This is to certify that the

thesis entitled

ANALYSIS OF TWO VARIATIONS OF CONTROL QUESTION POLYGRAPH TESTING UTILIZING EXCLUSIVE AND NONEXCLUSIVE CONTROLS

presented by

John Joseph Palmatier

has been accepted towards fulfillment
of the requirements for

MS degree in Crim. Justice

MASTER OF SCIENCE

Major professor

Date May 16,1991

ANALYSIS OF TWO VARIATIONS OF CONTROL QUESTION POLYGRAPH TESTING UTILIZING EXCLUSIVE AND NONEXCLUSIVE CONTROLS

By

John Joseph Palmatier

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

School of Criminal Justice

1991

ABSTRACT

ANALYSIS OF TWO VARIATIONS OF CONTROL QUESTION POLYGRAPH TESTING UTILIZING EXCLUSIVE AND NONEXCLUSIVE CONTROLS

By

John Joseph Palmatier

Two variations of control question (CQ) polygraph testing, ZOC and MGQT, and two types of control questions, exclusive and nonexclusive, were studied. Subjects were equally assigned (male=80; female=40) to treatments as innocent or guilty of a mock theft. The theft was denied by subjects during polygraph testing. Results showed the accuracy of blind evaluator decisions classifying subjects as guilty or innocent wasn't significantly related to the variation of CQ test used, gender, or subject status ($p > .05$). The accuracy of decisions was significantly related to the type of control question used ($p = .0014$). Nonexclusive control questions were superior to exclusive controls, the results across all treatments were respectively 77% and 45% correct, 15% and 29% wrong, and 8% and 26% inconclusive decisions. Exclusive control questions gave a higher proportion of false positive error compared to nonexclusive controls, respectively the results were 22% and 7%. When inconclusive results were excluded the proportion of error was 30% and 7% respectively.

ACKNOWLEDGMENTS

This study was made possible, in part, by a graduate research assistantship and additional research support funding from the American Polygraph Association's Research Center in the School of Criminal Justice at Michigan State University. All statements, opinions and conclusions are those of the author and do not necessarily represent the opinions or official position of the staff of the APA Research Center or the American Polygraph Association and its members.

I am indebted to Mr. Max Wastl, Jr., Mr. Robert Pritchard and Mr. Chris Fausett, all of the Lafayette Instrument Company, Inc., Lafayette, Indiana, for their helpful advice and agreement to provide the instrumentation used in this research. I also owe a debt of gratitude to the members of my committee for their guidance in the completion of this thesis and advice regarding my graduate education. To Dr. Frank Horvath, my mentor, who allowed me to discover that the distance between the practitioner and the academic was far greater than I had perceived, but has continued to provide direction and to support me in my quest. To Dr. Vincent Hoffman and Dr. Jay Siegel, for their insightful review of this project and their willingness to be involved. I also wish to thank the Michigan Department of State Police, who

gave me the opportunity to become involved with the polygraph field, and to the Royal Canadian Mounted Police, who provided my introductory training.

Finally, but most importantly, I must express my deepest gratitude to my wife, Emiko, and daughter, Nikkole. Without their understanding and support, this journey I have undertaken would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE

CHAPTER I

Introduction

The relative accuracy of various field polygraph testing methods and the effects of altering them has stimulated increased empirical review (Barland, 1988; Driscoll, Honts & Jones, 1987; Forman & McCauley, 1986; Horvath, 1988; Lykken, 1981; Podlesny & Raskin, 1978). Many conclusions have been tendered based on the results of these endeavors. However, examination of one issue often results in the discovery of issues that are equally important. Consequently, additional research is required to collect objective data from which tenable explanations are developed. The present study is a product of this process.

Researchers have conducted a number of studies examining field polygraph testing methods in laboratory settings using mock-crime scenarios (Office of Technology Assessment [OTA], 1983). Laboratory based studies offer certain distinct advantages for research when compared to studies conducted in a field setting (McGrath, Martin, & Kulka, 1982). These advantages include, but are not limited to, precision in the control, measurement, and identification of various factors which may influence the classification of a subject as either innocent or guilty. For example, the type of mock-crime

committed by subjects assigned to a guilty condition, the physiological parameters recorded, the field polygraph testing method(s) employed, and the effect of different rewards and punishments can be manipulated and examined relative to ground truth, the knowledge of who is actually innocent or guilty. In field settings these factors are cited as difficult, if not impossible, to control (Barland, 1988; OTA, 1983; Podlesny & Raskin, 1977; Raskin, 1988).

Today the most frequently used field polygraph testing technique is the control question (CQ) test (Horvath, 1988; Honts, 1986; Raskin, 1982; Reid & Inbau, 1977). Accordingly, CQ testing in one form or another has been subjected to empirical review on more occasions than any other polygraph procedure (OTA, 1983). Two previous studies, one conducted by Podlesny and Raskin (1978), the other by Horvath (1988), used different variations of CQ test to examine in part the relative merits of two forms of control question, one of the primary types of question used for this procedure, within the same context. The results of Horvath's (1988) study and the conclusions he drew were quite different as compared to the results and conclusions drawn by Podlesny and Raskin (1978). Horvath (1988) said: "it is difficult to reconcile these findings. . . . that additional research of this topic is warranted" (p. 207-208). After examining both studies one difference between them was obvious, Horvath (1988) had used one variation of CQ test to examine the control question issue, while Podlesny and Raskin (1978) had used another

variation of CQ test. Even though these two variations of CQ test have seen use in other studies, their relative merits had never been examined within the same context. Accordingly, the purpose of this study was to evaluate in part, within the same setting, the comparative accuracy of the two variations of control question test and the two forms of control question examined in these earlier studies (Horvath, 1988: Podlesny & Raskin, 1978).

To enhance the reader's understanding of this subject, the introduction first presents an explanation of control question test theory. Next a description of numerical scoring, a method generally taught practitioners for analysis of recorded physiological data, is presented. This is followed by a description of the two variations of CQ testing and the two kinds of control questions that were examined. The introduction is then concluded with the purpose of this study. In the second chapter abstracts are presented and a brief review of the literature is given to show the findings of laboratory studies examining issues that are relevant to the validity and reliability of CQ testing and other issues relevant to field polygraph testing. A rationale for the present study concludes the second chapter. In the third chapter the methods used to conduct this study are presented. The forth chapter is used to present the results of the statistical analyses. The project is then concluded with the last chapter offering a discussion of the results and their significance.

Control Question Test Theory

There are basically three types of questions used for CQ testing, these are irrelevant, relevant and control questions. Irrelevant questions are simple questions of fact that are obviously true to both the examiner and the examinee. These questions, are "used as buffers and to establish norms" (Horvath, 1988, p. 198). For example, "Are you now in the United States?" or "Are you over 17 years of age?" Responses to them are not used per se to form opinions; responses to relevant and control questions are used for that purpose.

Relevant questions relate to the matter under investigation (Reid & Inbau, 1977) asking an examinee did s/he do something, does s/he know who is responsible or does s/he have any information that could facilitate the investigation. For example, in the theft of a money bag a relevant question could be "Did you steal that bag of money?"

Control questions lack the specificity of relevant questions and relate instead to "the motive for the offense" (Horvath, 1988, p. 198) under investigation. The form these questions take is a product of the interaction between the examiner and the examinee (Horvath, 1988). Each control question is constructed in such a manner that an answer of "no" is elicited from the examinee in response to it. In so doing the examinee is led to lie or at least doubt the veracity of his/her answer to the question. An example of a control question used for the theft of the money bag mentioned

above could be: "Did you ever steal anything?" or "Besides what you have told me, did you ever steal anything else?"

The premise on which control and relevant questions are based is the assumption that a person will fear that which holds the greatest immediate threat to his/her well being. For a person who has committed a crime the most immediate threat, and hence concern, attaches to answers given in reply to relevant questions. Consequently, a truthful person tested regarding a crime they did not commit should experience little concern when answering relevant questions honestly, but at the same time experience more concern for the accuracy of an answer given in response to the more general control questions (Horvath, 1988).

Control questions were developed by John Reid to be used in "a completely revised polygraph technique," (Reid & Inbau, 1977, p. 3) the control question (CQ) test. The advent of CQ testing is believed by field practitioners to be one of the most important contributions made to polygraph testing, because it allows an objective comparison between the responses evoked by asking relevant and control questions. The greatest benefit derived from using control questions is the increased accuracy they provide due to a reduction in the incidence of false positive errors, the incorrect classification of innocent (truthful) people as guilty (deceptive) (Horvath, 1988). One advocate of CQ testing (Raskin, 1979) theorizes that the accuracy of the procedure is a result of the increased signal value control questions would

have for the innocent by giving him/her an alternative on which to focus their attention, instead of the relevant or irrelevant questions. It's believed that when an innocent person is asked a control question it will elicit greater or at least equal physiological responses when compared to those elicited by the asking of a relevant question.

To formulate an opinion of innocent (truthful) or guilty (deceptive) CQ testing uses a comparative analysis of a person's recorded physiological responses to the control and relevant questions. If a person shows "more pronounced and more consistent responses to control questions than to relevant questions" (Horvath, 1988, p.198) a decision of truthful is rendered. On the other hand if greater physiological responses occur consistently to relevant questions then a decision of deceptive is given. If there is little difference in the magnitude of the physiological responses to both control and relevant questions then a decision of inconclusive is proffered.

Numerical Scoring

To accurately assess the status of an examinee as guilty or innocent requires that an examiner closely analyze the physiological data elicited from an examinee in response to the asking of control and relevant questions. To facilitate this process a semi-objective numerical scoring procedure created by Backster (1965) is often employed. Today, all

schools accredited by the American Polygraph Association to train examiners in field polygraph testing teach some form of this procedure (Honts & Driscoll, 1989). Using numerical scoring an examiner makes a relative evaluation of the response data recorded for each physiological measure in a control/relevant pair (Horvath, 1988) and then assigns a score to the greater response. When a response to a control question is seen as greater a positive score is given; but if the greater response is to a relevant question then a negative score is assigned. A scale of 0 to 3 is commonly used to assess the difference in magnitude of responses. If the difference in magnitude is large a score of 3 is assigned, whereas a score of 0 would indicate no difference or about equal magnitude. A comparison is made for each control/relevant pair, for each physiological measure, and a score assigned. This process is repeated for each of the three charts usually administered during a polygraph examination. The scores assigned each physiological measure on each question are then added together to give a total score for that particular question. The question total score for each question, on each of three charts, is then summed to yield an examination question total score. Next, on each of the charts administered, the scores across each physiological component are summed to yield a component total score. The component total score on each chart is then added to give a component examination total score. Finally, an examination total score is calculated by adding the component examination

total scores together. A simple check on addition is accomplished by adding together the examination question total scores, this number should equal the examination total score (see Appendices J and K). Decisions regarding a subject's status are then based on that score. The cutoff at which a subject is classified as innocent, guilty or inconclusive, may vary. For example, cutoffs of $\pm$ 6 (Podlesny & Raskin, 1978; Raskin & Hare, 1978; Honts, Hodes & Raskin, 1985) and $\pm$ 5 (Barland & Raskin, 1975; Horvath, 1988) have both been used for laboratory research.

The Zone Comparison Test

In field settings there are two major variations of CQ testing that are commonly used, these are the Zone Comparison test (ZOC) and the Modified General Question test (MGQT) (Barland, 1983; Horvath, 1988; Lykken, 1981; Reid & Inbau, 1977) A review of the literature (OTA, 1983), however, shows that the ZOC test has dominated the research environment. This variation of control question test was developed by Cleve Backster (1965) who incorporated changes that he believed would improve upon Reid's CQ test. Backster's (1965) modifications included the addition of two new types of questions, "symptomatic" and "sacrifice relevant" questions.

The symptomatic question was created by Backster (1965) because he believed "It is very important that the examiner has the subject's confidence regarding the avoidance of

questions relating to outside issues. . . ." (p. 11). If an examinee was worried about an outside issue Backster (1965) states this could "dampen" or "eliminate" physiological responses. For example, assume a fictional subject guilty of many crimes is caught stealing cigarettes from a market. Nevertheless, the subject claims the store owner is mistaking him for someone else and volunteers to submit to a polygraph examination to demonstrate the veracity of his denial. However, at the time of the polygraph examination the polygraph examiner isn't aware that this subject is wanted for a bank robbery in another state. Consequently, if the polygraph examiner chose to use control questions about stealing or committing other crimes, they could have more signal value for the subject than relevant questions about the stolen cigarettes. Accordingly, the subject may respond more to control questions, thus leading to a false negative error, a guilty person called innocent. A symptomatic question is believed to guard against such events by bringing forward salient issues the examinee has not discussed with the examiner. An example of a symptomatic question is; "Are you afraid I'll ask you about an issue we have not reviewed."

A sacrifice relevant question is described by Backster (1965) as "an 'icebreaker' question confining itself to the same target issue" (p. 10) as the other relevant questions. This type of question is supposed to provide an examinee with a buffer by asking him/her a question that appears to be relevant, but has no adverse consequences for them should it

elicit a large physiological response. A sacrifice relevant question, generally, isn't scored or used by examiners to formulate opinions. For example, the test about the cigarettes might use a sacrifice relevant that asks; "Will you answer truthfully each question about the stolen cigarettes?"

A ZOC test usually employs two or three relevant questions, each of them rephrased slightly to deal with a narrowly focused issue (Backster, 1969; Barland & Raskin, 1973, 1975). For example, questions for the cigarette theft could be; "Did you steal the cigarettes from the store?" "Are you the person who stole those cigarettes?" "Were the cigarettes stolen by you?" To formulate an opinion, an analysis is conducted by pairing each relevant question with the control question that was asked prior to it, creating a "zone" for comparison. The same questions are asked on each of the three tests ordinarily administered during an examination. However, the question sequence is rotated for each test. This is done so that after the three tests are completed each control question will have been asked adjacent to each relevant question.

The Modified General Question Test

Another variation of CQ testing that is widely used in the field is the Modified General Question Test (MGQT) (Reid & Inbau, 1977). The MGQT has enjoyed popular support amongst practitioners for many years. Nonetheless, only two analog

studies (Horvath, 1988; Widacki & Horvath, 1978) to date have used this procedure to explore the efficacy of CQ testing. The MGQT in its most basic form uses only three types of questions. They are irrelevant questions, relevant questions and control questions (Horvath, 1988). Unlike the ZOC test the MGQT may deal with a broader range of questions pertaining to the issue, an attribute of great utility since the commission of almost any crime ordinarily engenders a sequence of related events. For example, in a fictional, yet typical, "breaking and entering," some checks were stolen and later cashed by forging the owner's name. Questions that could be asked using an MGQT are: "Did you break into that house?" "Did you steal any checks from that house?" "Do you know who forged the owner's name on those checks?" and "Did you cash even one check stolen from that house?" The MGQT generally employs only two control questions and four or five relevant questions. It is a common practice not to rotate the questions on the first two charts administered, but then modify the order in which the questions are presented for the last test to create a mixed question test. Additionally, because there are only two control questions used for the MGQT, the comparative analysis of control and relevant questions is performed using questions that are often not adjacent to one another.

Nonexclusive and Exclusive Control Questions

Field examiners generally agree that control questions must cover a length of time wide enough to insure a person's answer is either a lie or there is at least concern for the veracity of their answer (Horvath, 1988; Podlesny & Raskin, 1978; Raskin, Barland & Podlesny, 1978; Reid & Inbau, 1977). However, there are two "schools of thought" (Horvath, 1988, p. 199) concerning the amount of time a control question should encompass.

The first one is supported by examiners who state a control question should include the broadest time period possible. These examiners purposely word control questions to include the period during which the offense under investigation occurred. An example of this question is: "Did you ever steal anything," or "Besides what you have told me did you ever steal anything else?" This is the type of control question developed by Reid (Reid & Inbau, 1977) and is called a nonexclusive control.

Backster (1969) concluded that nonexclusive control questions may be misinterpreted as relevant questions by a deceptive subject, thereby, evoking equal or greater responses to a control question when compared to a relevant question. Consequently, the number of deceptive subjects called truthful (false negative errors) would be unacceptably large. To prevent this type of error Backster modified the format of Reid's nonexclusive control question by adding a time bar, a

period of time which purposely excludes the offense at examination. For example, a person 22 years of age examined about a recent theft might be asked: "Before your 20th birthday, did you ever steal anything," or "Besides what you have told me, before your 20th birthday, did you ever steal anything else?" These questions are called exclusive controls.

Advocates favoring the use of nonexclusive control questions argue that an exclusive control question cannot arouse the same level of concern within an examinee as can a nonexclusive control question. These advocates reason that exclusive control questions will, by limiting the scope of time considered by an examinee, logically lower his/her level of concern for them. These advocates feel that any decrease in concern experienced by a person in answering a control question would reduce their responses to it. This would be especially true for the innocent because of the inherently strong signal value of relevant questions. Consequently, the use of exclusive control questions for testing should result in a larger number of innocent people called deceptive (false positive errors).

Purpose of This Study

The results of past studies (e.g., Podlesny & Raskin, 1978; Horvath, 1988) have raised many questions concerning the accuracy of CQ testing. In light of the strong opposing

beliefs advocated by those using one type of control question or another, as well as the different variations of control question testing, the answers to such questions may be found only in a controlled laboratory setting. For example, given a CQ polygraph test, what type of control question would produce the greatest accuracy, exclusive or non-exclusive control questions? Next, if both types of control question yield equal accuracy, does the use of one produce a different type of error compared to the other? Also, would the variation of control question test used, ZOC or MGQT, produce different levels of accuracy depending on the type of control questions employed?

These questions are very important. For example, if the answer to any of these questions were yes, a practitioner would be well advised to examine carefully the different interests at stake before making a decision on which variation of CQ testing, and type of control question to use for a particular situation. Consequently, one purpose of this study was to examine the relative accuracy and utility of the MGQT and ZOC variations of CQ polygraph testing in the same context. These procedures are used many times each day by practitioners in many different countries (Barland, 1988). However, this study was the first to examine the relative merits of these procedures, one compared to the other, in a controlled setting.

The second purpose of this study was to examine, the relative merits of using exclusive and nonexclusive control

questions with each of the two variation of CQ testing described earlier. To date only two studies have examined the relative effectiveness of these questions. However, the first (Podlesny & Raskin, 1978) study used a ZOC variation of CQ test, while the second (Horvath, 1988) used an MGQT. The results of these studies were equivocal at best, thus prompting this investigation.

CHAPTER II

Review Of The Literature

Chapter Overview

The use of an instrument (a polygraph) to record physiological data from which veracity is assessed, the psychophysiological detection of deception, has been the subject of increasing debate ever since William Marston publicized the process in 1917 (Iacono and Patrick, 1987a; Lykken, 1974, 1981). The Office of Technology Assessment (1983) published a critical review and evaluation of prior research examining this subject and found in part that CQ testing was the technique most often analyzed by researchers. However, the number of studies deemed acceptable for review at that time was few. The OTA (1983) found only 14 mock-crime (analog) studies that used "actual field polygraph [PDD] techniques to detect deception or concealed information" (p. 62). Today, seven years later an examination of the literature that is generally available to the public finds there are now 20 laboratory studies (see Table 1, p. 18) that would meet the OTA's criteria; a status that has left many issues as yet unresolved.

Insert Table 1 about here

To date research has focused largely on the validity and reliability of CQ testing (Barland & Raskin, 1975) and the effects of various factors thought to influence it's accuracy, for instance, the testing of people diagnosed as psychopathic (Hammond, 1980; Patrick & Iacono, 1989; Raskin & Hare, 1978), using alcohol (Bradley & Ainsworth, 1984) or drugs (Gatchel, Smith & Kaplan, 1984), the effects of employing various countermeasures (Dawson, 1980; Honts, Hodes, & Raskin, 1985; Rovner, 1979), the comparable accuracy of one variation of CQ test versus another (Driscoll, Honts & Jones, 1987; Forman & McCauley, 1986) and the relative validity and reliability of different techniques that also employ the recording of physiology for the detection of deception (Horvath, 1988; Podlesny & Raskin, 1978). It is only through the conduct of analog studies like these that the majority of empirical knowledge, concerning such issues, is now available (OTA, 1983).

To facilitate an introduction to the literature the results of the available laboratory studies are capsulated in Table 2 (see pgs. 20 & 21) then abstracts are presented for a sample of these studies permitting additional review and comparison. A summary of the literature believed relevant to this project is then offered. Finally, the chapter is

Table 1

Laboratory Experiments Examining the Control Question Technique Using a Simulated or Mock Crime

Number	Experiments Author(s)
1.	Barland, G. H., & Raskin, D. C. (1975)
2.	Bradley, M. T., & Janisse, M. P. (1981)
3.	Bradley, M. T., & Ainsworth, D. (1984)
4.	Dawson, M. E. (1980)
5.	Driscoll, L. N., Honts, C. R. & Jones, D. (1987)
6.	Forman, R. F., & McCauley, C. (1986)
7.	Gatchel, R. J., Smith, J. E., & Kaplan, N. M. (1984)
8.	Ginton, A., Daie, N., Elaad, E., & Ben-Shakhar, G. (1982)
9.	Hammond, D. L. (1980)
10.	Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985) [Exp #1]
11.	Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985) [Exp #2]
12.	Honts, C. R. (1986)
13.	Kircher, J. C. (1983)
14.	Horvath, F. S. (1988)
15.	Patrick, C. J., & Iacono, W. G. (1989)
16.	Podlesny, J. A., & Raskin, D. C. (1978)
17.	Raskin, D. C., & Hare, R. D. (1978)
18.	Rovner, L. I. (1979)
19.	Widacki, J., & Horvath, F. (1978)
20.	Yankee, W. J., & Grimsley, D. L. (1986)

concluded with the rationale for the present study.

Insert Table 2 about here

Selected Studies

Barland and Raskin (1975)

The first laboratory research to utilize field procedures and a mock-crime, an analog study, was conducted by Barland and Raskin (1975). Data was collected from 72 subjects who were randomly assigned to innocent (n=36) and guilty (n=36) conditions. Subjects were further divided into three feedback groups, positive, negative and no feedback. Positive feedback subjects were led to believe that a lie could easily be detected by the equipment, negative feedback subjects were led to believe that the equipment was malfunctioning and not working effectively; the third group received no feedback. The mock crime committed by guilty subjects was the theft of $10 from a designated desk drawer. Innocent subjects were told all the details of the crime before being tested. Classifications overall were 53% correct, 12% wrong, and 35% inconclusive, when inconclusive results were excluded from analysis the decisions were 81% correct. There were a total of 9 errors made, of those 3 guilty subjects (8.3%) were

Table 2

Results By Number and Percent Correct, Wrong, Inconclusive and Correct Excluding Inconclusives For Laboratory Experiments Examining the Control Question Technique Using a Simulated or Mock Crime

Experiments Number	Author(s)	No.(%) Correct	No.(%) Wrong	No.(%) Inconcl	(%) Correct No Inconcl
1.	Barland & Raskin	38 (53%)	9 (13%)	25 (34%)	(81%)
2.	Bradley & Janisse [A]	114 (59%)	22 (11%)	56 (30%)	(84%)
3.	Bradley & Ainsworth	28 (70%)	7 (18%)	5 (12%)	(80%)
4.	Dawson [B]	18 (75%)	3 (13%)	3 (12%)	(86%)
5.	Driscoll, Honts, & Jones [C]	15 (75%)	2 (10%)	3 (15%)	(88%)
6.	Forman & McCauley	21 (55%)	11 (29%)	6 (16%)	(66%)
7.	Gatchel, Smith, & Kaplan	18 (64%)	1 (04%)	9 (32%)	(95%)
8.	Ginton, Daie, Elaad, & Ben-Shakhar	6 (40%)	1 (07%)	8 (53%)	(86%)
9.	Hammond [D]	47 (76%)	5 (08%)	10 (16%)	(90%)
10.	Honts, Hodes, [E] & Raskin [Exp #1]	34 (71%)	7 (15%)	7 (14%)	(83%)
11.	Honts, Hodes, [E] & Raskin [Exp #2]	29 (51%)	13 (23%)	15 (26%)	(69%)
12.	Honts[2]	56 (47%)	32 (27%)	32 (26%)	(67%)

Table 2 (cont'd)

Results By Number and Percent Correct, Wrong, Inconclusive and Correct Excluding Inconclusives For Laboratory Experiments Examining the Control Question Technique Using a Simulated or Mock Crime

Experiments Number	Author(s)	No.(%) Correct	No.(%) Wrong	No.(%) Inconcl	(%) Correct No Inconcl
13.	Kircher	87 (87%)	6 (06%)	7 (07%)	(94%)
14.	Horvath	32 (80%)	7 (18%)	1 (02%)	(82%)
15.	Patrick & Iacono	30 (63%)	11 (23%)	7 (14%)	(73%)
16.	Podlesny & Raskin	32 (80%)	4 (10%)	4 (10%)	(89%)
17.	Raskin & Hare	39 (81%)	1 (02%)	8 (17%)	(98%)
18.	Rovner	57 (79%)	8 (11%)	7 (10%)	(88%)
19.	Widacki & Horvath	76 (95%)	2 (03%)	2 (02%)	(97%)
20.	Yankee & Grimsley	56 (78%)	4 (06%)	12 (16%)	(93%)

[A] Results for only GSR parameter.
[B] Results for only initial answer test.
[C] Results for only CQ Tests.
[D] Results for only experienced examiner.
[E] Results include people trained in countermeasures.

called truthful (false negative errors) and 6 innocent subjects (16.7%) were called deceptive (false positive errors). The ability to classify subjects as innocent or guilty wasn't significantly effected by the type of feedback given. The authors concluded that their results showed the ability to accurately assess the status of subjects as innocent or guilty was possible "at levels well beyond chance" (Barland & Raskin, 1975, p. 327).

