. During the theft, an unwitting victim came upon the crime scene. At this point, the criminal pointed the gun at the victim and fired twice. While making sure the victim was dead, the criminal also stole the victim's wrist watch. After viewing the video, the subjects were then questioned by the investigator concerning the critical elements to insure that the guilty subject indeed had the guilty knowledge prior to the polygraph examination.

Subjects who were programmed innocent were shown a brief training film and asked questions concerning the content afterwards. All subjects were told that the purpose of the polygraph exam is to determine if a polygraph examiner could tell whether or not a subject witnessed the crime based solely on

their physiological activity. All subjects were strictly warned not to inadvertently alert examiner to which video they viewed. The subjects were told that if they did allow the examiner to 'guess' their condition prior to running charts, either verbally (admission) or nonverbally, they would be released from the investigation and returned immediately to their unit.

Once the subjects were programmed, they were taken to the polygraph room and introduced to the examiner. Only one examiner was used to run the polygraph exams. The investigator who programmed the subject remained in the room. Once in the polygraph room, the subject was briefed on what measures were being taken, how a polygraph works, what kind of question would be asked, and how they were expected to respond. The subjects were informed that they would be taking a polygraph, because they were suspected of having been an accomplice during a crime. The components were attached.

Subjects were seated in a typical polygraph chair, outfitted with the elongated arm rests. The subject was seated approximately 1 meter from a computer monitor and 30 cm in front of the Coulbourn equipment. The examiner sat at a computer terminal located next to the Coulbourn equipment and was therefore approximately 1 meter to the left of, but slightly behind, the subject. The arrangement was designed so that movement of the examiner would occur outside of the subjects' peripheral vision.

Subjects in the visual condition were told that the questions would be presented on the screen in front of them, while subjects in the aural condition were told that the questions would be presented via a tape recording. The subjects were given an example question (presented either visually or verbally on the tape recorder) to make sure they understood the instructions. The example question can be found in Appendix B. The example question was unrelated to the crime and the subjects were told this prior to the presentation of the example. The subjects were fully aware that the purpose of the example was to give them a chance to see what the actual testing would be like and to make sure they understood what they were supposed to do.

The visual stimuli were created and presented using Harvard Graphics version 2.0. Each character presented visually was approximately 2 cm in height. Subjects in the visual condition were questioned concerning clarity and those requiring reading glasses were requested to use them if necessary. After the presentation of the last question, subjects were required to read the last alternative out loud to ensure that the subject could see and read the word clearly. Subjects in the aural condition were asked if the volume was acceptable.

Each question was presented once. There were three cases in

which a question was interrupted during recording by the telephone or some one at the lab door. In these cases, the question was stopped immediately and the question was asked a second time. There was approximately 3 minutes between each question, while the examiner informed the examinee what the next question would be. Prior to the presentation of each question, the examiner said, "Please remain still, the test is about to begin." At this point, the data collection program was started and physiological recording began. Simultaneously, either the tape recorder was turned on (aural condition) or the program for the specific question (visual condition) was initiated. After a 20 second pause, the question was presented.

In both the aural and the visual conditions, there were 15 seconds between the presentation of the question and the first alternative, as well as between each subsequent alternative. In the visual condition, the question remained on the computer screen until the first alternative was presented. Each alternative also remained on the screen until the next alternative was presented. After the last alternative was presented in both the aural and visual conditions, there was a 15 second pause until the examiner said, "Now you can relax, this portion of the test is complete." During the aural condition the recorder was then turned off. The program in Harvard Graphics terminates automatically using the sideshow option of presentation.

Upon completion of the polygraph examination, the subjects were taken to another room and asked to fill out a questionnaire. The questionnaire was simply a copy of the GKT questions. A copy of this questionnaire can be found in Appendix C. In the questionnaire subjects were asked to identify the critical items for each question. The purpose of this task was to ensure the following: (a) that no programming mistakes were made; (b) the guilty subjects did remember what the critical items were; and (c) innocent subjects did not identify what the critical items were at a better than chance rate. Since all of the subjects were told not to discuss the nature of the study with anybody, the questionnaire might also reveal an innocent subject who had been given information about the crime from a buddy who served earlier.

The Confronter

A confronter was used to increase the accuracy of the examination. For all subjects, the computer screen was approximately 3 feet directly in front of the subject. All the subjects were told that during the recording they should focus on the computer in front of them. Subjects in the visual condition were told to watch the computer screen so they would not miss the presentation of the questions or alternatives while those in the aural condition were told to focus on the screen to prevent them from becoming distracted and looking about the room. The

investigator who programmed the subjects acted as the confronter. The confronter sat next to the computer screen. Subjects were told to focus on the computer screen while the questions and alternatives were presented but when they had to respond they were to look directly in the eyes of the confronter and say 'NO' just as if the confronter had asked the question.

The rationale behind the use of the confronter was to increase physiological responsivity. By increasing physiological responsibility, one would be more likely to observe differential responding which should, in turn, increase the overall accuracy.

Basically, this strategy should serve to make the guilty subjects more uncomfortable about lying. Perhaps lying to someone who knows you are lying is potentially far more disturbing than the simple act of lying alone. Requiring the examinee to look directly into the eyes of the confronter was designed to make the act of 'lying' a little more uncomfortable for the guilty person.

It may be true that simply looking at a stranger during this process would be uncomfortable for the innocent subjects as well, however, the guilty person also has to lie to a strange person who knows they are lying. It was hoped that this differential anxiogenic procedure would increase the accuracy of detecting the guilty subjects. If this did indeed increase the accuracy for detecting guilt then accuracy in establishing innocence would increase as well.

Previous piloting of this study, using field instruments and regular field polygraph examiners, rendered very poor accuracy, statistically around chance levels. Since the purpose of this study was to compare the accuracy between aural versus visual presentation of questions, it was decided to duplicate the conditions of a previous study conducted earlier in this lab (Richardson, Carlton & Dutton, 1990). This previous study used the same video, virtually identical questions and used a confronter. Since this earlier study obtained a high accuracy rate (76% - 80% for the original examiners) it was decided to include the confronter on this study.

Scoring

The skin conductance data were scored in following fashion: (1) by the original examiner upon completion of the polygraph examination; (2) by a blind evaluator; and (3) using a scoring system introduced by Lykken (1959) devised exclusively for scoring guilty knowledge tests.

(1) Original Examiner. The first author of this report served as the examiner who ran the polygraph test and, therefore, was the original examiner. After the subject was run the data files were printed out to get the hard copy. There were five

questions and each question was called a "chart". Scoring of the charts was subjective. A call of Deception Indicated (DI) or No Deception Indicated (NDI) was made based on these five charts alone. The original examiner used information derived from the electrodermal responses. The following physiological indices were used: (1) amplitude; (2) rise-time; (3) latency changes; (4) changes in frequency of responding.

Of the four indices, the examiner generally placed more weight on the amplitude information. If the largest response on a chart occurred after the presentation of the critical item, the chart was scored a 'hit'. A subject could be called DI if they hit 3/5 keys or more. However, on a few occasions only 2/5 keys were given a 'hit' designation if any or all of the following occurred: (a) rapid decrease in rise-time for response occurring at the key, but not at the other alternatives; (b) shorter latencies for responses occurring at the key and not elsewhere; and (c) the electrodermal activity diminished after the presentation of the key.

(2) Blind Evaluator. A blind evaluator was given information about how the guilty knowledge test was conducted and simply asked to render a decision.

(3) The Lykken Scoring System. The Lykken scoring system uses only the amplitude of the electrodermal responses for scoring purposes. For a given question, the subject's electrodermal responses for the first alternative are discarded while the remaining responses are ranked according to amplitude. If the largest response occurs at the key, the question is given a score of '2'. If the response is the second largest response on the question, the score of '1' is given. Since there are 5 questions, the largest score possible is 10. A subject was classified as deceptive if the total score was 6 or higher. The total score is referred to as a Lykken score.

Results

All of the statistical calculations were conducted using Crunch statistical software.

Questionnaire Results

Analyses were conducted on the questionnaires to address two issues. The first issue was concerned with the accuracy of guilty subjects, that is, to determine if the guilty subjects knew and remembered all of the critical items to each question. The results of the questionnaire showed that all of the guilty subjects correctly identified all of the critical items.

The second issue was to determine if the innocent subjects could correctly identify the critical items. This could occur if the incorrect alternatives were not adequate and the critical

item was too obvious or if the subject was given information about the crime by a buddy who served as a subject earlier in the study. Table 1 shows the probability distribution of correctly guessing the critical items, and the number of the innocent subjects who correctly guessed the specified number of critical items.

Table 1. Probability, frequency and expected frequency distributions of innocent subject currently identifying critical items.

<u># Correct</u>	<u>P(# Correct</u>	<u>N Observed</u>		<u>N Expected</u>
0	.328	16	13	
1	.410	13	16	
2	.205	10	8	
3	.051	1	2	
4	.006	--	--	
5	.0+	--	--	

Table 1 provides probability, observed, and expected frequency distribution for the number of critical items identified by innocent subjects. The number in parentheses is the number of subjects that would correctly identify that number of critical items by chance alone (out of 40).

The table shows that 16 or 40% of the innocent subjects could not correctly identify any of the critical items, 13 subjects (32.5%) correctly identified one critical item, 10 subjects (25%) could correctly identify two critical items and 1 subject (.025%) correctly identified three of the critical items. These two frequency distributions (observed and expected based on chance) are not statistically significantly different ($X^2 = 2.25$; $p = .05$).