Raskin and Hare (1978)

Three studies have examined the effectiveness of control question testing with psychopathic and non-psychopathic subjects. The first was conducted by Raskin and Hare (1978) who recruited subjects from a prison population. Data from 48 subjects was used for analysis. One half of the subjects were diagnosed as psychopathic and the other half as normal. Subjects from both groups were randomly assigned to innocent (n=24) and guilty (n=24) conditions. Guilty subjects were tasked with the theft of $20 from a designated area, while innocent subjects were just told about the crime. Subjects were tested to determine their status, innocent or guilty, after each one finished the task he was assigned. Classifications overall were 88% correct, 4% wrong and 8% inconclusive, when inconclusive results were not used for analysis the decisions were 96% correct. There were 2 errors (8.3%), both were innocent subjects classified as guilty

(false positive errors). Analysis showed there were no statistically significant differences in the accuracy of classifying psychopathic and non-psychopathic subjects as innocent or guilty. Overall the authors showed the accuracy of classifying subjects as innocent or guilty was highly significant.

Hammond (1980)

The second study was conducted by Hammond (1980) to test the effectiveness of CQ testing with both alcoholics and psychopaths. Data from a total of 62 subjects, 21 normal, 20 alcoholics and 21 psychopaths, was collected. Subjects were randomly assigned to innocent and guilty conditions. Guilty subjects were to steal $10 from a designated location and then undergo testing, innocent subjects were given only general information about the crime. Classifications overall were 40% correct, 3% wrong and 57% inconclusive, excluding inconclusive results decisions were 92.6% correct. Results showed there were 2 wrong decisions, both were innocent subjects (6.7%) classified as guilty (false positive errors). Analysis showed that the CQ test allowed discrimination between innocent and guilty alcoholics, psychopaths, and normals at a statistically significant level.

Patrick and Iacono (1989)

The last study exploring this topic was conducted by Patrick and Iacono (1989), replicating with some alterations, the Raskin and Hare (1978) experiment. The authors employed the services of two practitioners to collect physiological data from 48 male prisoners, using a Lafayette field polygraph instrument. A Beckman laboratory instrument was used by the principal author, Patrick, to simultaneously collect an additional set of physiological data for later comparison. One half of the prisoners were diagnosed as psychopathic (n=24) and the other half as non-psychopathic (n=24), these groups were equally divided and assigned further to innocent and guilty conditions. Subjects assigned to the guilty condition were tasked with stealing $20 from a jacket located in another room, while innocent subjects were only told about the crime. A ploy was used to increase the saliency of the laboratory context for all of the participants. Subjects were told that if 10 or more people failed to pass their tests by being classified as guilty, then all 48 subjects would loose a $20 bonus. Furthermore, when the experiment was finished a list would be given to all subjects showing those who passed and those who failed. No such list was in fact distributed and all subjects received the bonus for participating. Classifications overall, using the scores assigned by the original examiners, were 75% correct, 25% wrong and 0% inconclusive. There were 12 wrong decisions, 2 of the errors

were guilty subjects (8.3%) classified as innocent (false negative errors), while the other 10 errors were innocent subjects (41.7%) classified as guilty (false positive errors). When the scores assigned by a blind evaluator were used for classification the results were 63% correct, 23% wrong and 14% inconclusive. Using the blind evaluator's scores there were 11 wrong decisions, 3 of the errors were guilty subjects (13%) classified as innocent (false negative errors), while the other 8 errors were innocent subjects (33.3%) classified as guilty (false positive errors). Although the number of false positive errors was higher than in the majority of other studies the results of the analysis showed there were no statistically significant differences in classifying psychopathic and non-psychopathic subjects as innocent or guilty.

Bradley and Ainsworth (1984)

Alcohol has long been known to diminish a person's physiological and psychological functioning. Hence, the belief that it may also effect the accuracy of CQ testing (Reid & Inbau, 1977). The effect of alcohol intoxication on the accuracy of this technique was examined by Bradley and Ainsworth (1984). Data was collected from 40 white male college students who volunteered to submit to testing. There were 8 subjects assigned to an innocent condition and 32 subjects assigned as guilty. The task performed by guilty

subjects was a simulated shooting and robbery. The group of 32 guilty subjects were split in half so that 16 subjects were intoxicated at the time they committed the simulated shooting and the other 16 were sober. For testing purposes each group was again evenly divided so that 4 innocent subjects were intoxicated and 4 were sober. Of the 16 guilty subjects who were intoxicated at the time of the shooting, 8 subjects were again intoxicated, the other 8 remained sober. The 16 guilty subjects who were sober at the time of the shooting were divided so that 8 of them would be intoxicated and the last 8 guilty subjects again were sober. The amount of alcohol given each subject assigned to be intoxicated was calculated from his body weight to achieve about a .12% blood alcohol level. Overall classifications were 70% correct, 17% wrong and 13% inconclusive, when inconclusive results were excluded the decisions were 80% correct. Those subjects who were intoxicated at the time of the shooting were detected at a significantly lower rate than subjects in all other groups. In the other experimental conditions subjects were identified as innocent or guilty at levels that were statistically significant. The classification of sober subjects as innocent or guilty was 91.6% correct, 8.3% wrong and 0% inconclusive, while decisions about intoxicated subjects were 75% correct, 16.6% wrong and 8.3% inconclusive. Across all conditions the number of errors were different with 1 of 8 (12.5%) innocent subjects classified guilty (a false positive error) and 6 of 32 (18.8%) guilty subjects were classified as innocent (false

negative errors). Analysis showed that the ability to accurately classify subjects as guilty using a CQ test was greatly diminished if a person were intoxicated when they committed a crime. On the other hand, the differences in classifying innocent and guilty subjects who were intoxicated only at the time of testing, did not exceed the level of chance. Overall the CQ test again allowed the identification of innocent and guilty subjects with a level of accuracy that was statistically significant.

Gatchel, Smith and Kaplan (1984)

The ability of drugs to alter physiological responses also calls into question their effect on the accuracy of the psychophysiological detection of deception. To examine this question Gatchel, Smith and Kaplan (1984) studied the effects of "Propranolol," a beta-blocking drug used for the treatment of cardiovascular and anxiety-related disorders, on the accuracy of CQ testing. Data was collected from 28 male subjects who volunteered to participate in a "psychophysiological experiment" (p. 2). Subjects were randomly assigned to one of four equal groups, innocent with no drugs, innocent taking Propranolol, guilty with no drugs and guilty taking Propranolol. The simulated crime performed by guilty subjects was the theft of jewelry from a designated location. Innocent subjects were told that a mock crime was committed, but no details were given. Overall classifications

were 64% correct, 4% wrong and 32% inconclusive. When inconclusive results were excluded the decisions were 95% correct. The only error made in evaluating the data was a guilty person (7.1%) classified as innocent (false negative error). The CQ test discriminated between innocent and guilty subjects at levels exceeding chance, however, the proportion of subjects classified as inconclusive was larger than that found in most studies.

Rovner, Raskin and Kircher (1979)

Another area of research interest has been the susceptibility of the CQ test to intervention by subjects possessing detailed knowledge about the process (Rovner, Raskin & Kircher, 1978), or using countermeasures (Dawson, 1980; Honts, Hodes & Raskin, 1985). The effect of subjects having detailed knowledge about the CQ test was investigated by Rovner, Raskin and Kircher (1979). Data was collected from 72 male subjects, 36 innocent and 36 guilty, who volunteered to participate in a "lie detection experiment." Subjects assigned to guilty conditions were instructed to steal a ring from a designated location, while innocent subjects were informed that a crime had been committed. Prior to testing the two groups were each divided into three subgroups, STD, INFO and INFO + PRACTICE, and given additional instructions accordingly. The STD group received no information or practice to assist them in appearing innocent. The INFO group

received detailed information about the CQ test and suggestions on how to appear innocent while being tested. The third group, INFO + PRACTICE, received the same detailed information as the INFO group, but were also given two practice CQ tests and feedback on their performance. Classification of subjects in the STD and INFO groups were 88% correct, 4% wrong and 8% inconclusive. When inconclusive results were excluded decisions were 95% correct. Classification of subjects in the INFO + PRACTICE group were 62.5% correct, 25% wrong and 12.5% inconclusive. When inconclusive results were excluded decisions were 71% correct. There were 2 errors in the STD and INFO groups, 1 innocent subject in each group was classified as guilty, whereas the INFO + PRACTICE group had a total of 6 errors; 3 innocent subjects were classified as guilty (false positive errors), and 3 guilty subjects were classified as innocent. Overall there were 5 innocent subjects (13.9%) classified as guilty (false positive errors) and 3 guilty subjects (8.3%) classified as innocent (false negative errors). Analysis showed The CQ test was again highly effective for classifying subjects as innocent and guilty, especially those who were naive or had access only to information about the procedure. The accuracy of classifying subjects in the group that received information and practice, as innocent or guilty, also exceeded the level of chance, but at a level that was significantly lower than that recorded for the other two groups.

Dawson (1980)

The ability to appear innocent, when one is not, varies from person to person due to individual differences, life experience and even training. Actors are especially adept at controlling their emotions because of their experiences, training and a desire to succeed in their chosen field. One form of training, the "Stanislavsky method", teaches actors to use "personal memories of sensory experiences in order to recreate emotional states" (Dawson, 1980, p. 9). The effectiveness of the CQ test against this type of countermeasure was examined by Dawson (1980). Data was collected from 24 student actors. Six males and 6 females were randomly assigned as innocent, and 7 males and 5 females were assigned as guilty. All subjects were promised a monetary reward if they were classified as innocent. The crime committed by guilty subjects was the theft of $20 from a desk drawer, whereas innocent subjects were simply asked to submit to testing. All subjects were instructed to employ the techniques taught them at the Strasberg Theater Institute to appear innocent during their test. The series of questions pertaining to the theft of the money was repeated four times for each subject. For two of these repetitions, subjects were instructed to answer as soon as a question was finished. This was titled an immediate answer test (IAT). For each of the two remaining repetitions subjects were told to delay their answer to a question until a red light, which illuminated when

a question was asked, was extinguished eight seconds later. This procedure was titled a delayed answer test (DAT). Classifications of the IAT response overall were 75% correct, 12% wrong and 13% inconclusive. Excluding the inconclusive results, decisions were 86% correct. The eight second period between a question being asked and the signal to answer, allowed analysis of the physiological response to both the question and the answer. Classifications based on only DAT question responses overall were 84% correct, 8% wrong and 8% inconclusive. When the inconclusive results were excluded, decisions were 91% correct. Classifications based only on DAT answer responses overall were 29% correct, 8% wrong and 63% inconclusive. Excluding the inconclusive results, decisions were 78% correct. Classifications using both IAT responses and the DAT question responses overall were 88% correct, 8% wrong and 4% inconclusive. When inconclusive results were excluded, the judgments were 91.3% correct overall. The errors made in each condition were innocent subjects classified as guilty (false positive errors). There were 3 errors (25%) with IAT responses and 2 errors each (16.7%) with DAT question and DAT answer responses. Analysis showed the use of mental imagery by student actors was not successful in defeating the CQ test at levels exceeding chance. The ability to accurately classify subjects as innocent or guilty using the CQ test was again demonstrated to be statistically significant.

Honts, Hodes and Raskin (1985)

Two additional studies examining the effect of various countermeasures on the accuracy of CQ testing were published in 1985 by Honts, Hodes and Raskin. Data for experiment 1 was collected from 48 introductory psychology students. Subjects were randomly assigned to one of four groups, one group was innocent and the other three groups were guilty. Guilty subjects were instructed to commit a mock crime, the theft of a psychology final examination from a designated location. Innocent subjects were told only the general nature of the crime. The first group of guilty subjects was a control group, they were given no further information and tasked only with stealing the psychology exam. The second group of guilty subjects was told to use a pain countermeasure, which involved biting their tongue at appropriate times during the CQ test. The last group of guilty subjects was told to use a muscle countermeasure. This maneuver involved the pressing of their toes into the floor at appropriate times during the CQ test. The two counter-measure groups were given information about the CQ test and the differences between irrelevant, relevant and control questions so their countermeasure could be used to create a response each time a control question was asked. Examination of the physiological data was performed three times using semi-objective (numerical) scoring techniques (Barland & Raskin, 1975). The first analysis was conducted by the original examiner, the second by an independent evaluator

and the third, some years later, again by the original examiner using a system developed at the University of Utah (Raskin & Hare, 1978). Overall classifications of innocent and guilty subjects the first time were 67% correct, 17% wrong and 16% inconclusive. Excluding the inconclusive results, the decisions were 80% correct. A total of eight errors were made in classifying subjects. Five innocent subjects (41.7%) were classified as guilty (false positive errors) and 3 guilty subjects (7%) were classified as innocent (false negative errors). The classifications made by the independent evaluator overall were 71% correct, 15% wrong and 14% inconclusive. When inconclusive results were excluded, the decisions were 83% correct. There were a total of seven errors made by the independent evaluator; 4 innocent subjects (33.3%) were classified as guilty (false positive errors) and again 3 guilty subjects (7%) were classified as innocent (false negative errors). Classifications by the original examiner using the Utah system were 58% correct, 8% wrong and 33% inconclusive. Excluding the inconclusive results, judgements were 87.5% correct. Errors made using the Utah system were 2 innocent subjects (16.7%) classified as guilty (false positive errors) and 2 guilty subjects (4.7%) classified as innocent (false negative errors). Analysis of the data showed, in each of the evaluations, the accuracy of classifying subjects as innocent or guilty exceeded the level of chance. Additional analyses of the total semi-objective scores found a significant Group X Gender interaction.

Innocent female subjects had lower mean total numerical scores (M=0.8) than innocent male subjects (M=6.8); and pain countermeasure female subjects had higher mean total numerical scores (M=1.3) than pain countermeasure male subjects (M=-13.2). However, the accuracy of classifying subjects in the different guilty conditions, as guilty, did not exceed the level of chance. Overall the counter-measures employed did not prove effective in defeating the CQ test in this experiment.

Experiment 2 used data collected from 57 subjects enrolled in an introductory psychology course. Subjects were randomly assigned to one of three groups, innocent, guilty only and guilty with countermeasure instructions. Guilty subjects were instructed to commit a mock crime, the theft of a psychology final examination from a folder on top of a designated desk. Innocent subjects were told only the general nature of the crime. The first group of guilty subjects was again assigned to be a control group, they were given no further information and tasked only with stealing the psychology exam. The second guilty group was given instructions on both the countermeasures used in Experiment 1 and given a practice CQ test. During this test subjects were instructed to demonstrate use of the countermeasures for an assistant and told to practice them at home. Examination of the physiological data was performed by the original examiner and independent evaluator using semi-objective scoring techniques (Barland & Raskin, 1975) similar to those used in

the first experiment. Overall the classifications of innocent and guilty subjects, by the original examiner, were 65% correct, 21% wrong and 14% inconclusive. Excluding the inconclusive results, the decisions were 75% correct. The original examiner made a total of 12 errors; 5 innocent subjects (26.3%) were classified as guilty (false positive errors) and 7 guilty subjects (18.4%) were classified as innocent (false negative errors). The independent evaluator's classifications were 51% correct, 23% wrong and 26% inconclusive. When inconclusive results were excluded, decisions were 69% correct. There were a total of 13 errors made by this evaluator; 4 innocent subjects (21.1%) were classified as guilty (false positive errors) and 9 guilty subjects (23.7%) were classified as innocent (false negative errors). Analysis showed the accuracy of classifying subjects as innocent or guilty by the original examiner and the independent evaluator, exceeded the level of chance with no significant interactions. However, the accuracy of classifying subjects as guilty varied significantly between the guilty only group and the guilty with countermeasures group. Subjects who received instructions and practice in using countermeasures were detected at a rate significantly lower than those who were guilty and did not receive any type of countermeasure instruction. In their conclusion the authors said "some data from these two experiments suggest that relatively sophisticated training is required for countermeasures to be successful" (Honts et al., 1985, p. 185).

Questions concerning the comparative accuracy and utility of different variations of CQ testing and other psychophysiological detection of deception techniques often inspire debate among practitioners and others interested in the detection of deception. To date at least four published studies (Driscoll, Honts & Jones, 1987; Forman & McCauley, 1986; Horvath, 1988; Podlesny & Raskin, 1978) have examined this issue empirically.

Forman and McCauley (1986)

The relative accuracy of CQ testing, the Positive Control Test ([PCT] Reali, 1978) and the Guilty Knowledge (GK) test Lykken, 1959, 1974) was examined in a study conducted by Foreman and McCauley (1986). The difference between the PCT and CQ testing is that the PCT uses only irrelevant and relevant questions. The absence of questions that are commonly acknowledged as controls is offset by asking each relevant question twice when a PCT is administered. The first time a relevant question is asked its prefaced with instructions to tell the examiner a lie; for example, "Tell me a lie, did you steal the car?" The second repetition of the question is prefaced with instructions to tell the truth; "Tell me the truth, did you steal the car?" The assumption made in employing the PCT is that a subject will exhibit greater physiological arousal when telling a lie. If a subject displays greater responses in saying yes they

committed a crime then a decision of innocent is made. On the other hand if greater responses are given in denying the crime a decision of guilty is made.

The GK test (Lykken, 1974, 1981, 1988a) employs a series of multiple choice questions. In each, is embedded a piece of information that only authorities and a guilty subject(s) would have knowledge of. In each of these questions, the alternative choices appear to be equally plausible as the valid piece of information. The first alternative for each question is always false to guard against a significant physiological response often elicited by answering the first question in a test. Decisions of guilty or innocent are based on the number of times greater physiological arousal is seen in responding to the correct alternative for each multiple choice question.

To examine the above procedures, data was collected from 38 female subjects, 16 innocent and 22 guilty, by giving each subject in one session a PCT, a CQ test and a GK test. The authors did not explicitly state whether subjects received a break between each test or not, as in field applications of the procedure. Analysis of the data was performed by both the original examiner and a blind evaluator. Overall classifications of innocent and guilty subjects by the original examiner using the PCT, were 61% correct, 24% wrong and 16% inconclusive. Excluding the inconclusive results, decisions were 72% correct. There were a total of 9 errors in classification using the PCT; 3 innocent subjects (18.7%) were

classified as guilty (false positive errors) and 6 guilty subjects (27.3%) were classified as innocent (false negative errors). Classifications using the CQ test were 55% correct, 29% wrong and 16% inconclusive. Excluding inconclusive results decisions were 66% correct. There were a total of 11 classification errors using the CQ test; 8 innocent subjects (50%) were classified as guilty (false positive errors) and 3 guilty subjects (13.6%) were classified as innocent (false negative errors). Overall classifications by the original examiner, using the GK test, were 66% correct, 29% wrong and 5% inconclusive. Excluding inconclusive results the examiners decisions were 69% correct. There was a total of 11 classification errors made using the GKT; all 11 were guilty subjects (50%) classified as innocent (false negative errors).

The independent evaluator's classifications using the PCT were 63% correct, 21% wrong and 16% inconclusive. When inconclusive results were excluded, decisions were 75% correct. There was a total of 8 classification errors made; 1 innocent subject (6.3%) was classified as guilty (a false positive error) and 7 guilty subjects (31.8%) were classified as innocent (false negative errors). The evaluator's classifications employing the CQ test were 63% correct, 18% wrong and 18% inconclusive. Excluding inconclusive results decisions were 77% correct. There was a total of 13 classification errors made using the CQ test; 8 innocent subjects (50%) were classified as guilty (false positive errors) and 5 guilty subjects (22.7%) were classified as

innocent (false negative errors). Overall classifications by the blind evaluator using the GK test were 63% correct, 32% wrong and 5% inconclusive. Excluding the inconclusive results decisions were 67% correct. There was a total of 12 errors made; 1 innocent subject (6.3%) was classified as guilty (a false positive error) and 11 guilty subjects (50%) were classified as innocent (false negative errors). Results of the statistical analysis showed that the PCT and GK test achieved the same level (73%) of average accuracy, while the CQ test had a lower (65%) average accuracy. The GK test provided the highest detection of innocence and the CQ test the lowest. In summary, the authors concluded that the PCT "appears to be a promising innovation in polygraph testing" and was only "mildly biased in favor of the truthful where the CQ test was biased against the detection of truth" (p. 696). This conclusion, was disputed by the results of another study (Driscoll, Honts & Jones 1987) that also examined the PCT and CQ tests.

Driscoll, Honts and Jones (1987)

The relative validity and reliability of Positive Control Testing (PCT) and CQ testing was also examined by Driscoll, Honts and Jones (1987) using a simulated crime, the theft of a ring. Data was collected from 40 subjects who were randomly assigned to one of four groups; innocent (n=10) with the PCT administered first, innocent (n=10) with the CQ test

administered first, guilty (n=10) with the PCT administered first and guilty (n=10) with the CQ test administered first. Each type of test was given three different times to each subject with the question order altered slightly from test to test. Overall classifications using the PCT were 50% correct, 5% wrong and 45% inconclusive. Excluding inconclusive results the decisions were 91% correct. There was a total of 2 classification errors made using the PCT; 2 guilty subjects (10%) were classified as innocent (false negative errors). Overall the classification of subjects as innocent or guilty using the CQ test were 90% correct, 0% wrong and 10% inconclusive. Excluding inconclusive results the accuracy of decisions were 100% correct, with no classification errors.

Podlesny and Raskin (1978)

Podlesny and Raskin (1978) conducted a seminal study examining the relative validity and reliability of the GK test and CQ testing using a simulated crime, the theft of a ring. In conducting this investigation the authors used a zone comparison variation of control question test, as have the majority of other researchers. The authors included in their design, an analysis of the two types of control question used for CQ testing. Data for 60 subjects recruited from the community, was used for analysis. Subjects were randomly assigned to one of six treatment conditions, innocent or guilty taking a GK test, taking a CQ test using the exclusive

(time bar) type of control questions or taking a CQ test using the nonexclusive (no time bar) type of control questions. Overall classifications made by an independent blind evaluator using the GK test were 80% correct, 10% wrong and 10% inconclusive. Excluding the inconclusive results decisions were 89% correct. There were two errors made using the GK test; 2 guilty subjects (20%) were classified as innocent (false negative errors). Overall classifications using the CQ test with exclusive (time bar) control questions were 85% correct, 5% wrong and 10% inconclusive. Excluding inconclusive results the accuracy of decisions were 94% correct. There was only one error in classification made, 1 guilty subject (10%) was classified as innocent (a false negative error). Classifications using the CQ test with nonexclusive (no timebar) control questions overall were 75% correct, 15% wrong and 10% inconclusive. Excluding the inconclusive results the accuracy of decisions were 83% correct. There were a total of three classification errors made; 1 innocent subject (10%) was classified as guilty (a false positive error) and 2 guilty subjects (20%) were classified as innocent (false negative errors). Analysis showed the only significant main effect was for status, the classification of subjects as innocent or guilty. The differences in classification were not significantly effected by the type of control question used. Nevertheless, review of the data demonstrated that exclusive control questions were "superior in some respects" (p. 357) to nonexclusive control

questions. Analysis of the mean total numerical scores (Barland & Raskin, 1975) showed identification of innocent subjects was statistically significant using both types of control question, but only exclusive control questions provided classification of guilty subjects at a level exceeding chance.

Horvath (1988)

The relative validity and reliability of CQ testing and Relevant/Irrelevant (R/I) testing was examined by Horvath (1988) in a mock-crime (analog) study, the theft of money, using a Modified General Question Test (MGQT), another variation of control question testing. The author also reexamined the relative effectiveness of exclusive and nonexclusive control questions (Podlesny & Raskin, 1978) in the same context.

The R/I technique is the predecessor to CQ testing and utilizes only relevant and irrelevant questions. R/I testing assumes that subjects who are guilty (deceptive) will exhibit greater physiological responses to relevant questions then they will to irrelevant questions. Conversely, innocent (truthful) subjects are expected to present responses that are about equal to both types of question because their answers to each are truthful (Horvath, 1988; Weir, 1974). Although R/I testing is still used, the assumptions on which it's founded are strongly questioned by others interested in field

polygraph testing (Lykken, 1981; Reid& Inbau, 1977).