A partial item analysis on the correctly chosen critical item for innocent subjects showed that of the 35 correct answers given by innocent subjects, 26% (9) occurred on question 1, 26% (9) occurred on question 2, 31% (11) occurred on question 3, 6% (2) occurred on question 4, and 6% (4) occurred on question 5. These figures can be found in Table 2.

Table 2. Frequency of correctly identified critical items for each question.

<u>Question #</u>	<u># of correct</u>
1	9
2	9
3	11
4	2
5	5

Polygraph Examination Results

The decisions of the two examiners and the Lykken scores were all highly correlated. Table 3 shows the correlation matrix between the three evaluations.

Table 3. Inter-scoring system/evaluator matrix.

	Original Examiner	Blind Evaluator	Lykken Scores
Original Examiner	1.0	.6773	.6652
Blind Evaluator	--	1.0	.6691
Lykken Scores	--	--	1.0

The correlations between the original examiner and the blind evaluator and Lykken scores were .68 and .67, respectively. The correlation between the blind examiner and the Lykken scores was .67. All of the correlations were statistically significant with $p < 0.0001$.

Overall Accuracy

The accuracy levels for the original examiner, blind evaluator, and the Lykken scores are found in Figure 1.

Overall accuracy for the original examiner was 78%. This level of accuracy is highly statistically significant ($\chi^2 = 24.2$; $p < .0001$). The blind evaluator obtained an accuracy of 74%, also highly statistically significant ($\chi^2 = 18.05$; $p < .0001$). The Lykken scores showed an overall accuracy rate of 76%, again highly statistically significant ($\chi^2 = 22.05$, $p < .0001$).

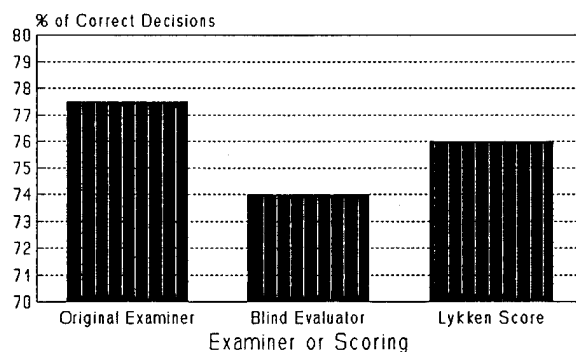


Figure 1.

Role

Figure 2 shows the accuracy levels of the original examiner, blind evaluator, and the Lykken scoring system for both guilty and innocent subjects. It shows that accuracy for the guilty subjects was 80% for the original examiner, 73% for the blind examiner, and 63% for the Lykken scores. Accuracy for the innocent subjects was 75% for both the original examiner and the blind evaluator and 90% for the Lykken scores.

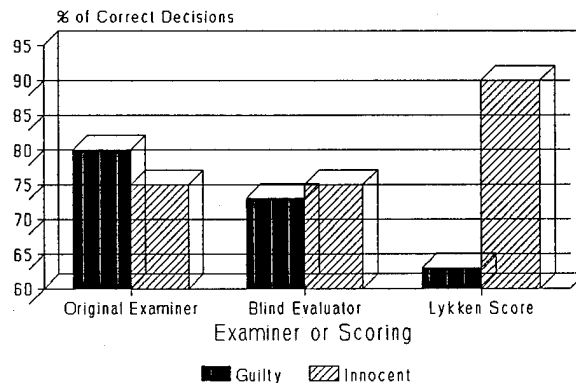


Figure 2.

Tables 4, 5 and 6 provide the χ^2 contingency tables for the decision of the original examiner, blind evaluator and Lykken scores.

Table 4 indicates that there is a significant association between role and the decision of the original examiner ($\chi^2 = 22.1$, $p < 0.0001$) in that 32 of the 40 guilty subjects were correctly identified as DI with only 8 false negative errors (guilty subjects called NDI), while 30 of the 40 innocent subjects were correctly identified as NDI with 10 false positive errors (innocent subjects called DI).

Table 4. Contingency table for role versus decision of the original examiner.

	DI	NDI	Total
Guilty	32	8	40
Innocent	10	30	40
Total	42	38	80

Table 5 indicates that there is a significant association between role and the decision on the blind examiner ($\chi^2 = 16.21$, $p < 0.0001$) in that 29 of the guilty subjects were correctly identified as DI with 11 false negatives and 30 innocent subjects were correctly identified as NDI with 10 false positive errors.

Table 5. Contingency table for role versus decision of the blind evaluator.

	DI	NDI	Total
Guilty	29	11	40
Innocent	10	30	40
Total	39	41	80

Table 6 indicates that there is a significant association between role and the decision made using the Lykken scoring system ($\chi^2 = 21.64$, $p < 0.0001$) in that 25 of the guilty subjects were correctly identified as DI with 15 false negative errors and 36 of the innocent subjects were correctly identified as NDI with 4 false positive errors.