Sixty white males were recruited to participate in a lie detection study. Subjects were randomly assigned to innocent (n=30) and guilty (n=30) conditions and again randomly assigned for testing with either the R/I test, the MGQT using exclusive control questions or the MGQT using nonexclusive control questions. Review of the test data and classification of subjects was made by two independent evaluators. The first evaluator's classifications using exclusive control questions, were 70% correct, 25% wrong and 5% inconclusive. Excluding inconclusive results decisions were 74% correct. There was a total of 5 errors; 3 innocent subjects (30%) were classified guilty (false positive errors) and 2 guilty subjects (20%) were classified innocent (false negative errors). The first evaluator's classifications using nonexclusive controls were 90% correct, 10% wrong and no inconclusive. There were 2 errors made; 1 innocent subject (10%) was classified guilty (a false positive error) and 1 guilty subject (10%) was classified innocent (a false negative error). The first evaluator's classifications using the R/I test data were 45% correct, 55% wrong and no inconclusive. There were a total of 11 errors made; all 10 innocent subjects (100%) were classified guilty (false positive errors) and 1 guilty subject (10%) was classified innocent (a false negative error). The second evaluator's classifications, using exclusive control questions differed slightly, they were 70% correct, 30% wrong and no inconclusive decisions. There was a total of 6 errors;

4 innocent subjects (40%) were classified guilty (false positive errors) and 2 guilty subjects (20%) were classified innocent (false negative errors). The second evaluator's classifications using nonexclusive controls again differed little, these were 80% correct, 15% wrong and 5% inconclusive decisions. Excluding inconclusive results decisions were 84% correct. There was a total of 3 errors made; 2 innocent subjects (20%) were classified guilty (false positive errors) and 1 guilty subject (10%) was classified innocent (a false negative error). The second evaluator's classifications of the R/I test data were 40% correct, 55% wrong and 5% inconclusive. Excluding inconclusive results decisions were 42% correct. The errors made by the second evaluator were the same as those made by the first; all 10 innocent subjects (100%) were classified guilty (false positive errors) and 1 guilty subject (10%) was classified innocent (a false negative error). Analysis showed the accuracy of discriminating between innocent and guilty subjects exceeded the level of chance using the CQ test, but not with the R/I test. Differences in the accuracy of classifying subjects as innocent or guilty, using the two types of control question, were not significant. Nonetheless, further analysis showed the mean total numerical score for the guilty tested with nonexclusive control questions (M=-24.5), was more extreme than the mean total numerical score for the guilty tested with exclusive control questions (M=-12.0). These results were opposite those of the Podlesny and Raskin (1978) study which

showed that the mean total numerical score for the guilty tested with exclusive control questions (M=-11.7), was more extreme than the mean total numerical score for the guilty tested with nonexclusive control questions (M=-6.3).

The Literature

The polygraph and lie-detection procedures now used day-to-day began their development almost one hundred years ago. Although the polygraph's historical evolution is viewed as important its beyond the scope of this project to give the subject fitting attention. Consequently, if a reader desires more information on this topic s/he should review one of the excellent narratives that are available on this subject (Reid & Inbau, 1977; Trovillo, 1939).

In the last twenty to twenty-five years the use of the polygraph, especially by local, state and federal government agencies, has increased dramatically (Lykken, 1981, 1988b; OTA, 1983). Correspondingly, during this same period the number of polygraph examiners being trained each year also grew. As a result of the polygraph's proliferation the number of scientists inquiring about the detection of deception, and the amount of information relevant to the subject, similarly increased, but at a much slower pace (OTA, 1983).

In a report to the Department of Defense a study group, that looked at training and facilities for the DOD, said a three fold increase in the number of examiners trained by the

federal government was required because of the need for security screening examinations (Barland, Herbold-Wootten, Orne & Yankee, 1987). In addition, the study group said that more material on lie detection had been published in the last ten years than in the previous sixty. However, the authors also said "it is unfortunate that a number of the studies -- some conducted by scientists, others by the polygraph community -- contain serious flaws which cloud the issue" (p. 268). The source of this problem was seen as the lack of scientific training possessed by those in the polygraph community and a lack of polygraph training for most scientists (Barland et al., 1987). When these elements, the expansion of polygraph services, increased empirical review, and the dissemination of information with a wide qualitative variance, are considered collectively, the ensuing controversies associated with this subject are somewhat easier to understand.

Field Polygraph Tests

If a person wanted to know the number of field polygraph techniques in existence today, s/he could expect a number of different answers depending on the source queried. The number of techniques are in fact few, compared to the number of field polygraph procedures currently in use. Although many of these procedures are called techniques, they are actually variations of an older technique. Often these new techniques

are created simply by altering the sequence of question presentation used by another procedure. These types of changes are usually accompanied by a claim of greater validity and reliability, but seldom are changes made because of an empirical review. In spite of these practices CQ testing appears to be "sufficiently robust that many of the minor differences in the manner in which that testing is structured have little effect" (Horvath, 1988, p. 208).

Three types of field polygraph test are discussed here. Each type of test was subjected to empirical review in one of the studies from which the present research evolved. The first two, relevant-irrelevant (R-I) testing and CQ testing, are different techniques (Horvath, 1988; Reid & Inbau, 1977). The third, guilty knowledge (GK) testing, is a variation of the peak of tension (POT) testing technique. The oldest procedure is the R-I test, this is followed by the CQ test and the GK test.

The literature shows that the R-I test was the base from which the others developed (Horvath, 1988; Larson, 1932). The use of this test requires that subjects be asked two types of questions: irrelevant questions and relevant questions. Irrelevant questions are, as explained earlier in the introduction, simple questions of fact, to which the examiner and examinee both know the answer is true. For example: "Are you 21 years of age?" Relevant questions on the other hand are those that deal with the issue in question, for example, an examiner may ask in the investigation of a theft: "Did you

steal the money?" The assumption made by those using the R-I test is that a deceptive (guilty) person will show greater physiological responses to the relevant questions, because they are not answered truthfully. Correspondingly, it's assumed that asking an innocent person irrelevant and relevant questions will elicit minor variations in his/her physiological responses, because the answers given to both types of question are truthful and should not evoke a large differential response (Horvath, 1988).

The premise on which the R-I test is based has been strongly questioned by people in and outside the polygraph field (Horvath, 1988; Lykken, 1981; Reid & Inbau, 1977). Those opposed to the R-I test contend that the relevant questions, and their significance relative to the issue being examined, should be obvious to the innocent and guilty alike; consequently, in looking for the guilty an unacceptable number of innocent people would incorrectly be called deceptive (false positive errors). Similarly, if a guilty person were physiologically unresponsive s/he would give about equal responses to both the irrelevant questions and the relevant questions. With no large differential responses evident a guilty person would appear to be truthful (a false positive error) (Reid & Inbau, 1988).

Recognizing the inherent problems with the R-I test John Reid, a Chicago polygraph examiner, developed what he later called the CQ test, introducing it to the polygraph field in 1947 (Horvath, 1988, Reid & Inbau, 1977). This technique is

now used for a majority of the field examinations conducted by practitioners (OTA, 1983). The CQ test, as described in the introduction, uses three types of questions; irrelevant, relevant and control questions. Unlike the R-I test, in which irrelevant questions are used for decision making, the CQ test uses irrelevant questions as buffers and to establish norms (Horvath, 1988). The relevant questions are directed at some specific issue related to the matter under investigation, while control questions relate generally to "the motive for the offense" (Horvath, 1988).

According to one advocate CQ test theory is posited on three assumptions: (1) that subjects who are mentally healthy know when they are telling the truth; (2) subjects who are deceptive (guilty) will display greater arousal to relevant questions, and (3) subjects who tell the truth (innocent) will display greater arousal to control questions (Jayne, 1986). These assumptions are soundly rejected by those opposed to CQ testing. For example, one opponent writing about this issue (Lykken, 1981) said it would require a ". . . simplistic, robotlike conception of human nature" (Lykken, 1981, p. 117) before one could accept such a theory. Opponents to CQ testing (Kleinmuntz & Szucko, 1982; Lykken, 1974, 1979, 1981, 1988b) argue that any relevant question will still be obvious. Therefore, as with the R-I test, a search for the guilty will continue to yield unacceptable numbers of innocent people incorrectly classified as deceptive (false positive errors). As Lykken (1974, 1981) has stated, there is no widely accepted

theory that supports the claims of those advocating the CQ test's use. Lykken (1988) states "that polygraph lie detection does not and, in the foreseeable future, probably cannot work well enough to justify its continued use in the field" (p. 125).

A field polygraph test that is advocated by even the strongest opponents to CQ testing (e.g., Lykken, 1981, 1988b), is the guilty knowledge (GK) test. This particular variation of POT field polygraph test was introduced as a result of research conducted by Lykken (1959, 1960). When one considers the proposal made by Lykken (1988b), regarding the future of polygraph lie detection, his work with the GK test would appear to present a paradox. However, Lykken states:

> "the GK test . . . does not attempt to determine whether the respondent is lying but, rather, whether he or she possesses guilty knowledge, that is, whether the respondent recognizes the correct answers, from among several equally plausible but incorrect alternatives, to certain questions relating to a crime" (1988b, p. 121).

The GK test is conducted by creating a series of multiple-choice questions concerning pieces of information that only a person committing a crime, and later the police, would know (Lykken, 1988b). For example, a suspect in a theft of money, that was kept in an envelope in a drawer, could be asked a series of questions on the amount of money, what the money was kept in, and where the money came from. If the suspect is innocent his/her responses to both the noncritical and critical pieces of information are expected to be about equal. On the other hand if the suspect is guilty greater

physiological responses to the critical items are expected compared to those for the non-critical items.

Lykken's (1981) variation to the POT test has never been subjected to empirical review in a field context. Nonetheless, results of other studies (e.g., Gustafson, Lawrence & Orne, 1964; Hikita & Suzuki, 1965) examining POT testing, in a number of different settings, supports the use of this procedure when possible. Although, practitioners have long acknowledged the usefulness of POT testing its use is often problematic, because of the difficulty in locating information that would be appropriate or wasn't known to the public in general (Lykken, 1974).

Polygraph In The Laboratory

A large part of the knowledge now available regarding field polygraph testing as its used for the detection of deception was accumulated in laboratories (OTA, 1983). As pointed out in the introduction, a laboratory setting offers the researcher an opportunity to manipulate a number of factors that otherwise could not be studied systematically in the field (Orne, 1972). Of these factors the one most relevant to the study of deception would be the control of ground truth, the knowledge of who is telling the truth and who is not.

Laboratory settings have allowed those interested in field polygraph testing to begin to understand and describe

many of the dynamics that exist in a lie-detection context. Orne (1972) explains that because there are no physiological responses singularly associated with lying, the term lie-detection is a misnomer. During a lie-detection test a polygraph instrument is used simply to record a person's emotional arousal in response to a question, that has psychological significance for him or her. This emotional arousal is not necessarily related to the act of telling a lie per se. While many knowledgeable practitioners would fully agree with Orne, it was laboratory research that best demonstrated this fact.

In a study conducted by Kugelmass, Lieblich, and Bergman (cited in Orne, 1972) a simple card test was used with subjects who were motivated to deceive. Subjects were instructed to select a numbered card from a group of several and then to answer "YES" each time they were asked through a tape recording if they picked a certain number, including the number they selected. In essence each subject was told to tell a series of lies and to tell the truth only once. In this setting subjects who were motivated to deceive were identified at greater than chance levels, even though they told the truth by answering "YES" to the number they chose.

The importance of laboratory research is obvious to scientists and many practitioners alike (Horvath, 1980; Orne, 1972, Reid & Inbau, 1977). However, many people questioned the value of studies like that above, cited by Orne (1972), saying they had little or no relevance because the procedures

employed did not reflect those used by practitioners in the field. Consequently, it's only in the last twenty to twenty-five years that some researchers (see Table 1, pgs. 18 & 19) have brought CQ testing into a laboratory setting.

CQ Testing and Error

Those opposed to CQ testing (Furedy, 1988; Lykken, 1974, 1981, 1988b; Kleinmuntz & Szucko, 1982) state that for the innocent the accuracy of a CQ test, as it's used in the field, is no better than the toss of a coin. Opponents usually disregard the results of laboratory studies saying the findings can be generalized "to the real world only if we succeed somehow in simulating the consequences that are associated with polygraph testing in the real world" (Lykken, 1988, p. 114).

The results of at least two laboratory studies appear, even in a controlled setting, to support the opponents thesis. The first study was conducted by Forman and McCauley (1986) and reported 50% false positive errors. The second study was conducted by Patrick and Iacono (1989) and reported 33% false positive errors (also see abstracts in this chapter). Table 3 (see pgs. 54 & 55) shows that if all the laboratory studies

---

Insert Table 3 about here

Table 3

Number of Subjects, False Positive and False Negative Errors For Laboratory Experiments Examining the Control Question Technique Using a Simulated or Mock Crime

Experiments Number	Author(s)	Number Of Innocent (N)	Number(%) Of False Positives	Number Of Guilty (N)	Number(%) Of False Negatives
1.	Barland & Raskin	36	6 (16%)	36	3 (08%)
2.	Bradley & Janisse [A]	96	9 (09%)	96	13 (14%)
3.	Bradley & Ainsworth	8	1 (13%)	32	6 (19%)
4.	Dawson [B]	12	3 (25%)	12	0 (00%)
5.	Driscoll, Honts, & Jones [C]	10	1 (10%)	10	1 (10%)
6.	Forman & McCauley	16	8 (50%)	22	3 (14%)
7.	Gatchel, Smith, & Kaplan	14	0 (00%)	14	1 (07%)
8.	Ginton, Daie, Elaad, & Ben-Shakhar	13	1 (08%)	2	0 (00%)
9.	Hammond [D]	30	4 (13%)	32	1 (03%)
10.	Honts, Hodes, [E] & Raskin [Exp #1]	12	4 (33%)	36	3 (08%)
11.	Honts, Hodes, [E] & Raskin [Exp #2]	19	4 (21%)	38	9 (24%)
12.	Honts [E]	20	3 (15%)	100	29 (29%)

Table 3 (cont'd)

Number of Subjects, False Positive and False Negative Errors For Laboratory Experiments Examining the Control Question Technique Using a Simulated or Mock Crime

Experiments Number	Author(s)	Number Of Innocent (N)	Number(%) Of False Positives	Number Of Guilty (N)	Number(%) Of False Negatives
13.	Kircher	50	3 (06%)	50	3 (06%)
14.	Horvath	20	4 (20%)	20	3 (15%)
15.	Patrick & Iacono	24	8 (33%)	24	3 (13%)
16.	Podlesny & Raskin	20	1 (05%)	20	3 (15%)
17.	Raskin & Hare	24	1 (04%)	24	0 (00%)
18.	Rovner	36	5 (14%)	36	3 (08%)
19.	Widacki & Horvath	60	1 (02%)	20	1 (05%)
20.	Yankee & Grimsley	36	0 (00%)	36	4 (11%)

[A] Results for only GSR parameter.
[B] Results for only initial answer test.
[C] Results for only CQ Tests.
[D] Results for only experienced examiner.
[E] Results include people trained in countermeasures.

were placed in order, according to the percent of false positive error in each one, these two studies would, at this time,define the upper limit on this type of error. On the other hand, the lower limit would be defined by the error rates reported in the studies done by Gatchel, Smith and Kaplan (1984), and Yankee and Grimsley (1986). These studies reported no false positive errors.

The literature shows that controversy concerning the CQ test's effectiveness has resulted in its receiving increased review and empirical study (OTA, 1983). As Horvath (1988) points out, two issues that have been of major interest are "the relative effectiveness of testing methods and the determinants of errors in CQ testing (Forman & McCauley, 1986; Honts & Hodes, 1982; Podlesny & Raskin, 1978)" (p. 199). Accordingly, Horvath (1988) focused on a more basic issue not addressed by prior research: "Does CQ testing protect against false positive errors that would be expected in testing without such controls" (p. 199)? To answer this and other questions Horvath (1988) conducted a study to compare the relative effectiveness of the R-I technique and the CQ test. The results of that study showed there were significantly fewer false positive errors when control questions were used for testing, as compared to tests that did not use control questions (see Horvath, 1988, abstract in this chapter).

However, more relevant to the present research was the distribution of errors that occurred when control questions were used. In Horvath's (1988) study, those subjects examined

with a CQ test, had it administered using one of the two types of control questions. Half of the subjects were tested using exclusive (time-bar) control questions, while the other half were tested using nonexclusive (no time-bar) control questions. Analysis showed there were more false positive errors using exclusive control questions than there were using nonexclusive controls. Additionally, the nonexclusive control questions gave "more effective identification of both guilty and innocent subjects. . ." (p. 208, Horvath, 1988). Horvath's findings were opposite those shown in the first study that examined this issue. Podlesny and Raskin (1978) examined the relative effects of exclusive and nonexclusive control questions (see abstract in this chapter) as part of a larger research project conducted to examine the relative effectiveness of the CQ test and the GK test within the same context. The results of their study demonstrated that both exclusive control questions and nonexclusive control questions gave significant identification of innocent subjects. However, only exclusive control questions gave significant identification of the guilty, while nonexclusive control questions did not. Therefore, exclusive control questions were judged by Podlesny and Raskin (1978) to be superior. In response to these differences Horvath (1988) concluded that "it is difficult to reconcile these findings. . . . " (p. 207) and ". . . it is clear that additional research of this topic is warranted" (p. 208).

Rationale for the Present Study

Laboratory studies have examined the validity and reliability of CQ testing relative to many different factors. Although the results of these studies vary from one to another, the ability to accurately identify the innocent and guilty generally is significant (see Table 3, pgs. 54 & 55). However, a review of the literature shows a number of issues that haven't been addressed at all or lack clarity. For instance, the contradictory results of Horvath's (1988) study, relative to Podlesny's and Raskin's (1978) findings, led this author to question in part the comparative accuracy of the MGQT and ZOC test. In the past all but two mock-crime (analog) studies have used a ZOC test to examine various issues relating to the accuracy of CQ testing. This is related more probably to the fact that the majority of mock crime studies have either been conducted by people trained to use the ZOC test or the study was patterned after one in which a ZOC test was used. This proposition finds support when the results of the two mock crime studies (Horvath, 1988; Widacki & Horvath, 1978) that have used an MGQT variation of CQ test are examined (see Table 2 and 3). Both studies showed essentially the same level of accuracy as studies using some variation of ZOC test (Horvath, 1988). The ZOC test and the MGQT are both widely used by field practitioners. However, issues concerning the relative accuracy of these two variations of CQ testing within the same context, or the

comparative effect(s) of any factor on them, was unknown. These two variations of CQ test had never been the subject of empirical review within the same context. Therefore, the present study was designed in part to examine the relative accuracy and utility of the MGQT and ZOC tests in the same setting. Further, the study provided an opportunity to analyze the effects of using exclusive (time bar) and nonexclusive (no time bar) control questions on each variation of CQ test within the same context. Moreover, due to the differences in philosophy (see introduction) of those choosing to use one type of control question or another, this project provided an opportunity to re-examine not only the type of errors made using exclusive control questions and nonexclusive control questions, but also the distribution of those errors for each type of control question one relative to the other.

The populations from which samples were drawn for past studies have varied, but researchers have generally relied upon on all male samples (Horvath, 1988; Podlesny & Raskin; 1978) or simply opted not to explore the possibility of an effect for gender (Dawson, 1980; Widacki & Horvath, 1978). Although the issue of gender effects has been explored in a few studies (Barland & Raskin, 1975; Honts, Hodes & Raskin, 1985; Yankee & Grimsley, 1986) the results are equivocal at best. One study (Honts, Hodes & Raskin, [Experiment 1] 1985) has reported significant effects related to gender differences, while other studies examining this issue (Honts, Hodes & Raskin, [Experiment 2] 1985; Barland & Raskin, 1975),

have found no significant differences attributed to gender. Because no substantive conclusion for the issue of gender effects has yet been demonstrated, the present study was designed to also examine the effect of gender relative the different CQ tests and control questions in the same setting.

CHAPTER III

Method

Subjects

A total of eighty white males (n = 80) and forty white females (n = 40) were selected from a pool of 176 volunteers recruited from introductory undergraduate criminal justice classes at Michigan State University. After being selected subjects were randomly assigned to treatment conditions. The subjects mean ages were, males, 19.8 years (SD = 1.65) and females, 19.4 years (SD = 1.22).

Polygraph Examiner

The author was a licensed police polygraph examiner with the State of Michigan at the time data was collected. He had fourteen years of police service, with three years of experience conducting polygraph exams in support of criminal investigations. Polygraph training was conducted by the Royal Canadian Mounted Police at the Canadian Police College in Ottawa, Canada, during the fall of 1983. The procedure taught all examiners was a ZOC test with exclusive control questions using Lafayette and Stoelting field polygraphs.

Chart Evaluator

Blind evaluation of the polygraph charts was performed by a person associated with the polygraph field for more than twenty-five years. His training included the relevant/irrelevant technique, variations of the control question technique, both ZOC and MGQT, and different methods of analysis, to include numerical scoring. This evaluator did not participate in either the Podlesny and Raskin (1978) or Horvath (1988) studies. The only information the evaluator had was that available from the polygraph charts. No other information was given until all evaluations were complete and the data subjected to analysis.

Apparatus

The polygraph examinations for this study were conducted in a small, quiet room, like those generally used for field examinations (Reid and Inbau, 1977). Subjects were seated in a standard Stoelting polygraph chair with adjustable wood arm rests. The recording of all physiological data was accomplished through the use of a field polygraph. The instrument was a Model #761-95GA, manufactured by the Lafayette Instrument Company of Lafayette, Indiana. Thoracic and abdominal respiration were recorded for each subject by means of two pneumograph tubes, one positioned around the thorax and the other around the abdomen. Skin resistance

response (SRR) was recorded through two stainless steel electrodes attached to the volar surface of the first and second fingers of the left hand and recorded in the DC mode. "Medi-Trace" conductivity gel (medical electrode paste) was used to insure the best possible contact. Cardiovascular activity was recorded by use of a standard pneumatic pressure cuff around the upper portion of the right arm. Pressure within the cuff was set to about 60 mm/hg and the cuff squeezed two or three times to equalize the air within it, generally resulting in a recording pressure of 40 to 55 mm/hg. The upper pneumograph and cardio tracings were both enhanced electronically.

Procedure

A pool of 176 students was recruited for this study during the winter and spring terms of 1986 at Michigan State University. There were 94 males and 82 females from introductory Criminal Justice (CJ110) classes who volunteered to take part. Instructors agreed in advance to grant extra credit to participants. Students were asked to take part in a "lie detection" experiment which would require them to take a polygraph examination. They were told that extra credit was available and a chance to earn a small cash reward if they chose to participate. A short writing assignment was available for those who wished extra credit, but did not want

to participate in the experiment.

To volunteer as a participant in this experiment students were instructed to sign a roster that was left with each instructor of the courses from which subjects were solicited. The students signed the roster with their name, telephone number and hours that would best fit their schedule. An assistant later used the rosters to contact each potential subject, asking if he/she still wished to participate in the study. Those who wanted to take part were then given a date and time to report to a location at the university for polygraph testing. Once an appointment was made the assistant randomly assigned subjects by gender to be innocent or guilty, tested with either the MGQT or ZOC test, and with either exclusive or nonexclusive control questions.

As subjects arrived at the specified location they were met by an assistant who gave them an Informed Consent Participation Agreement (see Appendix A). Subjects were asked to read the form to themselves as the assistant read it aloud. When the reading was over any questions the subject had were answered and he/she was asked to sign the agreement if he/she wished to continue. After the agreement was signed it was returned to the assistant. The assistant then gave each subject a cassette tape recorder with a set of pre-recorded instructions already in it. The tasks to be carried out by guilty and innocent subjects were on separate cassette tapes; these were maintained by the assistant to ensure the proper instructions were given to each subject. Subjects assigned to

the guilty condition were given the following instructions: "You have been randomly assigned to be a guilty subject. Your task, if you choose to participate, is to proceed from this office and go to the elevator in Baker Hall. Take the elevator to the fifth floor and exit. On that floor is the office of Dr. Horvath. Find it and locate a plastic mail slot on the outside of the door. You are to find a business size "airmail" envelope with red and blue markings around the edge and a large red "X" on each side. Once this envelope is found take it out of the slot and hide it but be careful. Lately there have been many thefts of mail from these locations and the Department of Public Safety has been running security checks. If someone stops you make whatever excuses are necessary and continue on your way. You must be careful not to be caught. Leave the floor and return here by whatever means or route you choose. Remember you have but 15 minutes to complete this portion of your assignment so do not waste time. When you return you will be given additional instructions and asked to submit to a polygraph examination. For no reason whatsoever are you to tell the person conducting the polygraph examination what you have done. Any questions related to the envelope, where it was taken from, or where you have been in the past hour <u>should not</u> (emphasis added) be answered truthfully. Deny all involvement with that "airmail" envelope. If you can successfully accomplish this task, that is if the polygraph examination shows that you were telling the truth then you will be rewarded with the contents of the

"airmail" envelope you took. It is important that you not speak with other students about your participation in this study and that you appear truthful at all times. Good luck, now carry out your instructions."

Innocent subjects were greeted and briefed in the same manner as those assigned to be guilty. After reading and then signing the Informed Consent Agreement, each innocent subject was asked to listen to the taped instructions which explained his/her task as follows: "You have been randomly assigned to participate in this study as an innocent person. Your task, once this tape is completed,is to leave the building and go for a short walk returning here in approximately 15 minutes. During the time you are out walking there will be a crime committed, but you will have no knowledge of what transpired. Once you return you will be asked to submit to a polygraph examination as a possible suspect in the crime because of your being in the area. You are to speak with no one about your participation in this study and to appear as you are, innocent. If the polygraph examination shows you are being truthful there will be a small cash reward in addition to your extra credit. Good luck, now carry out your instructions."

Subjects were told that when their assigned task was completed they were to return to the room where the assistant was located. Guilty subjects were instructed to display and tear open the envelope they took; they were then asked to remove it's contents. In all cases the"airmail" envelope contained three one dollar bills. The money was handed to the

assistant who asked the guilty subject to sign the "stolen" envelope. Innocent subjects simply returned to the room to await further instructions. Each subject was allowed a few minutes to relax, in the room where the assistant was located, before meeting the polygraph examiner.

Polygraph Testing Procedure

Once the polygraph examiner was available the waiting subject was taken to the polygraph suite and introduced to the examiner by the assistant. The experimenter was the only person conducting examinations and in all cases was blind to examinee's guilt or innocence.

Pretest Interview Phase

A pretest interview, similar to that practiced in the field, was conducted first. Each subject was told the reason for the examination, that an envelope with some money in it had been taken from a faculty member's mailbox and we wanted to know if he/she was the person that did it. Background information was then collected using a form (see Appendices B through E) standardized by the type of test and control question used (e.g. ZOC test with exclusive control questions). An explanation of the polygraph procedure and the instrument was then given and the control questions constructed. Excluding the presence or absence of a time bar,

two control questions were the same for all subjects, with an additional control question included for the administration of ZOC tests. The time-bar used to transform nonexclusive control questions into exclusive control questions, in all cases, excluded the last three years of a subject's life (Horvath, 1988). For example, if a subject were twenty years old an exclusive control question would begin "Before the age of 17...". Both exclusive and nonexclusive control questions were formulated in the same manner, adjusting them to elicit an answer of "no" from every subject. Once the control questions were completed the examiner would place the attachments on each subject and review the remaining questions. The format of the ZOC test limited the number of relevant questions used to three. Therefore, only three relevant questions were the same for all subjects in the study, with two additional relevant questions asked subjects administered MGQTs.

Instrumentation Phase

After the pretest interview was finished, the instrumentation phase began. During this portion of the test the polygraph instrument was used to collect physiological data. In the following pages is an explanation of the order in which tests and questions were presented to each subject. These varied according to the type of procedure used, MGQT (Horvath, 1988) or ZOC test (Podlesny & Raskin, 1978). In

either case a stimulation test was included in each examination. This test was used by the examiner to increase the amount of concern felt by a subject, when answering test questions, by telling him/her it would give the examiner an example of the subject's responses when telling the truth and when telling a lie. Each subject was instructed to choose a number from 3 to 6 and tell it to the examiner (Podlesny & Raskin, 1978). The number was then recorded on a piece of paper by the examiner as the subject watched. Subjects were then told each number between 1 and 7 would be used in a sentence asking " Did you tell me you chose the number. . .?" He/she was told to answer "no" each time the question was asked, even when it included the number they chose. Consequently, the subject would be lying to one question and telling the truth to the others, thus giving the examiner the examples he needed. Once this test was finished the examiner would review the chart briefly. Each subject was told the responses evoked by their truthful and deceptive answers were very clear and the differences easy to see whether this was true or not.

Administration of the MGQT.

The Modified General Question Test (MGQT) was given in a manner consistent with it's employment in the field using four irrelevant questions, five relevant questions, and two control questions (Horvath, 1988). The control questions used with

this procedure are generally of the nonexclusive type. When necessary these questions were prefaced with a time-bar to transform them into exclusive control questions. Excluding the time-bar modification, the questions, and their sequence, were the same for all subjects given an MGQT and replicated as close as possible the procedure used by Horvath (1988). The question sequence (using nonexclusive controls) asked at the time of testing was:

1. Do they call you [first name]?
2. Are you over [ ] years of age?
3. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?
4. Do you live in the United States?
5. Did you take that envelope containing three dollars?
6. Did you ever take something that did not belong to you?
7. Did you ever go to school?
8. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?
9. Did you write your name on that airmail envelope taken from Dr. Horvath's mail slot?
10. Did you ever tell a lie about something important?
11. Were you assigned to be a guilty person in this research?

In this list of questions, questions #1,2,4, & 7 were

irrelevant questions; questions #3,5,8,9, and 11 were relevant questions; questions #6 and 10 were control questions. All subjects were given a total of four tests. The first test was a reading of the above questions, in sequence, at about 20 second intervals. The examinee was instructed to respond verbally with either the answer "yes" or "no". The second test conducted was the stimulation test described above. Test three was a repetition of test one. The final test was a mixed question test in which the question order was changed. The question sequence (7, 4, 11, 8, 10, 1, 3, 6, 2, 5, 10, 9) for this "mixed question" test was the same for all subjects given an MGQT. When the mixed question test was finished each subject was instructed to report back to the assistant for further instructions.

Administration of the ZOC Test.

The Zone Comparison Test (ZOC) was also administered in a manner consistent with it's employment in the field using two irrelevant questions, a symptomatic question, a sacrifice relevant question, three relevant questions, and three control questions (Podlesny & Raskin, 1978). The control questions used with this procedure are generally of the exclusive type. When appropriate time-bars were eliminated to transform these questions into nonexclusive control questions. Excluding the time-bar modification, the questions, and their sequence, were the same for all subjects administered a ZOC test. The

question sequence (using exclusive controls) asked at the time of testing was:

1. Is your name [first name]?
2. Are you afraid I'll ask you a question we have not reviewed?
3. Do you intend to answer truthfully each question about the stolen envelope?
4. Before the age of [ ] did you ever take something that did not belong to you?
5. Did you take that envelope containing three dollars?
6. Before the age of [ ] did you ever tell a lie about something important?
7. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?
8. Are you now in Michigan?
9. Before the age of [ ] did you ever tell a lie to a person in authority?
10. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?

In this question sequence, questions #1 and 8 were irrelevant questions; question #2 was a symptomatic question; question #3 was a sacrifice relevant question; questions #5, 7, and 10 were relevant questions; questions #4, 6, and 9 were control questions. All subjects tested with the ZOC procedure were given a total of four tests. The first test in the

sequence was the stimulation test described earlier. The three remaining tests used the questions given above, but their order of presentation was changed from test to test. For each question examinees were instructed to respond verbally with either the answer "yes" or "no". The second test conducted was a reading of the questions in the sequence given above. The third test was administered with the questions in the following order: 8, 2, 3, 9, 5, 4, 7, 1, 6 and 10. For the final test the question order was: 1, 2, 3, 4, 10, 6, 5, 8, 9 and 7. When testing was completed each subject given the ZOC test was also told to report back to the assistant for further instructions.

Concern Ratings

The theory on which the control question test was founded is the assumption that people will fear that which holds the greatest immediate threat to their well being. Consequently, during a polygraph examination one would expect guilty subjects to have more concern for and display greater physiological arousal to relevant questions, while innocent subjects would have more concern for and show greater physiological arousal to the control questions. To test this theory, after each subject was finished with polygraph testing, he/she was asked to complete a short questionnaire (see Appendix F through I) listing only the relevant and control questions asked during the polygraph examination. To

complete the questionnaire subjects were instructed to rate their degree of concern for each relevant and control question, using a four point scale of 1, none, to 4, a lot of concern. Once finished with the questionnaire subjects returned it to the assistant and were excused.

Control, Use and Evaluation of Response Data

After each subject left, the assistant returned to the polygraph suite to collect the completed forms and polygraph charts made during the examination. Next, the assistant created a file and assigned a subject number to it. This number was placed on each form and polygraph chart, then recorded in a logbook for later reference. The charts, marked only with this number, were then separated from the file and kept together for later numerical scoring. The assistant maintained control of the polygraph charts until all of the subjects were examined. The charts were then scored by the examiner only for the purpose of determining who would receive a monetary reward. Rewards were given to subjects whose charts yielded total examination numerical scores of +6 or greater. When this task was finished the charts, along with the examiner's scores, were given to the assistant who notified the appropriate subjects. The assistant told these subjects to contact a secretary in the School of Criminal Justice who would make the necessary payments. The amount of each reward was $3.00, a sum equal to that found in each of

the envelopes.

A majority of the studies (see Chapter II) examining the psychophysiological detection of deception, specifically control question testing, have used "blind" scoring to estimate the accuracy of the technique. The importance of blind scoring was emphasized by Lykken (1981) who said that assessment of a polygraph test:

> "can be done only by having the charts scored independently by polygraphers who did not administer the tests in question and who are "blind" with respect to all knowledge of the subjects except for whatever information they can glean from the polygrams" (p. 74).

To that end an independent evaluator was solicited to numerically score (as described in the introduction), the polygraph charts produced for this study. For the purpose of scoring, a scale of ± 3 (Horvath, 1988; Podlesny & Raskin, 1978) was used in assigning scores for each parameter, and an inconclusive zone from of ± 5 (Podlesny & Raskin, 1978; Raskin & Hare, 1978; Patrick & Iacono, 1989) used to classify subjects as innocent or guilty. The charts were sent to the evaluator by mail in groups of ten to twenty with standardized score sheets (see Appendix J and K). As a group of charts was scored and returned another was sent to replace it. This process was repeated until all of the polygraph charts were scored. The control/relevant pairs the evaluator used for comparative analysis were different depending on the variation of control question test, ZOC or MGQT. ZOC tests were scored by comparing each relevant question to the adjacent control on

the left on each of the last three charts, allowing each control to be compared with each relevant question. MGQT charts were scored using the same comparisons Horvath (1988) used for his study. For the first and third charts these were 3/6, 5/6, 8/6, 9/10 and 10/11. The last chart, a mixed question test, used the following comparisons: 3/6, 5/10, 8/10, 9/10 and 11/10.

How Many Relevant Questions: Three vs. Five

The questions used for the ZOC and MGQT examinations varied by type and number as illustrated in Table 4 (see p. 77). The ZOC procedure used three relevant questions,

---

Insert Table 4 about here

---

while the MGQT used five. The effect of using a greater (5) or smaller (3) number of relevant questions in the same context was unknown. Therefore, the first three relevant questions used with the MGQT procedure were the same as the three relevant questions used with the ZOC procedure. In doing so, the scores given the last two relevant questions, used with the MGQT procedure, could be dropped and scores recalculated. Analyses were then repeated to examine the effect(s) of using three relevant questions versus five.

Table 4
Question Numbers and Type of Question Used For Each Variation of Control Question Test

Question Number	Test Variation: Zone Comparison	Test Variation: Modified General Question Test
1.	Irrelevant	Irrelevant
2.	Symptomatic	Irrelevant
3.	Sacrifice Relevant	Relevant [E]
4.	Control [A]	Irrelevant
5.	Relevant [B]	Relevant [B]
6.	Control [C]	Control [A]
7.	Relevant [D]	Irrelevant
8.	Irrelevant	Relevant [D]
9.	Control	Relevant
10.	Relevant [E]	Control [C]
11.	[Blank]	Relevant

[A] Same as MGQT question 6.
[B] Same as MGQT question 5.
[C] Same as MGQT question 10.
[D] Same as MGQT question 8.
[E] Same as MGQT question 3.

Analysis of Results

Analysis of the data was accomplished using both the SPSSX and SPSSPC+ computer programs. CROSSTABS, ANOVA, MANOVA and CORRELATION commands were employed, using a .05 rejection region, to determine significant relationships and to measure the strength of any association(s) between the independent and dependent variables.

The independent variables manipulated in this study were the variation of control question polygraph test, Zone Comparison (ZOC) test or Modified General Question Test (MGQT), the type of control question used, exclusive or nonexclusive, gender and subject status, guilty or innocent. The dependent variables were the number of correct, wrong and inconclusive decisions rendered by the blind evaluator, the numerical scores derived from his scoring of the polygraph charts and the concern ratings given by subjects to the relevant and control questions they were asked.

CHAPTER IV

Results

An analysis of the polygraph charts was conducted by the original examiner after all of the examinations were administered. This was necessitated by a need to make monetary rewards within a reasonable length of time. Those analyses resulted in 58% correct, 18% wrong and 24% inconclusive decisions. When inconclusive results were excluded decisions were 76% correct. The design and conduct of this study did not include techniques to measure, nor identify, the possible effect of extra-polygraphic cues or other knowledge available to the examiner while scoring the polygraph charts. For example, the knowledge that only college students were used as subjects, the information gathered from each subject in the one to one interviews, knowing the base rate of guilt and that the polygraph charts were the result of a laboratory experiment, are all factors posited to detract from the quality of the data collected (Lykken, 1981; Iacono & Patrick, 1988). Therefore, the results of the analysis on the examiner's scores are offered for information only and can be found in the Appendices L, M, and N.

Accuracy of Evaluator Decisions

The accuracy of the blind evaluator's decisions, using 3 relevant questions for ZOC tests and 5 relevant questions for MGQTs, is summarized in Table 5 (see p. 81). Across all

---

Insert Table 5 about here

---

treatment conditions the evaluator achieved 61% correct, 22% wrong and 17% inconclusive decisions. When inconclusive results were excluded decisions were 74% correct.

The CROSSTABS procedure (Norusis & SPSS Inc., 1988) taken from SPSS/PC+ was used to calculate a series of chi-square statistics to test for any significant relationship(s) between the accuracy of the evaluator's decisions and the independent variables. Analysis showed the accuracy of those decisions was not significantly related to the variation of control test used, ZOC or MGQT; subject status, innocent or guilty; or gender, male or female. The accuracy of the evaluator's decisions however, was significantly related to the type of control question used, exclusive or nonexclusive, $\underline{X}^2(2, \underline{N} = 120) = 13.17$, $\underline{p}=.0014$. Figure 1 (see p. 83) illustrates the distribution of the evaluator's decisions, including inconclusives, for each type of control question. The relationship between the accuracy of the evaluator's decisions

Table 5

Distribution of Evaluator's Decisions Based On All Relevant Questions

	Decisions By Number					
	No. Correct	No. Wrong	No. Inconc	No. False Positive	No. False Negative	Correct No Inconcl
Control Question						
Exclusive	27	17	16	13	4	27
Nonexclusive	46	9	5	4	5	46
Variation of Test						
ZOC	39	11	10	8	3	39
MGQT	34	15	11	9	6	34
Gender						
Female	23	10	7	5	5	23
Male	50	16	14	12	4	50

	Decisions By Percentage					
	% Correct	% Wrong	% Inconc	% False Positive	% False Negative	Correct No Inconcl
Control Question						
Exclusive	45%	29%	26%	22%	07%	61%
Nonexclusive	77%	15%	08%	07%	08%	84%
Variation of Test						
ZOC	65%	18%	17%	13%	05%	78%
MGQT	57%	25%	18%	15%	10%	69%
Gender						
Female	58%	26%	16%	13%	13%	70%
Male	63%	20%	17%	15%	05%	76%

and the type of control question used was examined further.

Insert Figure 1 about here

The data for subjects tested using nonexclusive control questions was subjected to analysis first. Results showed that the differences in the accuracy of the evaluator's decisions was not significantly related to subject status, guilty or innocent; the type of CQ test used, ZOC or MGQT; or gender, male or female.

Next, the data for subjects tested using exclusive control questions was analyzed. Results of chi-square tests showed the accuracy of the evaluator's decisions was significantly related to subject status, innocent or guilty, $\underline{X}^2(2, \underline{N} = 60) = 8.79$, $\underline{p}=.0124$, but not significantly related to the type of test used or subject gender.

Additional analyses were performed to examine the relationship between subject status, guilty or innocent, and the type of control question used for testing. Results showed the relationship between the accuracy of the evaluator's decisions for subjects assigned to the guilty condition and the type of control question used for testing was not significant, $\underline{X}^2(2, \underline{N} = 60) = 3.75$, $\underline{p}=.1532$. In other words, the accuracy of the evaluator's decisions classifying the guilty as guilty did not differ significantly by the type of

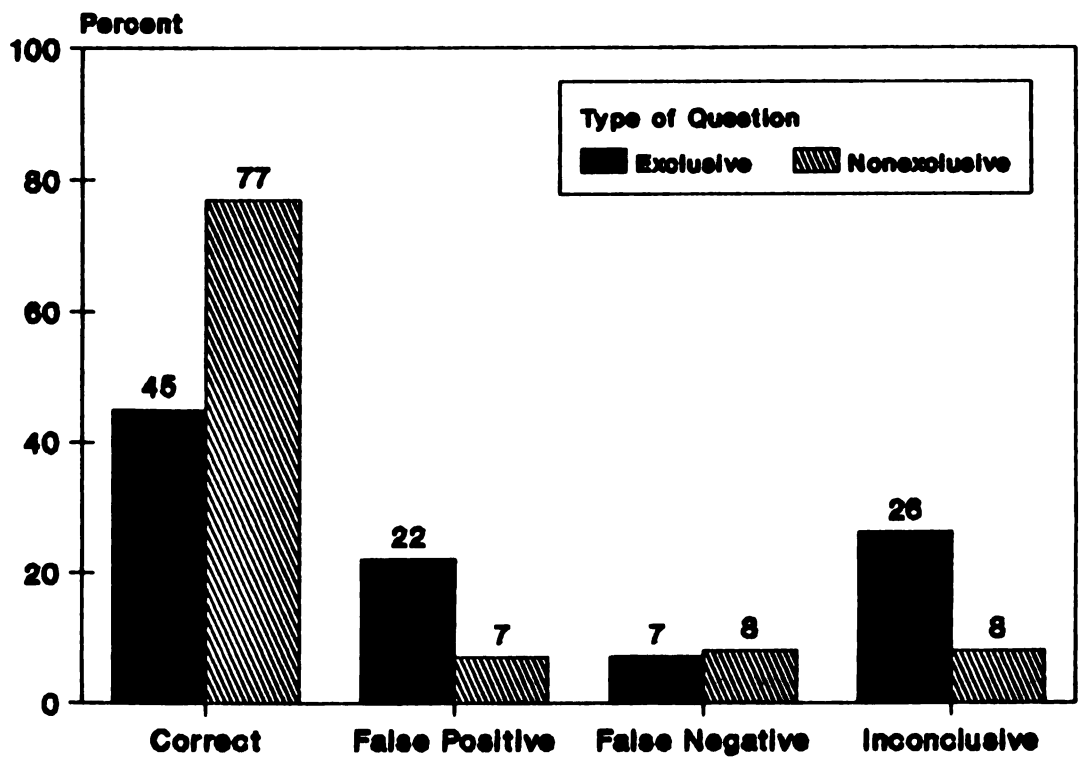
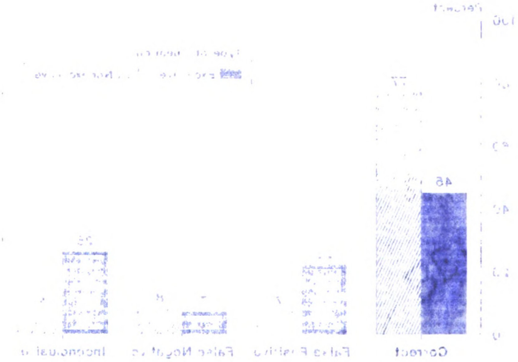


Figure 1

Distribution Of Evaluator's Decisions By

Type of Control Question With Inconclusives


Percent
100
80
60
40
20
0
Type of ...
77
46
25
Correct
False Positive
False Negative
Inconclusive

[illegible]

[illegible]

[illegible]

control question used for testing. This finding is evidenced by the proportion of false negative errors (guilty called innocent) that resulted using exclusive and nonexclusive control questions (see Figures 1 and 2, pgs. 81 and 84).

Results showed the accuracy of the evaluator's decisions classifying subjects in an innocent condition was significantly related to the type of control question used for testing, $\underline{X}^2(2, \underline{N} = 60) = 13.31$, $\underline{p}=.0013$. The accuracy of classifying the innocent as innocent, when nonexclusive control questions were used for testing, was equal to the accuracy shown for classifying the guilty as guilty using either exclusive or nonexclusive control questions. This wasn't true when exclusive control questions were used. Results showed that tests using exclusive control questions produced a significantly lower degree of accuracy for the classification of the innocent as innocent. Consequently, the proportion of false positive errors (innocent called guilty) also was larger. These findings are illustrated in Figures 1 and 2 (pgs. 83 and 86).

The literature review demonstrated that the accuracy of polygraphic decisions with and without an inconclusive zone is also of interest. Consequently, inconclusive decisions were excluded and the accuracy of only the evaluator's correct and wrong decisions was examined to see if the pattern of significant relationship(s) with independent variable(s) continued. Thus a series of 2 X 2 crosstabulations, with one (1) degree of freedom, was conducted. For these analyses a

Yates correction was employed in calculating chi-square statistics in spite of the controversy surrounding the necessity for it's use (Frude, 1987; Hays, 1981; Norusis & SPSS Inc., 1988).

This series of analyses showed that the accuracy of the evaluator's decisions was again significantly related only to the type of control question used for testing, exclusive or nonexclusive, $\underline{X}^2(1, \underline{N} = 99) = 4.17$, $\underline{p} = .0410$. Additional calculations were again conducted to test the relationship between the accuracy of the evaluator's decisions and the control question variable. These analyses were accomplished by first examining the data for subjects assigned to the guilty condition and then the data for the innocent. It was found that the accuracy of decisions for guilty subjects wasn't related significantly to the type of control question used, $\underline{X}^2(1, \underline{N} = 48) = .000$, $p = 1.000$, but the accuracy of decisions for innocent subjects and the type of control question used was once again related at a level exceeding chance, $\underline{X}^2(1, \underline{N} = 51) = 8.33$, $\underline{p} = .0037$. Figure 2 (see p. 86) illustrates the accuracy of the evaluator's decisions, excluding inconclusive results, using exclusive and nonexclusive control questions.

Insert Figure 2 about here

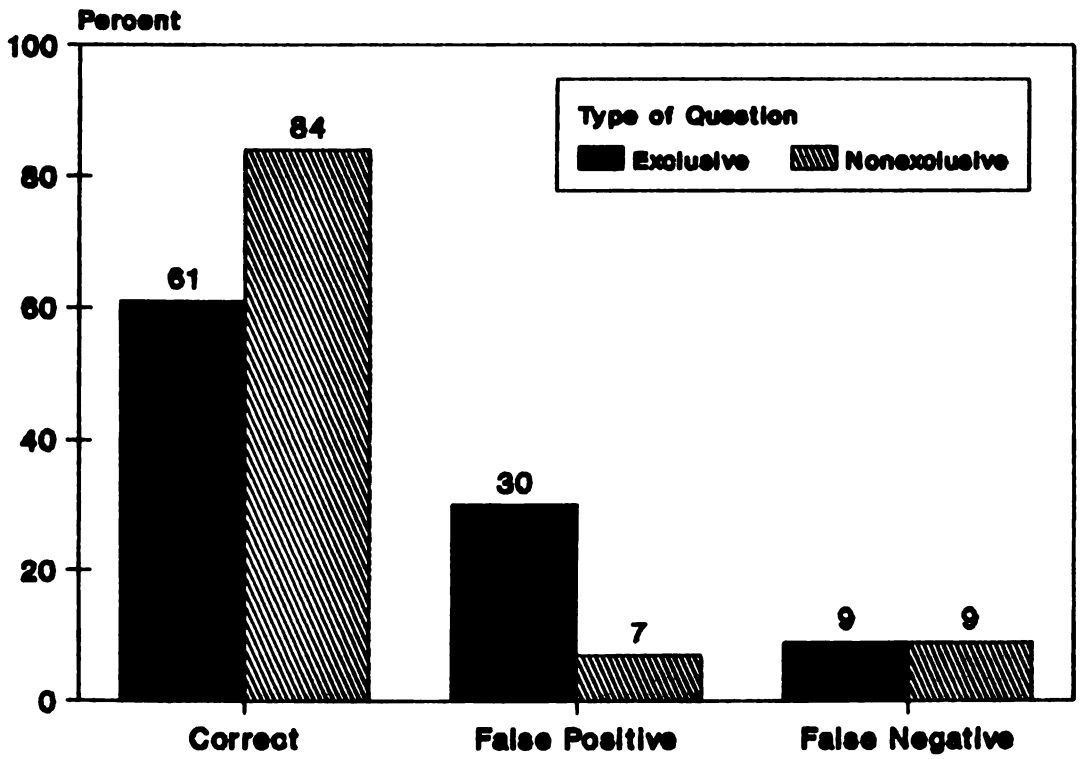
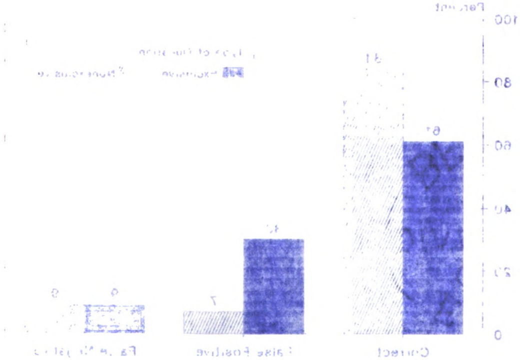


Figure 2

Distribution Of Evaluator's Decisions By

Type of Control Question Without Inconclusives


Percent
100
80
60
40
20
0
Type of conclusion
Exclusive
Nonexclusive
81
61
32
7
6
6
Correct
False Positive
False Negative

Results that are based on chi-square statistics alone are often difficult to interpret and many times misleading (Reynolds, 1984). "By itself, chi-square helps us only to decide whether our variables are independent or related. It does not tell us how strongly they are related" (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975, p. 224). The use of chi-square statistics in this study showed that in judging the status of innocent subjects, the evaluator's decisions were significantly related to the type of control question used. However, the strength of the relationship was unknown. To highlight this point further the reader's attention is directed to Figures 1 (see p. 83) and 2 (see p. 86). Even though the differences appear to be large, one can't determine from the chi-square statistics alone the meaning of such differences. Consequently, statistics used to measure the strength of association (Hays, 1981; Norusis, 1983; Reynolds, 1984) were calculated to examine this point. The measures selected were Kendall's Tau-c, for rectangular tables, and Tau-b, for square tables, because ". . .they will generally have lower numerical values than gamma" (Nie et al., 1975, p. 228), providing a more conservative estimate of the strength of any relationship.