Table 6. Contingency table for role versus decision of the Lykken scoring system.

	DI	NDI	Total
Guilty	25	15	40
Innocent	4	36	40
Total	29	51	80

Stimulus Mode

The accuracy levels for the original examiner, blind evaluator, and the Lykken system for the visual and the aural conditions are found in Figure 3.

Accuracy for subjects in the visual condition was 83% for the original examiner, 78% for the blind evaluator, and 70% for the Lykken scores. In the aural condition, accuracy rates were 73%, 70%, and 83% for the original examiner, blind evaluator, and Lykken scores, respectively.

To compare the stimulus modes, one way to organize such a comparison is compare stimulus mode on correct decisions and stimulus mode on errors. The first analysis indicates whether or not the types of correct calls are distributed differently by stimulus mode. The second analysis examines whether or not the types of errors are distributed differently for the two stimulus modes.

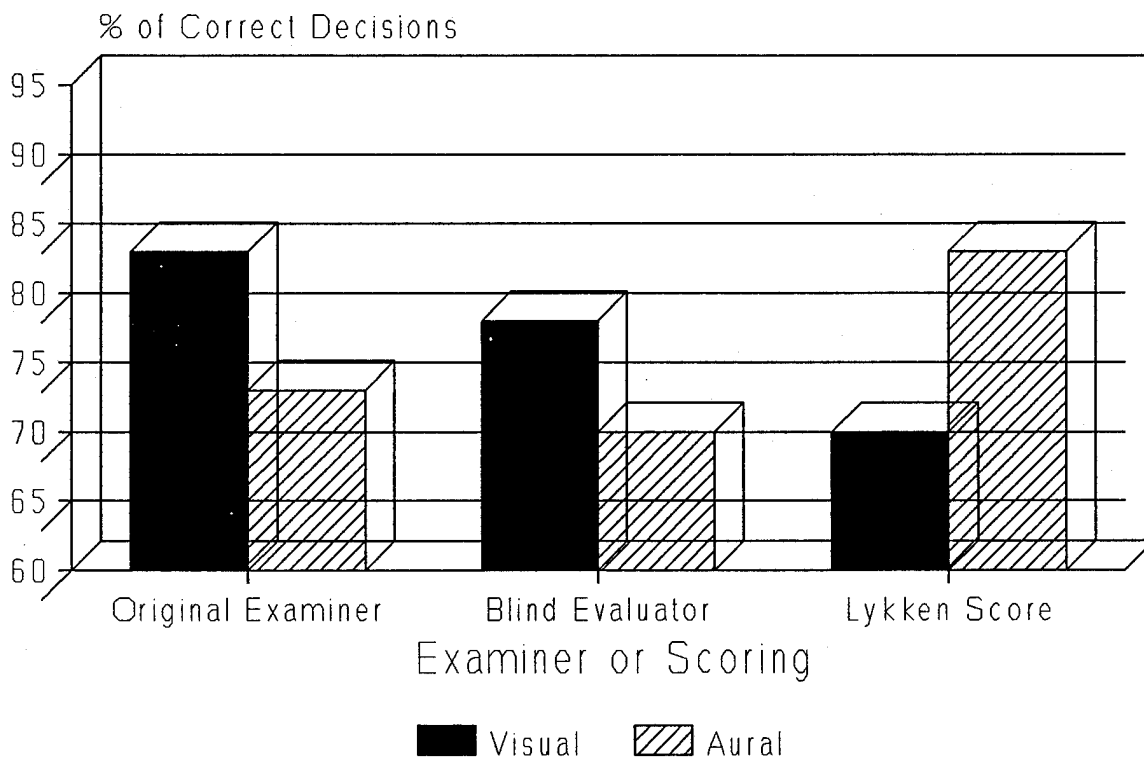


Figure 3

Distribution of correct calls as a function of stimulus mode

A decision x stimulus mode chi-square statistic was calculated on correct decisions for the original examiner, blind evaluator and Lykken score. The χ^2 contingency tables for these analyses can be found in Tables 7, 8, and 9, respectively. No significant associations were found between the type correct decision and the stimulus mode of question presentation for either and original examiner or the blind evaluator, or the Lykken scores on accuracy of decision ($\chi^2 = 0.6091$, $p < 0.4351$); $\chi^2 = 2.0378$; $p < 0.1534$; $\chi^2 = 1.0651$, $p < .3020$).

Table 7. Distribution of the correct original examiner decisions as a function of stimulus mode.

	True Negative	True Positive	Total
Aural	12	17	29
Visual	18	15	33
Total	30	32	62

Table 8. Distribution of the correct blind evaluator decisions as a function of stimulus mode.