Analysis showed the strength of the relationship between the accuracy of the evaluator's decisions, including inconclusives, and the two types of control question was tau $c = .33$, $p<.001$. Controlling for status, further analyses showed the strength of the relationship between the accuracy

of the evaluator's classifications and the two types of control question was, tau c = .47, p<.001, for innocent subjects. The strength of the relationship between the type of control question and accuracy of the evaluator's judgments regarding the status of guilty subjects was, tau c = .20, p>.05.

The exclusion of inconclusive results did not alter the trend seen in the results. The strength of the relationship between the type of control question used and the accuracy of the evaluator's decisions was, tau b = .23, p<.012. When status is controlled, the strength of the association between accuracy of the evaluator's decisions and the two types of control question again was, tau b = .45, p<.001, but only when classifying innocent subjects and not guilty subjects, tau b = -.06, p>.34.

It could be argued that the use of more relevant questions for MGQT examinations may in some way have contributed to the differences in accuracy. To explore this possibility, the effect(s) of using fewer relevant questions, as is the practice with ZOC examinations, was explored by dropping the two extra relevant questions that were used for MGQT tests. The number of correct, wrong and inconclusive decisions were then recalculated and the analyses repeated. The distribution of the evaluator's decisions is summarized in Table 6 (see p. 90). Overall, the use of just three relevant questions produced a smaller proportion of correct (58%) and wrong (19%) classifications, consequently resulting in a

Insert Table 6 about here

larger proportion of inconclusive decisions (23%). When inconclusive results were excluded there was a small increase in the accuracy of classification (76%) across all treatment conditions.

This data was then analyzed using the CROSSTABS procedure. Analysis showed that when fewer relevant questions were employed, the accuracy of classifying subjects as guilty or innocent was still significantly related to whether exclusive or nonexclusive control question were used for testing, $X^2(2, N = 120) = 11.46$, $p=.0032$, and that the strength of the relationship was, tau c = .29, $p<.001$. Again, the results showed there was a significant relationship between the accuracy of classifying subjects assigned to be innocent as innocent and the type of control question, exclusive or nonexclusive, used for testing, $X^2(2, N = 60) = 11.11$, $p=.0039$, tau c = .29, $p<.017$. At the same, time the accuracy of classifying guilty subjects as guilty did not differ significantly using either exclusive or nonexclusive control questions, $X^2(2, N = 60) = 5.03$, $p=.0810$. These results showed that the proportions of false positive error and inconclusive decisions was significantly larger when innocent subjects were tested using exclusive control questions. Figure 3 (see p. 92) illustrates the distribution

Table 6

Distribution of Evaluator's Decisions Based on the Same Three Relevant Questions in the Two Procedures

	Decisions By Number					
	No. Correct	No. Wrong	No. Inconc	No. False Positive	No. False Negative	Correct No Inconcl
Control Question						
Exclusive	26	16	18	12	4	27
Nonexclusive	44	6	10	2	4	46
Variation of Test						
ZOC	39	11	10	8	3	39
MGQT	31	11	18	6	5	34
Gender						
Female	24	9	7	4	5	23
Male	46	13	21	10	3	50

	Decisions By Percentage					
	% Correct	% Wrong	% Inconc	% False Positive	% False Negative	Correct No Inconcl
Control Question						
Exclusive	43%	27%	30%	20%	07%	62%
Nonexclusive	73%	10%	17%	03%	07%	88%
Variation of Test						
ZOC	65%	18%	17%	13%	05%	78%
MGQT	52%	18%	30%	10%	08%	74%
Gender						
Female	60%	23%	17%	10%	13%	73%
Male	58%	17%	25%	13%	04%	78%

Insert Figure 3 about here

of the evaluator's decisions for innocent and guilty subjects using exclusive and nonexclusive control questions.

When inconclusive results were excluded the relationship between accuracy and the type of control question used remained significant, (Yates-corrected) $X^2(1, N = 92) = 7.17$, $p=.0074$, with the strength of the association, tau b = .30, $p<.002$, but as before only for the classification of the innocent, (Yates-corrected) $X^2(1, N = 45) = 8.99$, $p=.0027$, tau b = .50, $p<.000$, and not the guilty, (Yates-corrected) $X^2(1, N = 47) = .005$, $p=.9401$, tau b = .07, $p=.3218$.

Field Numerical Scoring

Statistical analyses were conducted on the examination total numerical scores provided by the blind evaluator to each of the questions, 5 for MGQT and 3 for ZOC tests, to determine whether any effects produced by the different independent variables were significant. Results of a four way Analysis of Variance (ANOVA), type of test (ZOC/MGQT), status (Guilty/ Innocent), gender (Male/Female), and type of control question (Nonexclusive/Exclusive) showed only a significant effect [$F(1/104) = 32.156$, $p<.000$] for status (Guilty/ Innocent), with mean total numerical scores of guilty = -10.63 and

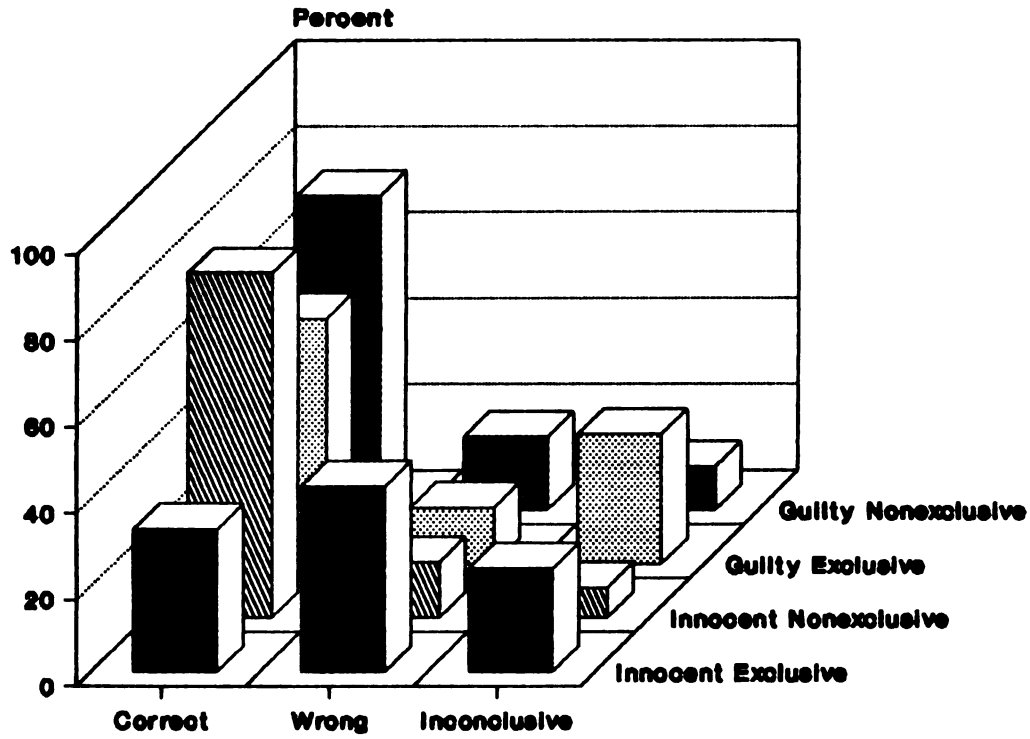


Figure 3

Distribution Of Evaluator's Decisions

For Innocent and Guilty By Type of Control Question

innocent = 6.83. Table 7 (see p. 94) displays a summary of the ANOVA and shows only one other significant finding, an interaction between status (innocent/guilty) and the type of

---

Insert Table 7 about here

---

control question used (Nonexclusive/Exclusive) [$F(1/104) = 5.413$, $p=.022$]. Examination of the mean total numerical scores for each control question group (see Figure 4, p. 95) showed that innocent subjects tested with exclusive control

---

Insert Figure 4 about here

---

questions had a lower mean score compared to the mean score produced by innocent subjects tested with nonexclusive controls. These findings show that exclusive control questions did not elicit the same level of concern from innocent subjects as did nonexclusive control questions. Although the difference was not statistically significant, the mean total numerical score for guilty subjects tested using nonexclusive control questions was more extreme compared to the mean total numerical score for guilty subjects tested using exclusive control questions. On the surface these

Table 7

ANOVA Summary For Examination Total Numerical Scores:

Using All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	9498.304	4	2374.576	8.343	.000*
Type Test					
ZOC/MGQT (A)	10.800	1	10.800	.038	.846
Status					
Inn/Guilty (B)	9152.533	1	9152.533	32.156	.000*
Gender					
Male/Female (C)	105.338	1	105.338	.370	.544
Control					
Excl/Nonexc (D)	229.633	1	229.633	.807	.371
2 Way Interactions	1871.746	6	311.958	1.096	.370
A x B	8.533	1	8.533	.030	.863
A x C	57.038	1	57.038	.200	.655
A x D	12.033	1	12.033	.042	.837
B x C	110.704	1	110.704	.389	.534
B x D	1540.833	1	1540.833	5.413	.022*
C x D	142.604	1	142.604	.501	.481
3 Way Interactions	551.513	4	137.878	.484	.747
A x B x C	242.004	1	242.004	.850	.359
A x B x D	86.700	1	86.700	.305	.582
A x C x D	47.704	1	47.704	.168	.683
B x C x D	175.104	1	175.104	.615	.435
4 Way Interactions	643.538	1	643.538	2.261	.136
A x B x C x D	643.537	1	643.537	2.261	.136

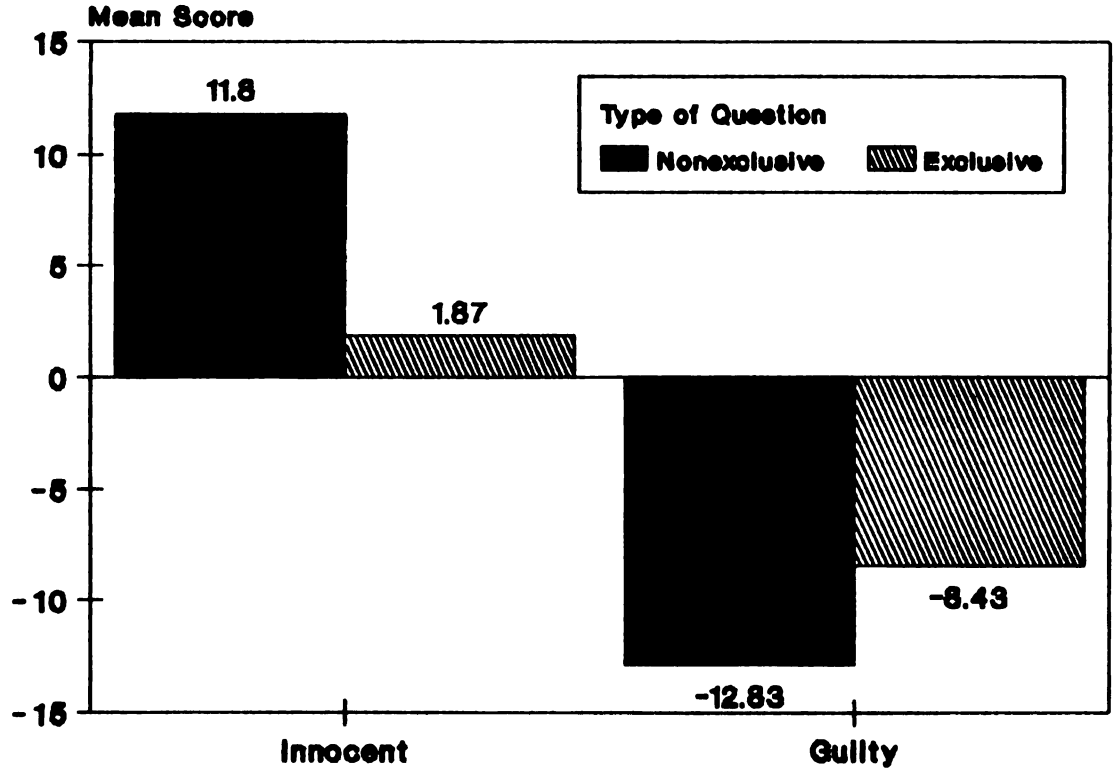


Figure 4

Total Mean Numerical Scores

For Guilty and Innocent By Type of Control Question

Using All Relevant Questions

findings suggest that exclusive control questions, perhaps due to their unique phrasing, are not as significant for the innocent and too distracting for the guilty, when compared to nonexclusive control questions in the same context. There were no other significant effects noted.

It is possible that the use of five relevant questions in one half of all the examinations conducted and only three relevant questions in the other half could have skewed the mean total numerical scores. Therefore, the two relevant questions that had been used for MGQT examinations and not used for ZOC examinations were dropped and the total examination scores recalculated. The four way ANOVA described above, type of test (ZOC/MGQT), status (Guilty/ Innocent), gender (Male/Female), and type of control question (Nonexclusive/Exclusive), was repeated to see once again if there were any significant differences. Table 8 (see p. 97)

Insert Table 8 about here

---

is a summary of the ANOVA and shows that using fewer relevant questions to calculate the examination total scores did not effect the pattern of significant results. The only main effect again shown to be significant was subject status (guilty/innocent) [$F(1/104) = 40.661$, $p<.000$]; a result once more moderated by a significant interaction [$F(1/104) = 8.967$,

Table 8

ANOVA Summary For Examination Total Numerical Scores:

Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	7247.296	4	1811.824	10.794	.000*
Type Test					
ZOC/MGQT (A)	49.408	1	49.408	.294	.589
Status					
Inn/Guilty (B)	6825.208	1	6825.208	40.661	.000*
Gender					
Male/Female (C)	192.604	1	192.604	1.147	.287
Control					
Excl/Nonexc (D)	180.075	1	180.075	1.073	.303
2 Way Interactions	2040.137	6	340.023	2.026	.069
A x B	255.208	1	255.208	1.520	.220
A x C	124.704	1	124.704	.743	.391
A x D	27.075	1	27.075	.161	.689
B x C	14.504	1	14.504	.086	.769
B x D	1505.208	1	1505.208	8.967	.003*
C x D	113.438	1	113.438	.676	.413
3 Way Interactions	506.654	4	126.664	.755	.557
A x B x C	78.204	1	78.204	.466	.496
A x B x D	95.408	1	95.408	.568	.453
A x C x D	31.538	1	31.538	.188	.666
B x C x D	301.504	1	301.504	1.796	.183
4 Way Interactions	451.004	1	451.004	2.687	.104
A x B x C x D	451.004	1	451.004	2.687	.104

p=.003] with the type of control question, exclusive or nonexclusive, used for testing. The mean total numerical scores for each control question group were examined. Review showed when total scores were calculated using only three relevant questions, all control question groups had slightly lower mean total scores (see Figure 5 p. 99) and the large difference between mean total scores for innocent subjects

---

Insert Figure 5 about here

---

innocent subjects tested with exclusive and nonexclusive control questions persisted.

Field Scoring of Physiological Measures

Using all 5 relevant questions on MGQT tests and 3 relevant questions on ZOC tests, the effectiveness of each physiological measure was evaluated by subjecting the examination total numerical scores for each component to analysis. A series of ANOVA's was used to see if any of the four independent variables, type of test (ZOC/MGQT), status (Guilty/Innocent), gender (Male/ Female), and type of control question (Nonexclusive/ Exclusive), produced significant effect(s). Those analyses showed there were significant effects. However, these varied for each physiological

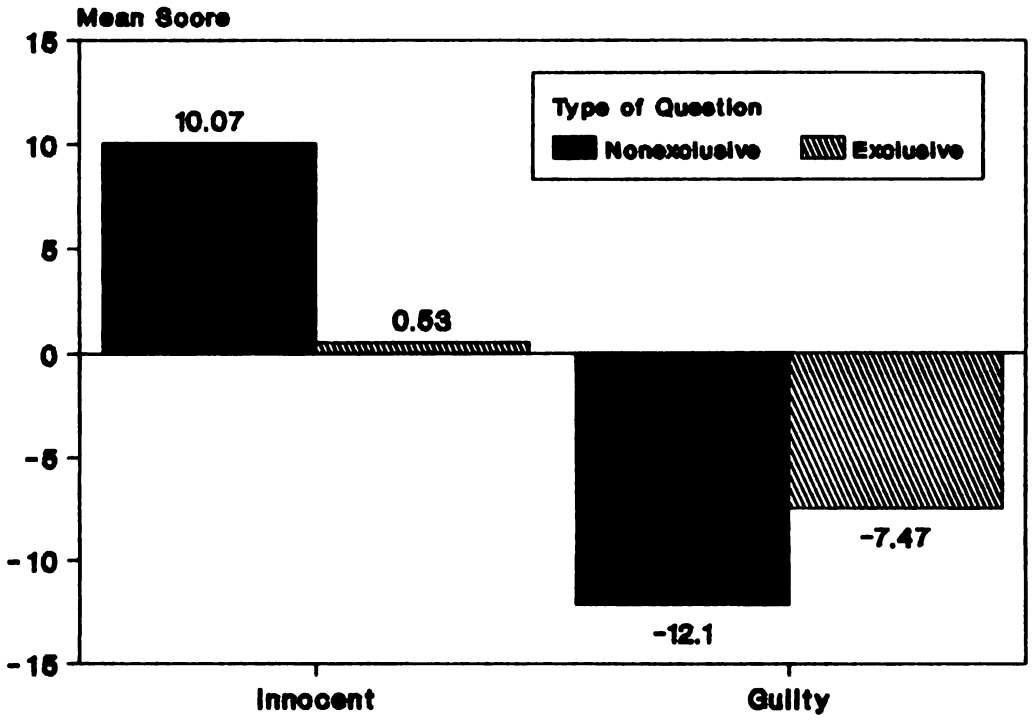
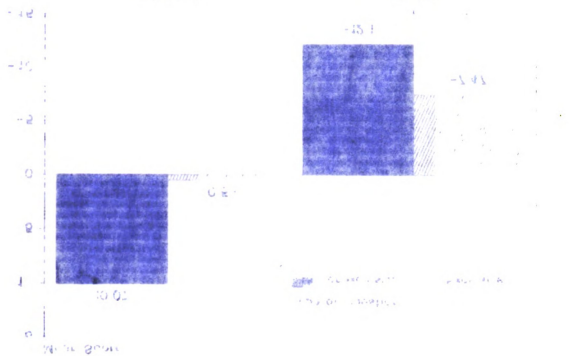


Figure 5

Total Mean Numerical Scores

For Guilty and Innocent By Type of Control Question

Using Just Three Relevant Questions



measure. The results of these analyses are offered, by measure, in the following order, upper pneumograph, lower pneumograph, SRR, and cardio.

Table 9 (see p. 101) summarizes the ANOVA conducted on

Insert Table 9 about here

---

upper pneumograph (thoracic respiration) scores and shows a significant effect for status (Guilty/Innocent) [F(1/104) = 7.099, p<.009], with a mean score of [M=.92] for innocent subjects and [M=-1.60] for guilty. Results showed there were also three significant interactions. The first was a status (Guilty/Innocent) X type of control question (Nonexclusive/ Exclusive) interaction [F(1/104) = 6.546, p=.012]; both innocent and guilty subjects tested with nonexclusive control questions produced mean numerical scores [M=2.50 & M=-2.43] that were more extreme than those produced by innocent and guilty subjects [M=-.67 & -.77] tested with exclusive control questions. The next significant result was a type of test (ZOC/MGQT) X status (Guilty/Innocent) X gender (Male/ Female) interaction [F(1/104) = 4.499, p=.036]. Examination of the mean scores showed that all subjects, both male and female, tested with ZOC procedures, all males tested with MGQT procedures, and innocent females tested with MGQT procedures, produced mean scores (see Table, 10 p. 103) that were in the

Table 9

ANOVA Summary For Upper Pnuemograph Total Numerical Scores:

Using All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	284.625	4	71.156	2.658	.037*
Type Test					
ZOC/MGQT (A)	21.675	1	21.675	.810	.370
Status					
Inn/Guilty (B)	190.008	1	190.008	7.099	.009*
Gender					
Male/Female (C)	56.067	1	56.067	2.095	.151
Control					
Excl/Nonexc (D)	16.875	1	16.875	.630	.429
2 Way Interactions	320.842	6	53.474	1.998	.073
A x B	.008	1	.008	.000	.986
A x C	60.000	1	60.000	2.242	.137
A x D	.208	1	.208	.008	.930
B x C	70.427	1	70.417	2.631	.108
B x D	175.208	1	175.208	6.546	.012*
C x D	15.000	1	15.000	.560	.456
3 Way Interactions	262.908	4	65.727	2.455	.050*
A x B x C	120.417	1	120.417	4.499	.036*
A x B x D	10.208	1	10.208	.381	.538
A x C x D	.267	1	.267	.010	.921
B x C x D	132.017	1	132.017	4.932	.029*
4 Way Interactions	.817	1	.817	.031	.862
A x B x C x D	.817	1	.817	.031	.862

Insert Table 10 about here

expected direction (positive for innocent and negative for guilty). However, guilty females tested with MGQT procedures yielded a mean score that was about the same magnitude as that produced by guilty males tested with the same procedure, but in the wrong (innocent) direction. The last result found significant was a status (Guilty/ Innocent) X gender (Male/ Female) X type of control question (Nonexclusive/Exclusive) interaction [$F(1/104) = 4.932$, $p=.029$]. The mean scores for each group (see Table, 11 p. 104) showed that males tested with nonexclusive control questions and females tested with

Insert Table 11 about here

exclusive control questions, to a much lesser degree, yielded mean scores in the correct directions (positive for innocent and negative for guilty). At the same time guilty females tested with nonexclusive control questions produced a mean score that was in the wrong (innocent) direction, while innocent males tested with exclusive control questions produced a mean score that was in the direction of guilt.

The effects of using only three relevant questions on

Table 10

Upper Pneumograph Mean Scores for Type of Test X Status X Gender Interaction: Using All Relevant Questions (5 in MGQT and 3 in ZOC)

	Gender	
Status	Male	Female
	MGQT	
Innocent	1.60	0.80
Guilty	-3.40	3.30
	ZOC	
Innocent	0.35	0.80
Guilty	-1.85	-2.40

p=.036

Table 11

Upper Pneumograph Mean Scores for Status X Gender X Type of Control Question Interaction: Using All Relevant Questions (5 in MGQT and 3 in ZOC)

	Gender	
Status	Male	Female
	Nonexclusive	
Innocent	3.05	1.40
Guilty	-4.45	1.60
	Exclusive	
Innocent	-1.10	0.20
Guilty	-0.80	-0.70

p=.029

component total mean scores was examined by again dropping the two additional relevant questions that were used in MGQT examinations and not used in ZOC examinations. The component total scores were recalculated and subjected to analysis. Table 12 (see p. 106) shows a summary of the ANOVA conducted on these upper pneumograph (thoracic respiration) scores.

Insert Table 12 about here

The analysis showed a significant main effect for status [F(1/104) = 6.676, p=.011] with a mean score of [M=.50] for innocent subjects and [M=-1.57] for guilty. Results showed also that with three relevant questions there were only two significant interactions. The first of these was again a status (Guilty/Innocent) X type of control question (Nonexclusive/ Exclusive) interaction [F(1/104) = 8.510, p=.004]; both innocent and guilty subjects tested with nonexclusive control questions produced mean numerical scores [M=2.00 & M=-2.40] that were more extreme and in predicted directions (positive for innocent and negative for guilty). The mean scores produced by subjects [M=-1.00 & -.73] tested with exclusive control questions were smaller and for innocent subjects in the wrong (guilty) direction relative those produced by subjects tested with nonexclusive control questions. The last result found significant was a three way,

Table 12

ANOVA Summary For Upper Pneumograph Total Numerical Scores:

Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	189.667	4	47.417	2.470	.049*
Type Test					
ZOC/MGQT (A)	6.533	1	6.533	.340	.561
Status					
Inn/Guilty (B)	128.133	1	128.133	6.676	.011*
Gender					
Male/Female (C)	41.667	1	41.667	2.171	.144
Control					
Excl/Nonexc (D)	13.333	1	13.333	.695	.406
2 Way Interactions	254.617	6	42.436	2.211	.048*
A x B	6.533	1	6.533	.340	.561
A x C	45.067	1	45.067	2.348	.128
A x D	.000	1	.000	.000	1.000
B x C	20.147	1	20.417	1.064	.305
B x D	163.333	1	163.333	8.510	.004*
C x D	19.267	1	19.267	1.004	.319
3 Way Interactions	152.567	4	38.142	1.987	.102
A x B x C	50.417	1	50.417	2.627	.108
A x B x D	13.333	1	13.333	.695	.406
A x C x D	.000	1	.000	.000	1.000
B x C x D	88.817	1	88.817	4.627	.034*
4 Way Interactions	8.817	1	8.817	.459	.499
A x B x C x D	8.817	1	8.817	.459	.499

status (Guilty/Innocent) X gender (Male/ Female) X type of control question (Nonexclusive/Exclusive), interaction [$F(1/104) = 4.627$, $p=.034$]. Examination of the mean scores for each group (see Table 13, p. 108) showed that males tested

---

Insert Table 13 about here

---

with nonexclusive control questions were the only ones to yield mean scores in both predicted directions (positive for innocent and negative for guilty). All innocent subjects tested with exclusive control questions produced mean scores in the wrong (guilty) direction. At the same time guilty females tested with nonexclusive controls produced a mean score in the innocent direction.

Table 14 (see p. 109) presents a summary of the ANOVA

Insert Table 14 about here

---

conducted on lower pneumograph (abdominal respiration) scores. This analysis shows a significant effect for status [$F(1/104) = 15.032$, $p<.000$], with mean scores for innocent and guilty of [M=1.43 & M=-1.97] respectively. There was one significant interaction for status (Guilty/Innocent) by type of control

Table 13

Upper Pneumograph Mean Scores for Status X Gender X Type of Control Question Interaction: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Gender	
Status	Male	Female
	Nonexclusive	
Innocent	2.20	1.60
Guilty	-4.00	0.80
	Exclusive	
Innocent	-1.45	-0.10
Guilty	-0.55	-1.10

p=.034

Table 14

ANOVA Summary For Lower Pnuemograph Total Numerical Scores:

Using All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	436.483	4	109.121	4.730	.002*
Type Test					
ZOC/MGQT (A)	45.633	1	45.633	1.978	.163
Status					
Inn/Guilty (B)	346.800	1	346.800	15.032	.000*
Gender					
Male/Female (C)	40.017	1	40.017	1.734	.191
Control					
Excl/Nonexc (D)	4.033	1	4.033	.175	.677
2 Way Interactions	229.100	6	38.183	1.655	.140
A x B	9.633	1	9.633	.418	.520
A x C	41.667	1	41.667	1.806	.182
A x D	4.800	1	4.800	.208	.649
B x C	18.150	1	18.150	.787	.377
B x D	124.033	1	124.033	5.376	.022*
C x D	30.817	1	30.817	1.336	.250
3 Way Interactions	128.883	4	32.221	1.397	.240
A x B x C	64.067	1	64.067	2.777	.099
A x B x D	4.800	1	4.800	.208	.649
A x C x D	9.600	1	9.600	.416	.520
B x C x D	50.417	1	50.417	2.185	.142
4 Way Interactions	9.600	1	9.600	.416	.520
A x B x C x D	9.600	1	9.600	.416	.520

question (Nonexclusive/Exclusive) [$F(1/104) = 5.376$, $p=.022$]. Each group of subjects produced mean scores that were in the predicted directions (positive for innocent and negative for guilty), however, subjects tested with nonexclusive control questions had more extreme scores [innocent, M=2.63 & guilty, M=-2.80] than those subjects tested with exclusive control questions, [innocent, M=0.23 & guilty, M=-1.13].

Table 15 (see p. 111) is a summary of the ANOVA conducted

Insert Table 15 about here

---

on the lower pneumograph (abdominal respiration) component scores calculated using just three relevant questions. That analysis shows two significant findings. There was a significant main effect for status (Guilty/Innocent) [$F(1/104) = 16.119$, $p<.000$] with a mean score for innocent subjects of [M=.80] and [M=-1.98] for guilty subjects. The other significant finding was a status (Guilty/Innocent) by type of control question (Nonexclusive/Exclusive) interaction [$F(1/104) = 7.644$, $p=.007$]; innocent subjects tested with exclusive controls produced a mean score that was in the wrong (guilty) direction. Whereas, all subjects tested using nonexclusive control questions [innocent, M=1.93 & guilty, M=-2.77] and guilty subjects tested with exclusive control questions [M=-1.20] produced mean scores in predicted

Table 15

ANOVA Summary For Lower Pnuemograph Total Numerical Scores:

Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	279.296	4	69.824	4.843	.001*
Type Test					
ZOC/MGQT (A)	10.208	1	10.208	.708	.402
Status					
Inn/Guilty (B)	232.408	1	232.408	16.119	.000*
Gender					
Male/Female (C)	33.004	1	33.004	2.289	.133
Control					
Excl/Nonexc (D)	3.675	1	3.675	.255	.615
2 Way Interactions	188.537	6	31.423	2.179	.051
A x B	.075	1	.075	.005	.943
A x C	34.504	1	34.504	2.393	.125
A x D	5.208	1	5.208	.361	.549
B x C	7.004	1	7.004	.486	.487
B x D	110.208	1	110.208	7.644	.007*
C x D	31.538	1	31.538	2.187	.142
3 Way Interactions	103.054	4	25.764	1.787	.137
A x B x C	40.837	1	40.837	2.832	.095
A x B x D	8.008	1	8.008	.555	.458
A x C x D	10.004	1	10.004	.694	.407
B x C x D	44.204	1	44.204	3.066	.083
4 Way Interactions	12.604	1	12.604	.874	.352
A x B x C x D	12.604	1	12.604	.874	.352

directions (positive for innocent and negative for guilty).

Table 16 (see p. 113) is a summary of the ANOVA conducted

Insert Table 16 about here

---

on the SRR (skin resistance response scores. That analysis shows there were two significant findings attributed to this measure. The first significant finding was a main effect for status (Guilty/Innocent) [F(1/ 104) = 33.987, p<.000]; innocent subjects produced a mean score of [M=2.75], while guilty subjects produced a mean score of [M=-5.18]. The other significant effect was a four way interaction, type of test (MGQT/ZOC) X status (Guilty/Innocent) X gender (Male/ Female) X type of control question (Nonexclusive/Exclusive) [F(1/104) = 4.801, p=.031]. The mean scores for each group are given in Table 17 (see p. 114) to assist in understanding the nature of

---

Insert Table 17 about here

---

of this interaction. All subjects tested with nonexclusive control questions produced mean scores in the predicted directions (positive for innocent and negative for guilty). The most extreme mean scores were produced by male subjects

Table 16

ANOVA Summary For SRR Total Numerical Scores:

Using All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	2031.671	4	507.918	9.143.	.000*
Type Test					
ZOC/MGQT (A)	80.033	1	80.033	1.441	.233
Status					
Inn/Guilty (B)	1888.133	1	1888.133	33.987	.000*
Gender					
Male/Female (C)	39.204	1	39.204	.706	.403
Control					
Excl/Nonexc (D)	24.300	1	24.300	.437	.510
2 Way Interactions	113.012	6	18.835	.339	.915
A x B	8.533	1	8.533	.154	.696
A x C	15.504	1	15.504	.279	.598
A x D	.833	1	.833	.015	.903
B x C	.504	1	.504	.009	.924
B x D	76.800	1	76.800	1.382	.242
C x D	10.837	1	10.837	.195	.660
3 Way Interactions	109.279	4	27.320	.492	.742
A x B x C	17.604	1	17.604	.317	.575
A x B x D	.133	1	.133	.002	.961
A x C x D	34.504	1	34.504	.621	.432
B x C x D	57.038	1	57.038	1.027	.313
4 Way Interactions	266.704	1	266.704	4.801	.031*
A x B x C x D	266.704	1	266.704	4.801	.031*

Table 17

SRR Total Mean Scores for Type of Test X Status X Gender X Type of Control Question Interaction: Using All Relevant Questions (5 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	1.10	6.20
Guilty	-4.30	-9.80
ZOC		
Innocent	6.40	2.80
Guilty	-5.50	-3.80
	Exclusive	
MGQT		
Innocent	3.10	-4.60
Guilty	-5.40	-5.20
ZOC		
Innocent	2.20	3.00
Guilty	-4.10	-4.80

p=.031

tested with the ZOC procedure and female subjects tested with the MGQT procedure both employing nonexclusive control questions. Only one group of subjects had a mean score that did not evolve in a predicted direction; innocent female subjects tested with the MGQT procedure using exclusive control questions produced a mean score that was in the wrong (guilty) direction.

Table 18 (see p. 116) shows a summary of the ANOVA

Insert Table 18 about here

---

conducted on the SRR (skin resistance response component scores calculated using three relevant questions. That analysis showed the same pattern of results as when all the relevant questions had been used to calculate component scores. A significant main effect was found for status (guilty/innocent) [$F(1/104) = 51.908$, $p<.000$]; innocent subjects produced a mean score of [M=2.60], while guilty subjects produced a mean score of [M=-4.47]. The other significant effect was once more a four way interaction of, type of test (MGQT/ ZOC) X status (Guilty/Innocent) X gender (Male/ Female) X type of control question (Nonexclusive/ Exclusive) [$F(1/104) = 4.472$, $p=.037$]. Table 19 (see p. 118) shows that when just three relevant questions were scored the mean scores of subjects tested with the MGQT procedure were

Table 18

ANOVA Summary For SRR Total Numerical Scores:

Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	1561.167	4	390.292	13.523	.000*
Type Test					
ZOC/MGQT (A)	34.133	1	34.133	1.183	.279
Status					
Inn/Guilty (B)	1498.133	1	1498.133	51.908	.000*
Gender					
Male/Female (C)	8.067	1	8.067	.279	.598
Control					
Excl/Nonexc (D)	20.833	1	20.833	.722	.397
2 Way Interactions	144.267	6	24.044	.833	.547
A x B	58.800	1	58.800	2.037	.156
A x C	.267	1	.267	.009	.924
A x D	1.633	1	1.633	.057	.812
B x C	.267	1	.267	.009	.924
B x D	80.033	1	80.033	2.773	.099
C x D	3.267	1	3.267	.113	.737
3 Way Interactions	55.367	4	13.842	.480	.751
A x B x C	29.400	1	29.400	1.019	.315
A x B x D	.033	1	.033	.001	.973
A x C x D	19.267	1	19.267	.668	.416
B x C x D	6.667	1	6.667	.231	.632
4 Way Interactions	129.067	1	129.067	4.472	.037*
A x B x C x D	129.067	1	129.067	4.472	.037*

Insert Table 19 about here

slightly smaller, but the pattern of results mirrored that seen in Table 15 (see p. 111). Innocent female subjects tested with exclusive control questions and the MGQT procedure had a mean score in the wrong (guilty) direction. The other groups all produced mean scores in the predicted directions (positive for innocent and negative for guilty). The most extreme mean scores were still those of male subjects tested with the ZOC procedure and female subjects tested with the MGQT procedure both using nonexclusive control questions.

Table 20 (see p. 119) shows a summary of results for the ANOVA conducted on cardio (cardiovascular activity) component scores using all relevant questions (five for MGQT and three for ZOC). Whereas, Table 21 (see p. 120) is a summary of

Insert Tables 20 and 21 about here

results for the ANOVA conducted on the same component scores excluding the two relevant questions used for MGQT examinations and not used for ZOC examinations. Each analysis revealed only one significant effect and that was for status (guilty/ innocent); these were using the number of relevant

Table 19

SRR Total Mean Scores for Type of Test X Status X Gender X Type of Control Question Interaction: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	1.00	5.40
Guilty	-3.70	-7.00
ZOC		
Innocent	6.40	2.80
Guilty	-5.50	-3.80
	Exclusive	
MGQT		
Innocent	1.40	-2.00
Guilty	-3.70	-4.00
ZOC		
Innocent	2.20	3.00
Guilty	-4.10	-4.80

p=.037

Table 20

ANOVA Summary For Cardio Total Numerical Scores:

Using All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	449.708	4	112.427	6.668	.000*
Type Test					
ZOC/MGQT (A)	33.075	1	33.075	1.962	.164
Status					
Inn/Guilty (B)	392.408	1	392.408	23.272	.000*
Gender					
Male/Female (C)	7.350	1	7.350	.436	.511
Control					
Excl/Nonexc (D)	16.875	1	16.875	1.001	.319
2 Way Interactions	63.175	6	63.175	.624	.710
A x B	9.075	1	9.075	.538	.465
A x C	7.350	1	7.350	.436	.511
A x D	.675	1	.675	.040	.842
B x C	8.067	1	8.067	.478	.491
B x D	37.408	1	37.408	2.219	.139
C x D	.600	1	.600	.036	.851
3 Way Interactions	20.492	4	5.123	.304	.875
A x B x C	.600	1	.600	.036	.851
A x B x D	12.675	1	12.675	.752	.388
A x C x D	2.400	1	2.400	.142	.707
B x C x D	4.817	1	4.817	.286	.594
4 Way Interactions	25.350	1	25.350	1.503	.223
A x B x C x D	25.350	1	25.350	1.503	.223

Table 21

ANOVA Summary For Cardio Total Numerical Scores:

Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	380.183	4	95.046	7.347	.000*
Type Test					
ZOC/MGQT (A)	48.133	1	48.133	3.721	.056
Status					
Inn/Guilty (B)	300.833	1	300.833	23.255	.000*
Gender					
Male/Female (C)	20.417	1	20.417	1.578	.212
Control					
Excl/Nonexc (D)	10.800	1	10.800	.835	.363
2 Way Interactions	86.133	6	14.356	1.110	.362
A x B	30.000	1	30.000	2.319	.131
A x C	.817	1	.817	.063	.802
A x D	2.700	1	2.700	.209	.649
B x C	8.067	1	8.067	.624	.432
B x D	43.200	1	43.200	3.339	.071
C x D	1.350	1	1.350	.104	.747
3 Way Interactions	28.983	4	7.246	.560	.692
A x B x C	.600	1	.600	.046	.830
A x B x D	9.633	1	9.633	.745	.390
A x C x D	3.750	1	3.750	.290	.591
B x C x D	15.000	1	15.000	1.160	.284
4 Way Interactions	11.267	1	11.267	.871	.353
A x B x C x D	11.267	1	11.267	.871	.353

questions [F(1/104) = 23.272, p<.000] and using just three relevant questions [F(1/104) = 23.255, p<.000]. When all relevant questions were used innocent subjects produced a mean score of [M=1.73], while guilty subjects produced a mean score of [M=-1.88]. The use of three relevant questions changed the cardio mean scores little; innocent subjects produced a mean score of [M=1.40], and guilty subjects produced a mean score of [M=-1.77]. No other results were found significant.

Subject Concern Ratings

After the polygraph tests subjects rated their level of concern for each relevant and control question asked using a scale from 1 to 4. The higher the score the greater the concern. A four way MANOVA, type of test (MGQT/ZOC) X status (Guilty/ Innocent) X gender (Male/Female) X type of control question (Nonexclusive/ Exclusive), was conducted treating subject's concern ratings on relevant and control questions as a repeated measure. Table 22 (see p. 122) is a summary of the MANOVA examining the effect(s) of the independent variables with the concern ratings. Analysis showed one significant Between-Subject effect, that was for status (Guilty/Innocent)

Insert Table 22 about here

Table 22

MANOVA Summary For Subject Concern Ratings (A Repeated Measure) For Relevant and Control Questions By Type Test, Status, Gender and Type Control question

Analysis Of Variance

Tests of Between-Subject Effects.

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
WITHIN CELLS	72.65	104	.70		
CONSTANT	943.04	1	943.04	1350.07	.000*
Type Test					
ZOC/MGQT (A)	.90	1	.90	1.29	.259
Status					
Inn/Guilty (B)	23.85	1	23.85	34.15	.000*
Gender					
Male/Female (C)	1.32	1	1.32	1.89	.172
Control					
Excl/Nonexc (D)	.88	1	.88	1.27	.263
A x B	1.75	1	1.75	2.51	.116
A x C	.74	1	.74	1.05	.307
A x D	.04	1	.04	.06	.802
B x C	1.26	1	1.26	1.80	.182
B x D	.03	1	.03	.04	.845
C x D	.01	1	.01	.01	.910
A x B x C	.57	1	.57	.82	.367
A x B x D	2.08	1	2.08	2.98	.087
A x C x D	1.72	1	1.72	2.46	.120
B x C x D	.04	1	.04	.05	.826
A x B x C x D	.05	1	.05	.15	.703

Table 22 (cont'd)

Tests involving 'CONCERN' Within-Subject Effect.

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
WITHIN CELLS	33.24	104	.32		
CONCERN (E)	.06	1	.06	.19	.660
A x E	.31	1	.31	.96	.330
B x E	31.52	1	31.52	98.61	.000*
C x E	.27	1	.27	.86	.357
D x E	.01	1	.01	.03	.868
A x B x E	.92	1	.92	2.87	.093
A x C x E	.21	1	.21	.67	.415
A x D x E	.05	1	.05	.15	.703
B x C x E	1.01	1	1.01	3.17	.078
B x D x E	.19	1	.19	.60	.440
C x D x E	.49	1	.49	1.55	.217
A x B x C x E	.58	1	.58	1.82	.180
A x B x D x E	.39	1	.39	1.21	.275
A x C x D x E	.11	1	.11	.36	.552
B x C x D x E	.32	1	.32	1.00	.319
A x B x C x D x E	.15	1	.15	.46	.499

[F(1/104) = 34.15, p<.000]. The only significant Within-Subject effect was the interaction between status (Guilty/Innocent) and concern for the type of question asked (Relevant/Control) [F(1/104) = 98.61, p<.000]. That is results showed innocent subjects had more concern for control questions and less concern for relevant questions, whereas guilty subjects had more concern for relevant questions and less concern for the controls. The interaction is clearly seen by examining the mean concern ratings for each group (Guilty/Innocent) as presented in Figures 6 and 7 (see pgs. 124 and 125). The results of the MANOVA are displayed by the

---

Insert Figures 6 and 7 about here

---

type of control question (Exclusive/Nonexclusive), for illustrative purposes only. The differences in concern rating by type of control question was not statistically significant.

Effect of Varying Numerical Cutoff Scores

Other studies (Barland & Raskin, 1975; Patrick & Iacono, 1989; Raskin & Hare, 1978) have found that varying the numerical cutoff score at which inconclusive decisions are made will yield different levels of accuracy. The effect on accuracy of manipulating the size of the inconclusive zone was

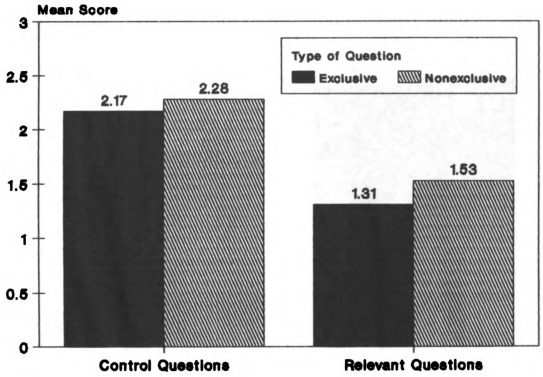


Figure 6

Innocent Subject's Mean Concern Ratings

For Relevant and Control Questions

By Type of Control Question

1.63
1.31

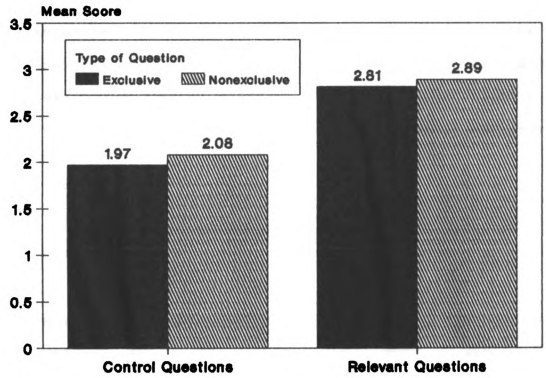


Figure 7

Guilty Subject's Mean Concern Ratings

For Relevant and Control Questions

By Type of Control Question

Mean Score
Type of Question
Relevant Questions
Control Questions

examined in this study by computing the number right, wrong and inconclusive using cutoff scores ranging from zero to ±12. Earlier analyses, however, showed that accuracy was significantly related to the type of control question used (Exclusive/Nonexclusive). Therefore, to examine this issue thoroughly results were computed separately for exclusive and nonexclusive control question groups. Figure 8 (see p. 127) shows the percent of accurate and inconclusive decisions at

---

Insert Figure 8 about here

---

each cutoff score using exclusive and nonexclusive control questions to classify guilty subjects as deceptive. The percentage of correct classifications differed by the type of control question used for testing. Using a cutoff score of zero would result in 83% of the guilty subjects tested with nonexclusive control questions and 67% of the guilty subjects tested with exclusive control questions classified correctly. There were no inconclusive decisions attributed to examinations using nonexclusive control questions with a cutoff score of zero. However, examinations using exclusive control questions had an inconclusive rate of 3% (1/30).

Figure 9 (see p. 129) shows the percent of accurate and inconclusive decisions at each cutoff score using exclusive and nonexclusive control questions to classify innocent

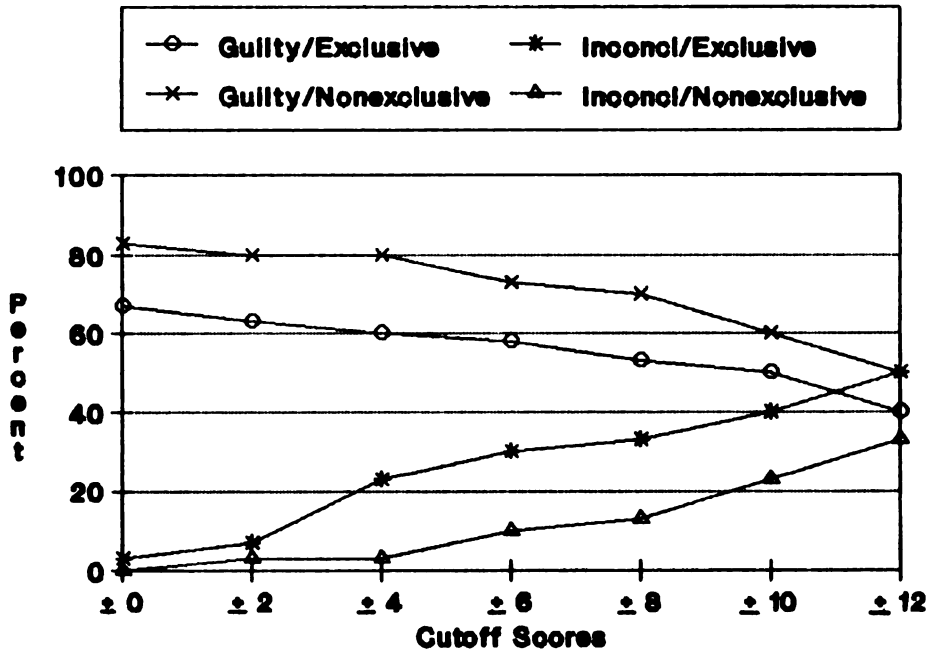


Figure 8

Accuracy of Guilty Decisions & Percent Inconclusive By Type of Control Question For Different Inconclusive Ranges

Insert Figure 9 about here

subjects as truthful. There were no inconclusive decisions when the cutoff score was zero; 83% of the innocent subjects tested with nonexclusive control questions and 47% of those tested with exclusive control questions were classified correctly.