	True Negative	True Positive	Total
Aural	11	17	28
Visual	19	12	31
Total	30	29	59

Table 9. Distribution of the correct Lykken scoring decisions as a function of stimulus mode.

	True Negative	True Positive	Total
Aural	17	16	33
Visual	19	9	28
Total	36	25	66

Distribution of error-type as a function of stimulus mode

Due to much smaller expected frequencies per cell, the association between the type of error in decisions and the stimulus mode was calculated using a Fisher's exact test. The contingency tables for error-type by stimulus mode for original examiner, blind evaluator, and the lykken scores are found in Tables 10, 11 and 12, respectively.

Table 10. Distribution of the errors made by the original examiner as a function of stimulus mode.

	False Negative	False Positive	Total
Aural	3	8	11
Visual	5	2	7
Total	8	10	18

No significant association was found between the role of subject and the stimulus mode of presentation for the original examiner or the Lykken scores on type of error (Fisher's exact test, two-tailed, $P_2 = 0.1448$; $P_2 = .2451$, respectively). There was a significant association found between the role of the subject and stimulus mode on error type for the blind evaluator (Fisher's exact test, two-tailed, $P_2 = p < .001$).

Table 11. Distribution of the errors made by the blind examiner as a function of stimulus mode.

	False Negative	False Positive	Total
Aural	3	9	11
Visual	8	1	9
Total	11	10	21

Table 12. Distribution of the errors made using the Lykken scoring system as a function of stimulus mode.

	False Negative	False Positive	Total
Aural	4	3	7
Visual	11	1	12
Total	15	4	19

Discussion

The stimulus mode in which the questions are presented appears to have very little influence on the rate of detection of the GKT. This was true for both subjective decisions of the original examiner and the blind evaluator as well as the more objective scoring system described by Lykken, when examining the accuracy of the decisions. These results support the earlier finding of Beijk, (1980).

It appears that the stimulus mode in which the question is presented also has little effect on the type of error in decision that is made at least for the original examiner and the more objective Lykken scoring system. The finding of a significant association between the type of error in decision and stimulus mode for the blind evaluator is somewhat puzzling. It is interesting to note that more false positive errors were made for subjects in the aural condition than in the visual condition. This relationship is reversed for false negative errors. More false negative errors were made for subjects in the visual condition than in the aural condition (See Tables 7, 8 and 9). This distribution of errors was found for all of the scores from both the original examiner, the blind evaluator and the Lykken scores, however, the association was significant for the blind evaluator alone. Perhaps with a larger sample size this distribution might be significant for the original examiner and the Lykken system. There are a couple of possible explanations for this result.

It is possible that there is a type of confronter effect. The confronter sat next to the computer during the polygraph examination. Therefore, she could not see each alternative as it was presented. She was aware of the presentation of each alternative by the click sound of the event marker used by the examiner, but she could not see which alternative was presented. However, in the aural condition, the confronter could hear each alternative as it was presented. It is possible that the confronter inadvertently reacted when the critical item was presented. If the confronter did react strongly enough for the subject to respond this would only have affected innocent people in the aural condition as the confronter would not have known (for all subjects and all questions) when the critical item was presented. So one possibility is that the confronter somehow elicited a larger response from innocent subjects when the critical item was presented in the aural condition.

This does not explain why there are more false negatives in the visual condition than in the aural condition, unless one makes a couple of assumptions about how the confronter affects the subjects. Perhaps the important element is that the confronter must know the following to have any effect: (a) that the subject is lying and (b) exactly when the subject is lying.

During the visual condition even though the confronter knew the subject would be lying, she was unaware of the exact moment that the subject was lying.

Another possible explanation for the higher false positive rate in the aural condition could be that the inflection in the voice of the person asking the questions could have caused the reactions. The tape of the questions was made by the examiner who ran the polygraph examination. Therefore, when the questions were being recorded, the examiner may have accidentally, through some tone or inflection, made the critical item more salient such that an innocent person could detect the difference. However, this is not supported by the questionnaire data.

The results of the questionnaire data indicate that innocent subjects were not aware of the critical items at the time the questionnaire was given to them after the exam. The distribution of correctly guessed critical item was not statistically different from what would be predicted from chance alone. This would mean if the confronter has any effect on the innocent subjects in the aural condition, the subject was unaware or not conscious of the effect. The innocent subjects in the aural condition did not know or learn what the critical item was in the questions, and, therefore, the reasons underlying false positive errors are unknown.

Table 2 provides a distribution of the number of times innocent subjects correctly chose the critical item for all of the questions. Although it is apparent that questions 1 through 3 were more often correctly guessed than were 4 and 5, this does not provide much insight to the problem. To examine whether or not this distribution is unusual would require a complete item analysis of the questionnaire data. The purpose of the questionnaire data was to ensure that the guilty subjects could correctly identify the critical items and that the innocent subjects could not do so at a better than chance level. Both of these assurances were maintained.