The results of past studies (Barland & Raskin, 1975; Patrick & Iacono, 1989; Raskin & Hare, 1978) shows that accuracy varied from one cutoff score to another. Using all of the relevant questions, 5 for MGQT tests and 3 for ZOC tests, the highest rate of accuracy obtained was using a cutoff score of zero and nonexclusive control questions; 83% of both the innocent and guilty subjects were identified correctly with no inconclusive decisions. These results are similar to those reported by Raskin and Hare (1978) who found a cutoff score of ± 2 to be the most productive for classifying subjects as innocent or guilty, achieving an accuracy rate exceeding 95%. These results are challenged by Patrick and Iacono (1989). They found the accuracy of classifying innocent subjects was at best 59%, this using a cutoff score of $\pm$ 2. The reason for this finding they said was that "the scores for guilty subjects were strongly clustered toward the negative end of the continuum, [whereas] the scores for innocent subjects were more evenly distributed over a range of

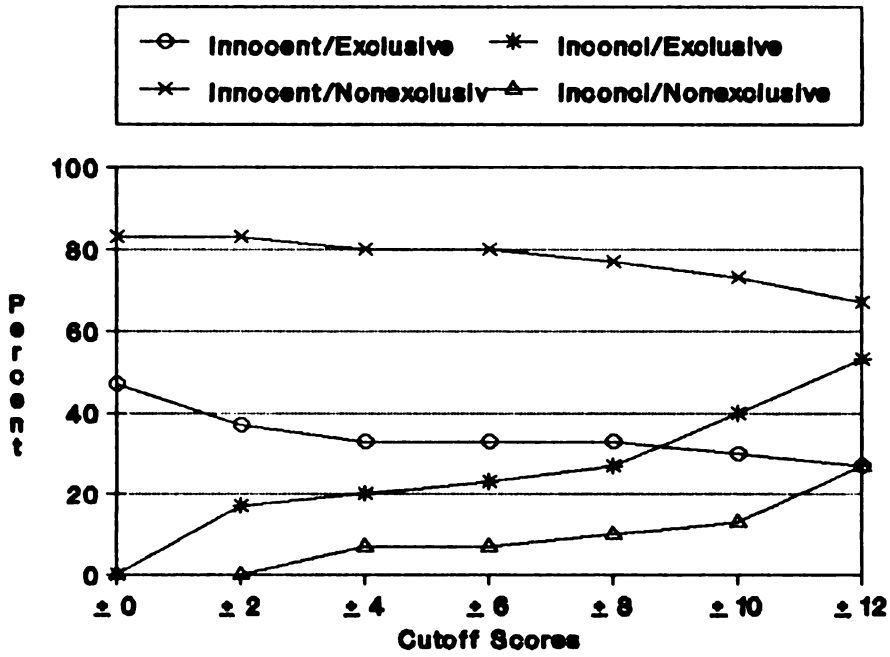


Figure 9

Accuracy of Innocent Decisions & Percent Inconclusive By Type of Control Question For Different Inconclusive Ranges

-26 to +22 (p. 351)." In the present study Figure 10 (see p. 131) shows that no matter which type of control question was

---

Insert Figure 10 about here

---

used for testing guilty subjects, the numerical scores cluster toward the negative end of the continuum. On the other hand Figure 11 (see p. 132) shows that only the numerical scores of innocent subjects tested with nonexclusive control questions

---

Insert Figure 11 about here

---

cluster toward the predicted (positive) end of the continuum, whereas, the scores of innocent subjects tested with exclusive control questions are, as seen in Patrick and Iacono's (1989) study, more evenly distributed across the continuum. Patrick and Iacono (1989) used exclusive control questions for the polygraph examinations given in their study.

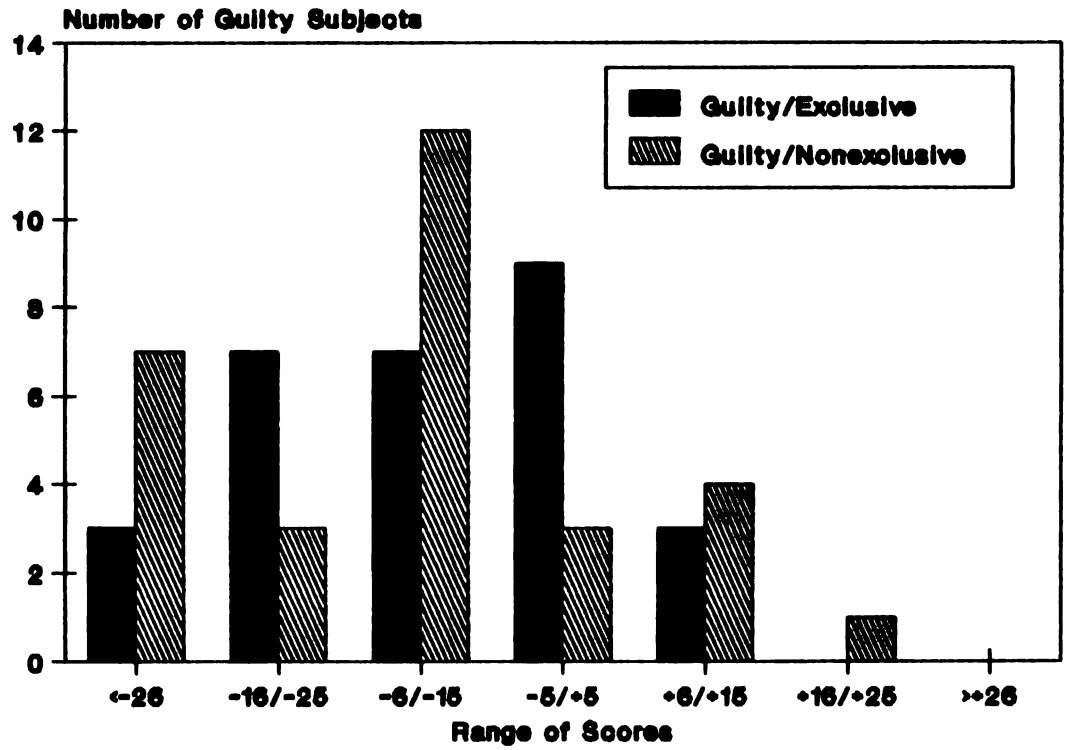
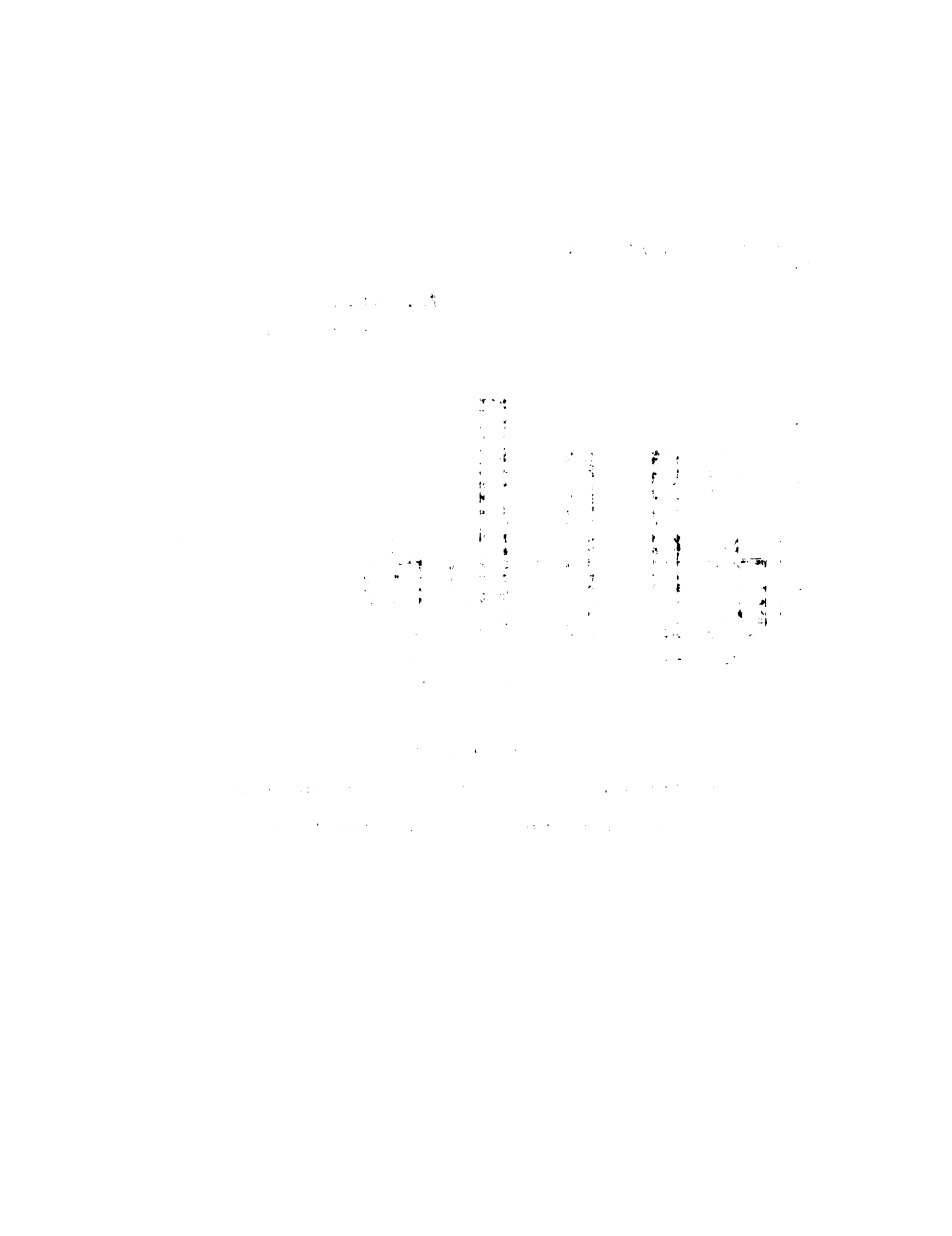


Figure 10

Distribution of Guilty Subject's Numerical Scores By Type of Control Question



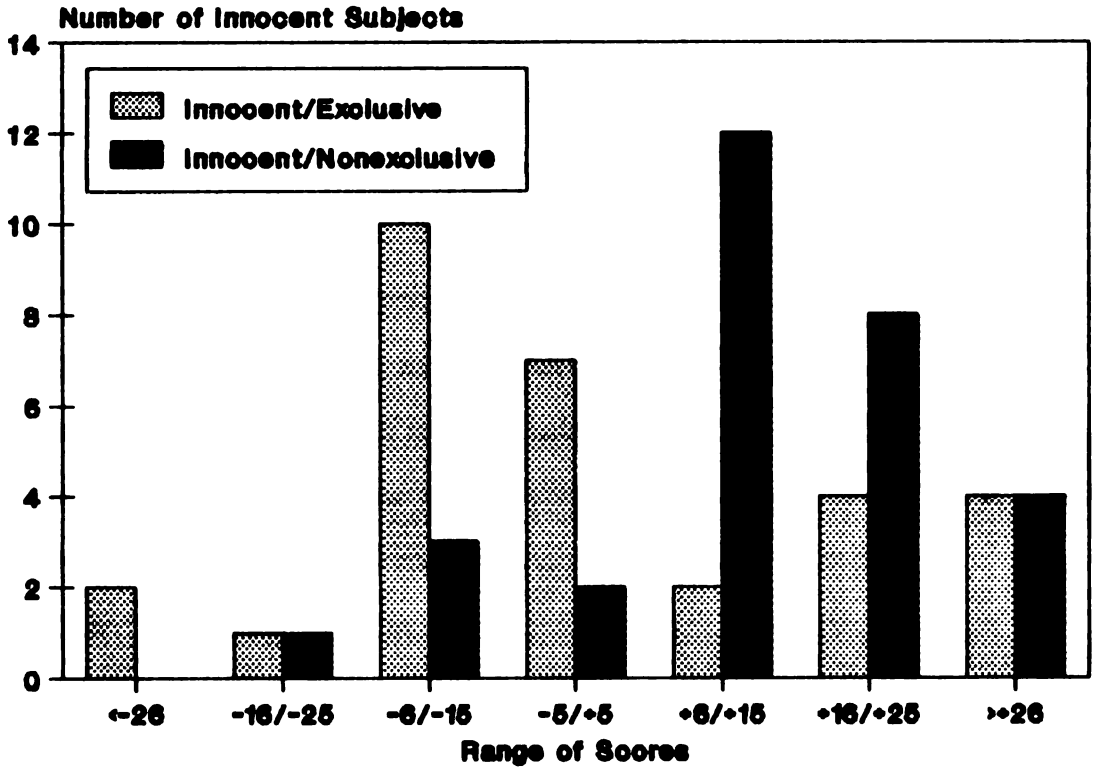


Figure 11

Distribution of Innocent Subject's Numerical Scores By Type of Control Question

CHAPTER V

Discussion

These findings generally support the theory posited by advocates of Control Question polygraph testing (Reid & Inbau, 1977), that in a specific issue context innocent (truthful) and guilty (deceptive) subjects will respond differentially to relevant and control questions. Overall, numerical scoring of the polygraph charts showed that innocent subjects were more responsive to control questions than to relevant questions, while guilty subjects were more responsive to relevant questions than they were to control questions. Correspondingly, the amount of concern for relevant and control questions experienced by subjects also varied according to status; the innocent subjects reported more concern for control questions, whereas guilty subjects reported more concern for relevant questions. These results strongly agree with earlier studies exploring this topic (Bradley and Janisse, 1981; Horvath, 1988).

The most interesting, and unexpected, results were those concerning the relative effectiveness of exclusive and nonexclusive control questions. The idea of using a time-bar to exclude the period in which a crime was committed, on its face seems valid. Additionally, the results of past research

(Podlesny & Raskin, 1978) have lent credibility to this practice. One of the conclusions reached by Podlesny and Raskin (1978) was that exclusive control questions were preferable to nonexclusive control questions. Their results showed that exclusive control questions provided significant identification of both innocent and guilty subjects, while nonexclusive control questions gave significant identification of only the innocent. However, the results of Horvath's (1988) study were in opposition to those of Podlesny and Raskin (1978).

Unlike the two earlier studies (Podlesny & Raskin, 1978; Horvath, 1988), the findings of this study show that differences in the accuracy of identifying both the innocent and the guilty is significantly related to the type of control question used.

The results of this study strongly disagrees with those who advocate the use of exclusive control questions in the form used for this project (Backster, 1965; Podlesny & Raskin, 1978). Further, the assumption that nonexclusive control questions may be mistaken as relevant questions by the guilty, thus leading to more false negative errors, was shown to be false. Results show that tests using exclusive control questions produced about the same proportion of false negative errors (guilty called innocent) as did tests using nonexclusive control questions. The proportions of false positive error (innocent called guilty) and inconclusive decisions made using exclusive control questions for

testing was three to four times larger than the proportion of error realized using nonexclusive control questions for testing. These findings suggest that the perceptions of innocent subjects regarding the comparable importance of exclusive control questions and relevant questions, vis-a-vis nonexclusive control questions and relevant questions, were significantly different (see Table 5, p. 81; Figures 1 and 2, pgs. 83 & 86).

Examination of the mean total numerical scores also demonstrates that subjects responded in a differential manner to the exclusive and nonexclusive control questions. The mean total numerical scores for subjects tested using nonexclusive control questions were more extreme, in both a positive and negative direction, compared to the mean total numerical scores of subjects tested using exclusive control questions (see Figures 4 and 5, pgs. 95 and 99). These findings show quite convincingly, that at least in this context, exclusive control questions do not perform as advocates would predict.

The reason(s) neither of the earlier studies (Podlesny & Raskin, 1978; Horvath, 1988) discovered a significant difference in the overall accuracy using one type of control question or the other may be attributed to variations in methodology. For example, the size or characteristics of the sample used, the testing procedures employed or the mechanics involved in their administration could have interacted together or in part affecting results and consequently the conclusions proffered.

The sensitivity of an experiment to detect treatment differences is related directly to the sample size used (Cohen, 1977; Hunter, Schmidt & Jackson, 1977; Hunter & Schmidt, 1990). "The greater the sample size, the greater the power and the more sensitive the experiment in detecting treatment differences in the population (Keppel, 1982, p. 70)." The sample size used in this study was three times larger (120/40) than that used in the earlier studies (Podlesny & Raskin, 1978; Horvath, 1988). Therefore, this experiment was more likely to detect significant differences in accuracy attributable to one or all of the treatments.

The efficacy of using samples drawn from college populations is another issue that has received scholarly attention (Honts, Hodes & Raskin, 1985; Horvath, 1988). The population sampled for this study was first and second year college students. It was posited by Honts, Hodes and Raskin (1985) that the use of such samples may be the reason other studies (Barland & Raskin, 1975; Bradley & Janisse, 1981; Szucko & Kleinmuntz, 1981) have reported lower levels of accuracy. Honts et. al. (1985) reason that college students may not take their participation in an experiment seriously. Consequently, the amount of concern induced is negligible or at least different than that felt by a participant from the general public or someone given a real examination (Podlesny & Raskin, 1977), contributing in some unique way to the number or type of errors found in mock crime experiments. If such a hypothesis were to be tenable it would mean that in a

laboratory setting college students assigned to a guilty condition would have less concern for relevant questions, since they "perceive the mock crime as more of a game and may thus be less stimulated by the situation in general Honts et. al., 1985, p. 186)." Nonetheless, one should not expect a frivolous attitude to decrease greatly the arousal evoked when a control question is asked; it may in fact increase the amount of concern felt by pseudo guilty participants in response to a control question because of any past indiscretion that would seem inappropriate to disclose during the course of some "game." Accordingly, in such a context an equal or larger number of false negative errors (guilty subjects called innocent), compared to the number of false positive errors (innocent subjects called guilty), would be expected. The findings of this and Horvath's (1988) study, both using college students and nonexclusive control questions, strongly contradict this hypothesis.

To date studies examining the detection of deception have generally used all male samples (Raskin & Hare, 1978; Rovner, Raskin & Kircher, 1979; Podlesny & Raskin, 1978; Horvath, 1988). Although, some studies have included both male and female (Dawson, 1980; Widacki & Horvath, 1978) participants, they typically have not looked for possible gender effects. The findings of one polygraph study that examined in part the effect(s) of gender (Honts et al., 1985), showed that significant effects were attributable to differences between male and female subjects. The results of this study, however,

show there were no main effects attributed to differences in gender. Still, when data for each physiological parameter was analyzed separately, the results for two of the four parameters indicated that higher order interactions involving gender were evoked (see Tables 8 to 17). It is not known why these effects weren't seen in the other two parameters, hence in the results overall. Studies examining the differences in gender and the control question polygraph examination are relatively new; the first study to do so, using a mock-crime setting, was Honts et. al. (1985). Consequently, many more studies are needed before we can begin to understand the nature of these interactions relative to the differences in gender.

The first laboratory study to examine the merits of control question testing, using a mock crime scenario and field polygraph techniques, was conducted by Barland and Raskin (1975). The variation of control question test used was the Zone Comparison test, because:

> "the federal government considers the Backster zone comparison technique to be the technique of preference for criminal cases where only one major aspect of the crime is probed with a S[ubject]" (Barland & Raskin, 1975, p. 323).

Accordingly, researchers to date (see Table 1, p. 18) have used for all but two mock crime laboratory studies (Horvath, 1988; Widacki & Horvath, 1978), some variation of the Zone Comparison test. Even though the ZOC test incorporates many modifications said to improve upon the perceived shortcomings of the MGQT, the findings of this study do not support the use

of one procedure over the other. In fact the accuracy of the Zone Comparison test was statistically equivalent to that of the Modified General Question Test. Consequently, this result bolsters Horvath's (1980, 1988) proposal that:

> "CQ testing is sufficiently robust that many of the minor differences in the manner in which that testing is structured have little effect" (p. 208).

This hypothesis is further sustained by the results obtained by manipulating the number of relevant questions used for numerical scoring. A ZOC test usually employs three relevant questions for scoring (Barland & Raskin, 1975; Podlesny & Raskin; 1978), whereas a MGQT generally uses five (Horvath, 1988; Reid & Inbau, 1977). The findings of this study show that examinations using five relevant questions produced slightly larger mean total numerical scores compared to those calculated using just three relevant questions (cf. Figures 4 and 5, pgs. 93 & 97). Still, the differences in accuracy did not exceed the level of chance.

An alternative hypothesis regarding the differences in results, between this and the earlier studies, concerns the procedures used to administer the polygraph examinations for each study. The accuracy of classifying subjects as innocent or guilty can easily be influenced by the setting in which an examination is conducted (Reid & Inbau, 1977; Horvath, 1988; Podlesny & Raskin, 1977). Therefore, field examinations are usually administered by one examiner in a quiet room alone with the examinee, such a setting is described by Reid and Inbau (1977) who state:

> Polygraph tests should be conducted in a quiet, private, semi-soundproof room. Extraneous noises, such as. . .the conversation of persons outside the examination room, or the presence of. . .other spectators in the room itself would induce disturbances and distractions that in turn would distort the various physiological recordings and seriously interfere with a satisfactory polygraph diagnosis (p. 6).

The polygraph tests administered in this and Horvath's (1988) study were conducted in settings equivalent to those described by Reid and Inbau (1977). The earlier study conducted by Podlesny and Raskin (1978) however, used three examiners operating together to administer each examination, while each examinee was placed in an isolation booth for remote testing by means of an intercom. The consequences of using these procedures are unknown and certainly warrant additional study.

Conclusion

In summary, these results illustrate the effectiveness of Control Question testing in discriminating between people assigned to innocent and guilty conditions. The results also contribute to a growing body of knowledge devoted to the objective evaluation of these procedures.

The most meaningful finding was that tests using nonexclusive control questions were significantly more accurate for classifying the innocent than tests using exclusive control questions. Tests using exclusive control questions produced a significantly larger proportion of false positive errors compared to tests using nonexclusive control

questions. This finding is opposite that of other earlier research (Podlesny & Raskin, 1978).

Another important finding was that the accuracy of the ZOC and MGQT procedures was comparable when used in the same context. Although many critics (Furedy, 1986, 1988; Kleinmuntz & Szucko, 1982; Lykken, 1974, 1979, 1981) oppose control question testing, the results of this endeavor and other empirical studies (see Tables 1 through 3, pgs. 18, 20 & 54) demonstrate consistent findings supporting the use of CQ testing. This experiment shows that CQ testing is sufficiently robust (Horvath, 1980, 1988) that small differences in structure will have little effect on accuracy. It also shows that the type of control question used for CQ testing is not a small difference. Consequently, the large proportions of false positive error seen in many studies, and often noted by critics (Lykken, 1974, 1979, 1981), may be the result of using exclusive control questions rather than a fundamental flaw in CQ testing.

Finally, as with most laboratory based studies is the issue regarding the appropriateness of generalizing these findings to the field. This issue has been debated in the literature (e.g., Horvath, 1984; Lykken, 1979, 1981; Patrick & Iacono, 1989). However, there is no one correct answer. It suffices to say that the results of any study must be examined closely. Then, as consistent findings evolve they should be moved into the field for further study.

APPENDICES

Appendix A

INFORMED CONSENT FORM

"Lie Detection Experiment"

I, _____ voluntarily agree to participate in a detection of deception ("lie detection") experiment to be carried out by both John Palmatier, Graduate Assistant, and Dr. Frank Horvath, Professor, School of Criminal Justice, Michigan State University. I have been told and understand the nature of this experiment and that my participation is completely voluntary and will require approximately two hours of my time. I know that I am free to discontinue my involvement in the experiment at any time without any penalty or recrimination except for the loss of extra credit and the "reward" as described in my instructions.

I understand that my identity and all information related to my participation will be held in strict confidence by the researchers regardless of whether or not I choose to complete the experiment as instructed. I also understand that if I choose to complete the experiment as instructed I will be given extra credit toward my course grade in _____ consistent with the procedure outlined in that class. I understand that if I have any questions during the experiment I may ask either of the researchers for clarification. I have been told that a copy of the results will be available for my perusal upon completion of the experiment through Dr. Horvath's office.

I further understand that in the unlikely event of physical injury resulting from research procedures, Michigan State University, its agents, and employees will assume that responsibility as required by law. Emergency medical treatment for injuries or illness is available where the injury or illness is incurred in the course of an experiment. I have been advised that I should look toward my own health insurance program for payment of said medical expenses.

NAME:_____**DATE:**_____

COURSE:_____

INSTRUCTOR:_____

Appendix B

INTERVIEW & QUESTION FORM
MGQT (Nonexclusive)

EXAMINEE:_____
DATE:_____
ADDRESS:_____
TX:_____
AGE:_____RACE:_____
MAJOR:_____
YRS IN SCHOOL:_____MARITAL
STATUS:_____
PHYSICAL/MEDICAL/MEDICATION:_____
_____

QUESTIONS

1. Do they call you _____?

2. Are you over _____ years of age?

3. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?

4. Do you live in the United States?

5. Did you take that envelope containing three dollars?

6. Did you ever take something that did not belong to you?

7. Did you ever go to school?

8. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?

9. Did you write your name on that airmail envelope taken from Dr. Horvath's mail slot?

10. Did you ever tell a lie about something important?

11. Were you assigned to be a guilty person in this research?

TEST #1 1 2 3 4 5 6 7 8 9 10 11

TEST #2 1 2 3 4 5 6 7 8 9 10 11

TEST #3 7 4 11 8 10 1 3 6 2 5 10 9

Appendix C

INTERVIEW & QUESTION FORM
MGQT (Exclusive)

EXAMINEE:_____
DATE:_____
ADDRESS:_____
TX:_____
AGE:_____RACE:_____
MAJOR:_____
YRS IN SCHOOL:_____MARITAL
STATUS:_____
PHYSICAL/MEDICAL/MEDICATION:_____
_____

QUESTIONS

1. Do they call you _____?

2. Are you over _____ years of age?

3. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?

4. Do you live in the United States?

5. Did you take that envelope containing three dollars?

6. Before the age of_____did you ever take something that did not belong to you?

7. Did you ever go to school?

8. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?

9. Did you write your name on that airmail envelope taken from Dr. Horvath's mail slot?

10. Before the age of_____ did you ever tell a lie about something important?

11. Were you assigned to be a guilty person in this research?

TEST #1 1 2 3 4 5 6 7 8 9 10 11

TEST #2 1 2 3 4 5 6 7 8 9 10 11

TEST #3 7 4 11 8 10 1 3 6 2 5 10 9

Appendix D

INTERVIEW & QUESTION FORM
ZOC (Exclusive)

EXAMINEE:_____
DATE:_____
ADDRESS:_____
TX:_____
AGE:_____RACE:_____
MAJOR:_____
YRS IN SCHOOL:_____MARITAL
STATUS:_____
PHYSICAL/MEDICAL/MEDICATION:_____
_____

QUESTIONS

1. Is your name_____?

2. Are you afraid I'll ask you a question we have not reviewed?

3. Do you intend to answer truthfully each question about the stolen envelope?

4. Before the age of_____did you ever take something that did not belong to you?

5. Did you take that envelope containing three dollars?

6. Before the age of_____did you ever tell a lie about something important?

7. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?