The question of intonation is an empirical question. However, it is a question that this investigation was not designed to answer. Given that accuracy was not significantly better for aural versus visual presentation, clearly a way to negate the debate is to rely on more visual presentations during polygraph exams.

In spite of the results concerning the types of errors found in this study, the fact remains that there was no significant association between the stimulus mode of question presentation on accuracy. This interpretation does support a greater role for visual stimuli in the polygraph test. In spite of this, subsequent research must address the potential differences found

in error type before questions may be presented visually during a polygraph exam.

An interesting observation gleaned from the results are the differences between the two subjective scoring systems and the more objective scoring system proposed by Lykken. It should be pointed out that the Lykken system is objective only in that it uses amplitude as the scoring criterion and attempts to apply a numerical scoring system. However, the cut-off point is arbitrary. Perhaps manipulating different cut-offs for the scores would prove to be a very informative exercise and should be done in subsequent research.

In this study the cut-off score of 6 resulted in a very high false negative rate. This is consistent with what is generally assumed about the GKT. Due to the probabilities involved, it is reasonable to assume that most of the errors should be false negatives. It should be very difficult to reach a false positive result due to chance alone. This investigation would support this notion as there were only 4 false positives and 15 false negatives when using the scoring system developed by Lykken.

Lastly, another interesting result of this investigation is the confronter issue. Although no firm conclusions may be stated, it is curious that the pilot studies for this investigation rendered very poor results (around chance) when using field polygraphs and field polygraph examiners. The decision was then made to use the Coulbourn equipment with one examiner and the confronter. After this decision was made, the accuracy for the investigation increased dramatically with overall accuracies ranging between 74% to 78%.

It is difficult to maintain that the equipment alone is responsible for this increase in accuracy. It is possible that the conductance recordings from the Coulbourn coupler were superior to the resistance recordings on the field polygraphs. Since this variable was not included in the design or even manipulated, no conclusion on this issue may be reached.

It is also possible that changing from multiple examiners to one examiner also played some role in the increase in accuracy. Even though the base rate of 50/50 was common knowledge to all four examiners, that did not necessarily relate to the base rate for any one examiner. There was no attempt to ensure that all of the examiners were given equal numbers of innocent and guilty subjects. This would have violated the random assignment to conditions since the schedules of the examiners varied from day-to-day and week to week.

Another consideration related to multiple examiners is that the examiners used during the pilot phase were all federally licensed polygraph examiners with no experience running GKTs in

the field. The examiner who ran the GKT for this study is not a polygraph examiner, but does have some experience with a GKT in laboratory situations. Perhaps the more experienced examiners maintained a peak of tension bias as that is a technique they are familiar with and is most similar to the GKT that somehow interfered with the running of the GKT.

A related possibility is that even though all of the examiners were given scripts to follow for the pre-test and testing, simply by virtue of differences in experience in the field, the examiners would not necessarily handle the subjects in the same way. The switch from several examiners to one examiner would eliminate any differences due to variability between examiners. However, if this is true, there are certain implications on accuracy in the field, where there is no attempt to require examiners to treat all suspects the same and the base rates also vary by examiner. This would mean that overall accuracy in the field would suffer simply due to differences between examiners.

That would leave the confronter issue as a primary candidate for explaining the differences in accuracy rates. How the use of the confronter increases accuracy is an empirical question. One possible explanation is that it increases the accuracy of detecting the guilty subjects simply by making the subject more uncomfortable during a lie. It is logical that if accuracy improves for the guilty subjects, the accuracy for the innocent would also improve.

The confronter issue is certainly one that should be addressed in subsequent research. This issue could affect many aspects of physiological detection of deception. It has ramifications on future research, both theoretical and applied, as well as on how examinations may be conducted in the future.

REFERENCES

- Andreassi, J. L. (1989). Psychophysiology, 2nd Edition, Lawrence Erlbaum Associates: New Jersey.
- Beijk, J., (1980). Experimental and procedural influences on differential electrodermal activity. Psychophysiology, 17(3), 274 - 278.
- CODAS Software by DATACQ Inc., Release Level 4, 825 Sweitzer Ave., Akron, Ohio 44311.
- Harvard Graphics. SPC Software Publishing Corp., Version 2.0. 1901 Landings Drive, Mountain View, CA 94039-7210.
- Lykken, D. T. (1959). The GSR in the detection of guilt. Journal of Applied Psychology, 43(6), 385 - 388.
- Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. Psychological Bulletin, 84, 782-799.
- Richardson, D. C., Carlton, B. L., & Dutton, D. (1990). Assessingsystolic time intervals. Presented at the annual meeting of American Psychological Society, June 1990, Dallas, TX.