8. Are you now in Michigan?

9. Before the age of_____did you ever tell a lie to a person in authority?

10. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?

TEST #1 1 2 3 4 5 6 7 8 9 10

TEST #2 8 2 3 9 5 4 7 1 6 10

TEST #3 1 2 3 4 10 6 5 8 9 7

Appendix E

INTERVIEW & QUESTION FORM
ZOC (Nonexclusive)

EXAMINEE:_____
DATE:_____
ADDRESS:_____
TX:_____
AGE:_____RACE:_____
MAJOR:_____
YRS IN SCHOOL:_____MARITAL
STATUS:_____
PHYSICAL/MEDICAL/MEDICATION:_____
_____

QUESTIONS

1. Is your name_____?

2. Are you afraid I'll ask you a question we have not reviewed?

3. Do you intend to answer truthfully each question about the stolen envelope?

4. Did you ever take something that did not belong to you?

5. Did you take that envelope containing three dollars?

6. Did you ever tell a lie about something important?

7. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?

8. Are you now in Michigan?

9. Did you ever tell a lie to a person in authority?

10. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?

TEST #1 1 2 3 4 5 6 7 8 9 10

TEST #2 8 2 3 9 5 4 7 1 6 10

TEST #3 1 2 3 4 10 6 5 8 9 7

Appendix F

POST TEST QUESTIONAIRE: ZOC
(Nonexclusive)

NAME:_____

To complete your participation in this research we ask you to fill out this questionaire as honestly as you can. Your answers will in no way affect your extra credit.

Listed below are some of the questions asked during the polygraph examination. For each, please indicate in the appropriate space the amount of concern you felt when asked that particular question during the exam... that is, to what degree do you believe you responded "physiologically" when asked this question. Please answer as honestly as possible and with no consideration as to whether you were "innocent" or "guilty".

	Degree of Concern			
	None	Slight	Moderate	A lot
1. Did you ever take something that did not belong to you?	1	2	3	4
2. Did you take that envelope containing three dollars?	1	2	3	4
3. Did you ever tell a lie about some thing important?	1	2	3	4
4. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
5. Did you ever tell a lie to a person in authority?	1	2	3	4
6. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?	1	2	3	4

7. Did you do or attempt to do anything 1 (YES) 2 (NO) during the polygraph testing to try to "beat the test"?

If you answered "yes", please explain briefly what you did (e.g. tried to control thoughts, control emotions, control physiological responces, and so forth) use the back of this form for your comments.

Appendix G

POST TEST QUESTIONAIRE: ZOC
(Exclusive)

NAME:_____

To complete your participation in this research we ask you to fill out this questionaire as honestly as you can. Your answers will in no way affect your extra credit.

Listed below are some of the questions asked during the polygraph examination. For each, please indicate in the appropriate space the amount of concern you felt when asked that particular question during the exam... that is, to what degree do you believe you responded "physiologically" when asked this question. Please answer as honestly as possible and with no consideration as to whether you were "innocent" or "guilty".

	Degree of Concern			
	None	Slight	Moderate	A lot
1. Before the age of_____Did you ever take something that did not belong to you?	1	2	3	4
2. Did you take that envelope containing three dollars?	1	2	3	4
3. Before the age of_____Did you ever tell a lie about something important?	1	2	3	4
4. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
5. Before the age of_____Did you ever tell a lie to a person in authority?	1	2	3	4
6. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?	1	2	3	4

7. Did you do or attempt to do anything 1 (YES) 2 (NO) during the polygraph testing to try to "beat the test"?

If you answered "yes", please explain briefly what you did (e.g. tried to control thoughts, control emotions, control physiological responces, and so forth) use the back of this form for your comments.

Appendix H

POST TEST QUESTIONAIRE: MGQT
(Exclusive)

NAME:_____

To complete your participation in this research we ask you to fill out this questionaire as honestly as you can. Your answers will in no way affect your extra credit.

Listed below are some of the questions asked during the polygraph examination. For each, please indicate in the appropriate space the amount of concern you felt when asked that particular question during the exam... that is, to what degree do you believe you responded "physiologically" when asked this question. Please answer as honestly as possible and with no consideration as to whether you were "innocent" or "guilty".

	Degree of Concern			
	None	Slight	Moderate	A lot
1. Did you take that airmail envelope out of Dr. Horvath's mail box Baker Hall?	1	2	3	4
2. Did you take that envelope containing three dollars?	1	2	3	4
3. Before the age of_____Did you ever take something that did not belong to you?	1	2	3	4
4. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
5. Did you write your name on that air airmail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
6. Before the age of_____did you ever tell a lie about something important?	1	2	3	4
7. Were you assigned to be a guilty person in this research?	1	2	3	4

8. Did you do or attempt to do anything 1 (YES) 2 (NO) during the polygraph testing to try to "beat the test"?

If you answered "yes", please explain briefly what you did (e.g. tried to control thoughts, control emotions, control physiological responces, and so forth) use the back of this form for your comments.

Appendix I

POST TEST QUESTIONAIRE: MGQT
(Nonexclusive)

NAME:_____

To complete your participation in this research we ask you to fill out this questionaire as honestly as you can. Your answers will in no way affect your extra credit.

Listed below are some of the questions asked during the polygraph examination. For each, please indicate in the appropriate space the amount of concern you felt when asked that particular question during the exam... that is, to what degree do you believe you responded "physiologically" when asked this question. Please answer as honestly as possible and with no consideration as to whether you were "innocent" or "guilty".

	Degree of Concern			
	None	Slight	Moderate	A lot
1. Did you take that airmail envelope out of Dr. Horvath's mail box in Baker Hall?	1	2	3	4
2. Did you take that envelope containing three dollars?	1	2	3	4
3. Did you ever take something that did not belong to you?	1	2	3	4
4. Did you remove three dollars from an airmail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
5. Did you write your name on that air-mail envelope taken from Dr. Horvath's mail slot?	1	2	3	4
6. Did you ever tell a lie about some-thing important?	1	2	3	4
7. Were you assigned to be a guilty person in this research?	1	2	3	4

8. Did you do or attempt to do anything 1 (YES) 2 (NO) during the polygraph testing to try to "beat the test"?

If you answered "yes", please explain briefly what you did (e.g. tried to control thoughts, control emotions, control physiological responces, and so forth) use the back of this form for your comments.

NUMERICAL EVALUATION SHEET (ZOC)	POLYGRAPH EXAM:
PERSON EXAMINED:	COMPLAINT NO:
EXAMINER:	DATE:

CHART #1	5	7	10	COMPONENT TOTAL
PNEUMO				
PNEUMO				CHART NUMBER ONE
G.S.R.				
CARDIO				
				TOTAL
(COMPARE 4/5 6/7 9/10)				

CHART #2	5	7	10	COMPONENT TOTAL
PNEUMO				
PNEUMO				CHART NUMBER TWO
G.S.R.				
CARDIO				
				TOTAL
(COMPARE 9/5 4/7 6/10)				

CHART #3	5	7	10	COMPONENT TOTAL
PNEUMO				
PNEUMO				CHART NUMBER THREE
G.S.R.				
CARDIO				
				TOTAL
(COMPARE 4/10 6/5 9/7)				

SPOT TOTALS				EXAMINATION TOTAL	

OPINION: ☐ TRUTHFUL ☐ DECEPTIVE ☐ INCONCLUSIVE

Appendix K

NUMERICAL EVALUATION SHEET (MSQT)	POLYGRAPH EXAM:
PERSON EXAMINED:	COMPLAINT NO:
EXAMINER:	DATE:

CHART #1	3	5	8	9	11	COMPONENT TOTAL
PNEUMO						
PNEUMO						
G.S.R.						
CARDIO						
	(COMPARE 3/6	5/6	8/6	9/10	11/10)	

CHART NUMBER ONE

TOTAL

CHART #2	3	5	8	9	11	COMPONENT TOTAL
PNEUMO						
PNEUMO						
G.S.R.						
CARDIO						
	(COMPARE 3/6	5/6	8/6	9/10	11/10)	

CHART NUMBER TWO

TOTAL

CHART #3	3	5	8	9	11	COMPONENT TOTAL
PNEUMO						
PNEUMO						
G.S.R.						
CARDIO						
	(COMPARE 3/6	5/10	8/10	9/10	11/10)	

CHART NUMBER THREE

TOTAL

SPOT TOTALS						EXAMINATION TOTAL	

OPINION: ☐ TRUTHFUL ☐ DECEPTIVE ☐ INCONCLUSIVE

Appendix L

Table 23

Distribution of Examiner's Decisions Based On All Relevant Questions

	Decisions				
	No.(%) Correct	No.(%) Wrong	No.(%) Inconclusive	No.(%) False Positive	No.(%) False Negative
Control Question					
Exclusive	33 (55%)	10 (16%)	17 (29%)	8 (13%)	2 (03%)
Nonexclusive	37 (62%)	12 (20%)	11 (18%)	9 (15%)	3 (05%)
Variation of Test					
ZOC	40 (67%)	7 (12%)	13 (21%)	4 (07%)	3 (05%)
MGQT	30 (50%)	15 (25%)	15 (25%)	13 (22%)	2 (03%)
Gender					
Female	25 (63%)	6 (15%)	9 (22%)	6 (15%)	0 (00%)
Male	45 (56%)	16 (20%)	19 (24%)	11 (14%)	5 (06%)

Appendix M

Table 24

Distribution of Examiner's Decisions Based on only Three Relevant Questions

	Decisions				
	No.(%) Correct	No.(%) Wrong	No.(%) Inconclusive	No.(%) False Positive	No.(%) False Negative
Control Question					
Exclusive	31 (52%)	9 (15%)	20 (33%)	7 (12%)	2 (03%)
Nonexclusive	34 (57%)	7 (11%)	19 (32%)	5 (08%)	2 (03%)
Variation of Test					
ZOC	40 (67%)	7 (12%)	13 (21%)	4 (07%)	3 (05%)
MGQT	25 (42%)	9 (15%)	26 (43%)	8 (13%)	1 (02%)
Gender					
Female	22 (55%)	4 (10%)	14 (35%)	4 (10%)	0 (00%)
Male	43 (54%)	12 (15%)	25 (31%)	8 (10%)	4 (05%)

Table 25

ANOVA Summary For Examiner's Examination Total Numerical Scores:

For All Relevant Questions (5 in MGQT and 3 in ZOC)

Analysis Of Variance

Source of Variation	Sum of Squares	DF	Mean Square	F	Signific of F
Main Effects	5706.292	4	1426.573	11.029	.000*
Type Test					
ZOC/MGQT (A)	1038.408	1	1038.408	8.028	.006*
Status					
Inn/Guilty (B)	4526.408	1	4526.408	34.996	.000*
Gender					
Male/Female (C)	141.067	1	141.067	1.091	.299
Control					
Excl/Nonexc (D)	.408	1	.408	.003	.955
2 Way Interactions	716.442	6	119.407	.923	.482
A x B	161.008	1	161.008	1.245	.267
A x C	248.067	1	248.067	1.918	.169
A x D	5.208	1	5.208	.040	.841
B x C	6.667	1	6.667	.052	.821
B x D	138.675	1	138.675	1.072	.303
C x D	156.817	1	156.817	1.212	.273
3 Way Interactions	81.442	4	20.360	.157	.959
A x B x C	32.267	1	32.267	.249	.619
A x B x D	1.008	1	1.008	.008	.930
A x C x D	4.817	1	4.817	.037	.847
B x C x D	43.350	1	43.350	.335	.564
4 Way Interactions	476.017	1	476.017	3.860	.058
A x B x C x D	476.017	1	476.017	3.860	.058

Table 26

Upper Pneumograph Mean Total Numerical Scores: For All Relevant Questions (5 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	3.40	1.20
Guilty	-4.80	4.60
ZOC		
Innocent	2.70	1.60
Guilty	-4.10	-1.40
	Exclusive	
MGQT		
Innocent	-0.20	0.40
Guilty	-2.00	2.00
ZOC		
Innocent	-2.00	0.00
Guilty	0.40	-3.40

Table 27

Upper Pneumograph Mean Total Numerical Scores: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	1.70	1.60
Guilty	-3.90	3.00
ZOC		
Innocent	2.70	1.60
Guilty	-4.10	-1.40
	Exclusive	
MGQT		
Innocent	-0.90	-2.00
Guilty	-1.50	-4.00
ZOC		
Innocent	-2.00	-0.20
Guilty	0.40	-3.40

Table 28

Lower Pneumograph Mean Total Numerical Scores: For All Relevant Questions (5 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	2.80	3.80
Guilty	-4.90	2.40
ZOC		
Innocent	2.40	1.60
Guilty	-3.70	-2.00
	Exclusive	
MGQT		
Innocent	1.80	1.00
Guilty	-1.60	0.80
ZOC		
Innocent	-1.80	0.40
Guilty	-0.40	-3.60

Table 29

Lower Pneumograph Mean Total Numerical Scores: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	1.20	2.80
Guilty	-4.50	1.80
ZOC		
Innocent	2.40	1.60
Guilty	-3.70	-2.00
	Exclusive	
MGQT		
Innocent	0.60	0.00
Guilty	-1.50	0.20
ZOC		
Innocent	-1.80	0.40
Guilty	-0.40	-3.60

Appendix S

Table 30

Cardio Mean Total Numerical Scores: For All Relevant Questions (5 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	1.30	1.80
Guilty	-1.60	-3.00
ZOC		
Innocent	3.70	4.20
Guilty	-2.70	-0.80
	Exclusive	
MGQT		
Innocent	0.30	0.60
Guilty	-2.40	-1.80
ZOC		
Innocent	0.20	3.20
Guilty	-0.80	-2.00

Table 31

Cardio Mean Total Numerical Scores: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	0.60	1.00
Guilty	-2.00	-2.20
ZOC		
Innocent	3.70	4.20
Guilty	-2.70	-0.80
	Exclusive	
MGQT		
Innocent	-0.80	1.00
Guilty	-2.00	-1.20
ZOC		
Innocent	0.20	3.20
Guilty	-0.80	-2.00

Appendix U

Table 32

Examination Mean Total Numerical Scores: For All Relevant Questions (5 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	8.60	13.00
Guilty	-15.60	-5.80
ZOC		
Innocent	15.20	10.20
Guilty	-16.00	-8.00
	Exclusive	
MGQT		
Innocent	5.00	-2.60
Guilty	-11.40	-4.20
ZOC		
Innocent	-1.40	6.60
Guilty	-4.90	-13.80

Table 33

Examination Mean Total Numerical Scores: Using Three Relevant Questions (3 in MGQT and 3 in ZOC)

	Type Of Question	
	Nonexclusive	
	Gender	
Type Of Test	Male	Female
MGQT		
Innocent	4.50	10.80
Guilty	-14.10	-4.40
ZOC		
Innocent	15.20	10.20
Guilty	-16.00	-8.00
	Exclusive	
MGQT		
Innocent	0.30	-1.20
Guilty	-8.70	-3.80
ZOC		
Innocent	-1.40	6.60
Guilty	-4.90	-13.80

LIST OF REFERENCES

LIST OF REFERENCES

Ash, P. (1973). Survey of attitudes on the polygraph. Polygraph, 2, 200-223.

Backster, C. (1965). Standardized polygraph notepack and technique guide. New York: Backster Research Foundation.

Barland, G. H. (1983). A dual issue format. Polygraph, 12(1), 1-6.

Barland, G. H. (1988). The polygraph test in the U.S.A. and elsewhere. In A. Gale (Ed.), The polygraph test: lies, truth and science (pp. 73-95). London: Sage.

Barland, G. H., Herbold-Wooten, H., Orne, M. T., & Yankee, W. (1987). Improving polygraph research standards: A report to the Department of Defense by the Department of Defense Polygraph Institute study group. Polygraph, 16(4), 264-284.

Barland, G. H., & Raskin, D. C. (1973). Detection of deception. In W. F. Prokasy & D. C. Raskin (Eds.), Electrodermal activity in psychological research (pp. 417-477), New York: Academic Press.

Barland, G. H., & Raskin, D. C. (1975). An evaluation of field techniques in detection of deception. Psychophysiology, 12(3), 321-330.

Bradley, M. T., & Janisse, M. P. (1981). Accuracy demonstrations, threat and the detection of deception: cardiovascular, electrodermal, and pupillary measures. Psychophysiology, 18(3), 307-315.

Bradley, M. T., & Ainsworth, D. (1984). Alcohol and the psychophysiological detection of deception. Psychophysiology, 21(1), 63-71.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York: Academic Press.

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

Dawson, M. E. (1980). Physiological detection of deception: measurement of responses to questions and answers during countermeasure maneuvers. Psychophysiology, 17(1), 8-17.

Driscoll, L. N., Honts, C. R., & Jones, D. (1987). The validity of the positive control physiological detection of deception technique. Journal of Police Science and Administration, 15(1), 46-50.

Ekman, P. (1985). Telling lies. New York: W. W. Norton.

Forman, R. F., & McCauley, C. (1986). Validity of the positive control polygraph test using the field practice model. Journal of Applied Psychology, 71(4), 691-698.

Frude, N. (1987). A guide to SPSS/PC+. New York: Springer-Verlag.

Furedy, J. J. (1986). Lie detection as psychophysiological differentiation: Some fine lines. In M. Coles, E. Donchin, & S. Porges (Eds.), Psychophysiology: Systems, processes, and applications-A handbook. New York: Guilford.

Furedy, J. J. (1988). Validity of the lie detector: A psychophysiological perspective. Criminal justice and behavior, 15(2), 219-246.

Gatchel, R. J., Smith, J. E., & Kaplan, N. M. (1984). The effect of propranolol on polygraphic detection of deception. Unpublished manuscript, University of Texas, Health Science Center, Dallas.

Ginton, A., Daie, N., Elaad, E., & Ben-Shakhar, G. (1982). A method for evaluating the use of the polygraph in a real life situation. Journal of Applied Psychology, 67(2), 131-137.

Gustafson, L. A., & Orne, M. T. (1964). Effects of task and method of stimulus presentation on the detection of deception. journal of Applied Psychology, 48(6), 383-387.

Hammond, D. L. (1980). The responding of normals, alcoholics and psychopaths in a laboratory lie-detection experiment. Unpublished doctoral dissertation, The California School of Professional Psychology, CA.

Hays, W. L. (1981). Statistics (3rd ed.). New York: Holt, Rinehart and Winston.

Hikita, Y., & Suzuki, A. (1965). Construction and interviewing of probing peak of tension tests for some other crimes suspected, and their results. Research Materials No. 35, Polygraph Reports, National Research Institute of Police Science [Text in Japanese], 39-50.

Honts, C. R. (1986). Countermeasures and the physiological detection of deception: a psychophysiological analysis. Unpublished doctoral dissertation, University of Utah, Salt Lake City.

Honts, C. R. & Driscoll, L. N. (1989). Validity of the positive control polygraph test: comments on Forman and McCauley. Polygraph, 18(3), 158-167.

Honts, C. R., & Hodes, R. L. (1982). The effects of simple physical countermeasures on the detection of deception. Psychophysiology, 19, 564-565.

Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985). Effects of physical countermeasures on the physiological detection of deception. Journal of Applied Psychology, 70(1), 177-187.

Horvath, F. (1980). Polygraphy: Some comments on the state of the art. Polygraph, 9, 34-41.

Horvath, F. (1984). Detecting deception in eyewitness cases: Problem and prospects in the use of the polygraph. In G. L. Wells & E. F. Loftus (Eds.), Eyewitness testimony: psychological perspectives (pp. 214-255). New York: Cambridge University Press.

Horvath, F. (1985). Liars and lie detection: Job screening. Social Science and Society, 22, 43-46.

Horvath, F. (1987). Public attitudes on polygraph testing: A national survey. Polygraph, 16(1), 1-17.

Horvath, F. (1988). The utility of control questions and the effects of two control question types in field polygraph techniques. Journal of Police Science and Administration, 16(3), 198-209.

Hunter, J. E. & Schmidt, F. L. (1990). Methods of Meta-analysis. Beverly Hills, CA: Sage.

Hunter, J. E., Schmidt, F. L. & Jackson, G. B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Hurd, S., (1985). Use of the polygraph in screening job applicants. American Business Law Journal, 22, 529-549.

Iacono, W. G. & Patrick, C. J. (1987a). What psychologists should know about lie detection. In I. B. Weiner & A. Hess (Eds.), Handbook of forensic psychology (pp. 460-489). New York: Wiley.

Iacono, W. G. & Patrick, C. J. (1987b). Validity and reliability of the control question polygraph test: A scientific investigation. Society for Psychophysiological Research Abstracts, 24(5), 604-605.

Iacono, W. G. & Patrick, C. J. (1988). Assessing deception: Polygraph techniques. In R. Rogers (Ed.), Clinical assessment of malingering (pp. 205-233). New York: Guilford Press.

Jayne, B. C. (1986). Control question theory in the polygraph technique. Polygraph, 15(4), 245-254.

Keppel, G. (1982). Design and analysis: a researcher's handbook (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kircher, J. C. (1983). Computerized decision making and patterns of activation in the detection of deception. Unpublished doctoral dissertation, University of Utah, Salt Lake City.

Kleinmuntz, B. & Szucko, J. J. (1982). On the fallibility of lie detection. Law and society review, 17(1), 85-104.

Larson, J. (1932). Lying and its detection. Chicago: The University of Chicago Press.

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

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

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

Lykken, D. T. (1979). The detection of deception. Psychological Bulletin, 86(1), 47-53.

Lykken, D. T. (1981). A tremor in the blood: uses and abuses of the lie detector. New York: McGraw-Hill.

Lykken, D. T. (1988a). Detection of guilty knowledge: A comment on Foreman and McCauley. Journal of Applied Psychology, 73(2), 303-304.

Lykken, D. T. (1988b). The polygraph test in the U.S.A. and elsewhere. In A. Gale (Ed.), The polygraph test: lies, truth and science (pp. 73-95). London: Sage.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D. H. (1975). Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill.

Norusis, M. J. (1988). SPSSX introductory statistics guide. Chicago: SPSS Inc.

Norusis, M. J. & SPSS Inc. (1988). SPSS/PC+ V2.0. Chicago: SPSS Inc.

Office of Technology Assessment (1983). Scientific validity of polygraph testing: A research review and evaluation. (OTA-TM-H-15) Washington D. C.: U. S. Congress, Office of Technology Assessment.

Orne, M. T. (1972). Implications of laboratory research for the detection of deception. Polygraph, 2(3), 169-199.

Patrick, C. J. & Iacono, W. G. (1989). Psychopathy, threat, and polygraph test accuracy. Journal of Applied Psychology, 74(2), 347-355.

Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. Psychological Bulletin, 84(4), 782-799.

Podlesny, J. A., & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. Psychophysiology, 15(4), 344-359.

Raskin, D. C. (1979). Orienting and defensive reflexes in the detection of deception. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), The orienting reflex in humans (pp. 587-605). Hillsdale, New Jersey: Erlbaum.

Raskin, D. C. (1982). The scientific basis of polygraph techniques and their uses in the judicial process. In A. Trankell (Ed.), Reconstructing the past: the role of psychologists in criminal trials (pp. 319-371). Stockholm: Norstedt and Soners.

Raskin, D. C. (1988). Does science support polygraph testing. In A. Gale (Ed.), The polygraph test: lies, truth and science (pp. 96-110). London: Sage.

Raskin, D. C., Barland, G. H., & Podlesny, J. A. (1978). Validity and reliability of detection of deception (U.S. Department of Justice: 1978 0-267-969). Washington D.C.: U.S. Government Printing Office.

Raskin, D. C., & Hare, R. D. (1978). Psychopathy and detection of deception in a prison population. Psychophysiology, 15(2), 126-136.

Reali, S. F. (1978). Reali's positive control technique. Polygraph, 7, 281-285.

Reid, J. & Inbau, F. (1977). Truth and deception: The polygraph (lie-detector) technique. Baltimore: Williams & Wilkins.

Reynolds, H. T. (1984). Analysis of nominal data (2nd ed.). Sage University Paper series on Quantitative Applications in the Social Sciences, 07-007. Beverly Hills and London: Sage.

Rovner, L. I. (1979). The effects of information and practice on the accuracy of physiological detection of deception. Unpublished doctoral dissertation, University of Utah, Salt Lake City.

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

Weir, R. (1974). In defense of the relevant-irrelevant polygraph test. Polygraph, 3(1), 119-166.

Widacki, J., & Horvath, F. (1978). An experimental investigation of the relative validity and utility of the polygraph technique and three other common methods of criminal identification. Journal of Forensic Sciences, 24(3), 596-601.

Yankee, W. J., & Grimsley, D. L. (1986). The effect of a prior polygraph test on a subsequent polygraph test (Contract No. MDA 904-84-C-4249). Washington, DC: National Security Agency.