Appendix A

TO YOU, THE SUBJECT:

Welcome to the Department of Defense Polygraph Institute. This may be the first time you have been at the institute so we would like to provide you with some information concerning the purpose for being here today. We hope that you will enjoy the task we will give you today. We will not ask you any embarrassing questions or make you do anything that you are uncomfortable doing. Your participation is completely voluntary. If you have any questions, please feel free to ask the investigator who greets you today.

PART A -- EXPLANATION

1. **PROJECT TITLE:** The Effects of Aural Versus Visual Presentation of questions during a Detection of deception Task. This project is being conducted at the Department of Defense Polygraph Institute. Fort McClellan, Alabama.

2. **PRINCIPAL INVESTIGATOR:** Barbara L. Carlton, Ph. D.

3. **DISCUSSION:** Congress has directed the Department of Defense to conduct research to determine the effectiveness of the polygraph. You are being asked to volunteer to a polygraph test in an experiment. The question of interest to us is "Does it make a difference in the physiological responses if the questions are presented orally by a polygraph examiner or visually on a computer screen?"

Some of you will witness a crime before taking the polygraph test. You will witness this crime on a video tape. By showing you this crime we are giving you guilty knowledge, that is, information that you would not know unless you either witnessed or committed the crime. Having this information makes you guilty, just as if you had committed the crime yourself. Others of you will be shown a brief training video and will not know anything about the crime, that is, you will not have any guilty knowledge so you will be innocent. During the polygraph test you will be asked a series of questions about certain events that pertain to the crime. After the question is asked it will be followed by a series of possible answers. You will be instructed to respond to each of the possible answers with 'no'. An example of such a question and the alternatives (and required responses) would be:

"Do you know what color of shirt the person who committed the crime was wearing? Was it ... green? (NO); red? (NO); blue? (NO); white? (NO); yellow? (NO)"

This means that those of you who did, in fact, witness the crime will be lying. If the person was wearing a red shirt you

will say 'no' to this alternative as well as to all of the other alternatives. Only those who witness the crime will have the guilty knowledge and will know that s/he is lying to the correct alternative. Those who did not see the video will not be lying as they will not know which alternative is correct.

The polygraph examiner who will be testing you is an experienced member of the research department here at DODPI. The examiner will not know who is lying on the exam until after all the polygraph tests have been completed. It will be her job to give you a polygraph examination and, based on your responses, determine whether or not you have guilty knowledge.

Our legal system assumes innocence, therefore, the examiner will assume you are innocent. However, the examiner also knows some of you are guilty, that you have witnessed the crime. She will be looking for subjects who are guilty. You must maintain your innocence and at no time tell the examiner whether or not you have witnessed the crime. This is how you can best help us test the polygraph technique.

4. **DISCOMFORTS:** Some people find it difficult to sit still for several minutes at a time during the polygraph test, while psychophysiological measurements are being made from the body. Including programming and instruction, the total length of time required for your participation in this investigation will be approximately 1 and 1/2 hours, however, you will be here at the Institute for the entire day.

5. **RISKS:** There are few, if any, risks involved in this study.

6. **BENEFITS:** The main benefit is the satisfaction of knowing you contributed to a scientific study, important to national security. We need to know the accuracy of polygraph techniques in detecting deception and how to improve the techniques. An additional benefit to will come from the experience of having taken a polygraph exam. Although this kind of exam is not used in the field in this form, it may, however, be used in the future. You may be exposed to the polygraph for security clearance you will need. Having had prior experience will be a benefit to you, as you will know what to expect if you ever take one in the future.

7. **CONFIDENTIALITY OF RECORDS:** Since this is not a polygraph exam technique that requires you to give any information at all, no information will be requested. You will not be asked any questions of a personal nature or about previous criminal activity. All of the polygraph charts, score sheets, interview forms, examiner work sheets, and related documents associated with your examination will be used for research purposes only. Members of the Army Surgeon General's Human Subjects Research Review Board may inspect the records of the research in their capacity as reviewing officials, but your identity will be kept confidential.

8. **YOUR RIGHTS:** You have the right to ask any questions about any aspect of your participation in the study. If any problems arise at any time in conjunction with your involvement in the study, or if you have been injured in any way as a result of the study, the person to contact is the chief of the research division of the Defense Polygraph Institute. In the event that you do have questions or any of the above has occurred please contact Dr. Barland at (205) 848-4952. Should any question arise concerning study-related injury, you may contact COL Hegstrom, M. D., Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, telephone number (205) 848-2200.

9. **VOLUNTARY PARTICIPATION:** Your participation in this study is completely voluntary. If you would prefer not to participate, do not volunteer for it! Even if you decide to participate in the study, you may discontinue at any time without penalty or loss of benefits to which you are entitled. Should you decide not to participate please inform someone on the staff at the Defense Polygraph Institute, or if it occurs during the polygraph examination itself, inform the examiner and you will be released and returned to your unit.

10. **ADDITIONAL COMMENTS** If you have been given guilty knowledge (you are a witness to the crime) it is very important that you never tell that to the examiner. If you tell him that at any time during the polygraph test, you will be withdrawn from the study and returned to your unit.

VOLUNTEER AFFIDAVIT

This form is affected by the by the Privacy Act of 1974 .

1. AUTHORITY; 10 USC 3012, 44 USC 3101 AND 10 USC 1071.1087.
2. PRINCIPAL PURPOSE: To document voluntary participation in the Defense Polygraph Institute Research Program. Your name will be used for identification.
3. ROUTINE USES: The name will be used for identification and locating purposes. Information may be furnished to Federal, State, and local agencies.
4. MANDATORY OR VOLUNTARY DISCLOSURE: Your signature is necessary if you want to be included in this research. If you do not sign, you will not be able to serve in this study and you will be returned immediately to your Unit.

PERSONAL STATEMENT

I, _____, being at least 17 years old, do hereby volunteer to participate in a research study entitled "The effects of Aural versus Visual Presentation of Questions during a Detection of Deception Task", being conducted by the Department of Defense Polygraph Institute at Fort McClellan under the direction of Barbara L. Carlton, Ph. D.

The implication of my participation, the nature, duration and purpose, and the methods by which it is to be conducted: and the inconveniences and hazards to be expected have been thoroughly explained to be as described above. I have been given the opportunity to ask questions concerning this study, and any such question has been answered to my satisfaction. Should any question arise concerning my rights relating to study-related injury, I should contact COL. Hegstrom, M. D., Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, telephone number (205) 848-2200.

SIGNATURE

DATE

Print your name here

Permanent Address (street, city, state and zip)

WITNESS

Witness Name Printed

Appendix B

Examination Questions and Alternatives

GKT QUESTIONS

1. Do you know how entry was gained into the building?
Was it . . .

- a. Climbing through an open window?
- b. Entering an unlocked door?
- c. Crowbarring the door?
- d. Breaking the window?
- e. **Cutting the padlock on the door?**
- f. Climbing through an attic vent?

2. Do you know what the sign read on the door to the room that was entered? Was it . . .

- a. Cashier?
- b. Receptionist?
- c. Director?
- d. **Paymaster?**
- e. Supply?
- f. Secretary?

3. Do you know how the victim was killed? Was it . . .

- a. Choked with a scarf?
- b. **shot with a pistol?**
- c. Stabbed with a knife?
- d. Struck over the head?
- e. Drowned in the bath tub?
- f. Hit with a car?

4. Do you know what was removed from the body? Was it . . .

- a. Money?
- b. Dog Tags?
- c. **Watch?**
- d. Pocket knife?
- e. Ring?
- f. Keys?

5. In the room entered, there were two boxes with names on them.
Do you know what name was on the bottom box? Was it . . .

- a. William?
- b. Raymond
- c. Gordon
- d. **Charles**
- e. Matthew
- f. Steve

The Critical Item is in **bold** print.

Example of GKT Question Given to All Subjects

Do you know what kind of shoes that man was wearing? Were they....

- a. Tennis Shoes?
- b. Combat Boots?
- c. Loafers?
- d. Hiking Shoes?
- e. Dress Shoes?

The question was given to subjects via tape recording (aural condition) or a computer monitor (visual condition). Subjects were requested to respond to the alternatives just as if it was an actual test question. This question was not significant to any of the subjects. The question was not related to the mock crime witnessed by the guilty subjects and the subjects were informed of this fact.

Appendix C

RESEARCH QUESTIONNAIRE

The following questions relate to a crime that you may have witnessed. Please answer each question to the best of your ability. If you witnessed the crime you will know what the correct answer is for each question. If you have not witnessed the crime all you need to do is simply guess for each question.

1. How was entry gained into the building?
 - a. Climbing through an open window?
 - b. Entering an unlocked door?
 - c. Crowbarring the door?
 - d. Breaking the window?
 - e. Cutting the padlock on the door?
 - f. Climbing through an attic vent?
2. On the door to the room that was entered, did the sign read:
 - a. Cashier?
 - b. Receptionist?
 - c. Director?
 - d. Paymaster?
 - e. Supply?
 - f. Secretary?
3. How was the victim killed:
 - a. Choked with a scarf?
 - b. Shot with a pistol?
 - c. Stabbed with a knife?
 - d. Struck over the head?
 - e. Drowned in a bath tub?
 - f. Hit with a car?
4. What was removed from the body?
 - a. Money?
 - b. Dog Tags?
 - c. Watch?
 - d. Pocket Knife?
 - e. Ring?
 - f. Keys?

5. In the room entered, there were two boxes with names on them.
What was the name on the bottom box?

- a. William
- b. Raymond
- c. Gordon
- d. Charles
- e. Matthew
- f. Steve