Polygraph Examinations Using the CQT With Actual Truth Control Questions

by

Murray Cosmas Cullen

M.A. (Clin. Psyc.), Lakehead University, 1984

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the Graduate Academic Unit of Psychology
University of New Brunswick

Supervisor: M.T. Bradley, Ph.D, Psychology

Examining Board: R. Hill, Ph.D., Economics, Chair

G. Marquis, Ph.D., History

R. Nicki, Ph. D., Psychology

M.T. Bradley, Ph. D., Psychology

External Examiner: G. Barland, Ph.D.,
Department of Defense,
Polygraph Institute, U.S.A.

THE UNIVERSITY OF NEW BRUNSWICK

April 2000

Dedication

I dedicate this entire project to my loving wife Beth, who took on more than her share of parenting, housekeeping, and stresses so that I could pursue a dream to obtain a higher education.

Abstract

Participants guilty of a mock crime, innocent and informed of the details of the crime or innocent, and uninformed of the details were examined on the polygraph with a modified version of a Control Question test. Normally, this test contains questions that are incriminating, ambiguous, and likely to be answered with a lie. The modification involved the replacement of incriminating and ambiguous control questions with lie engendering control questions that were unambiguous and answered truthfully. In order to examine for the potential effects of the orientation response and habituation, the relative position of control and crime relevant questions were altered such that in one condition the control question was first and in the other it was second.

Chi Square analyses showed both guilty and innocent participants were classed as guilty when the crime relevant question was presented first. When control questions were first, guilty participants tended to be classed as guilty whereas innocent participants were classed mainly as innocent. Secondary analyses explained the classification results. Scores derived from skin resistance responses and blood volume differed over the order of presentation and conditions. They indicated guilt when the crime relevant question was in the first position, but only did so in the guilty condition when the crime relevant question was in the second position. An ANOVA on raw physiological scores indicated that the habituation of physiological measures occurs quickly.

Sections in the introduction deal with the history of interrogations with polygraphs, instrument development, measurement considerations, the theory of the control question test as well as other techniques of questioning.

Acknowledgements

I must acknowledge my mentor, Dr. Michael Bradley, who has invested his time, letters of recommendations, and inexhaustible degree of effort over the past 19 years. He allowed me to become a research assistant for him back in the fall of 1980. Soon after, he assisted in my efforts to attend graduate school in Ontario, which allowed me to eventually obtain a Masters Degree in Clinical Psychology and, in turn, become a Licensed Psychologist. Dr. Bradley was instrumental in my acceptance into the Doctoral Degree program at UNB. He has demonstrated immeasurable patience in slowly transforming me into a Clinical Psychologist, and I fully realise that was no small feat.

Acknowledgement must go to Dr. Paul Gendreau, who was greatly instrumental in my acceptance in the Doctoral program, and who allowed me to participate as a student on the front end of his Meta-Analysis research.

Acknowledgement must also go to Dr. Murray Goddard, who invested considerable effort in reviewing my paper and providing directions for improvement. Dr. Don Fields introduced me to Chi-Square analysis which, in turn, helped me to discover an important interaction effect in this research.

I must also thank Michael Black for helping me collect some of the data contained in this document. In fact, Mike, together with Vance MacLaren and Steve Carle (fellow graduate students) helped to create an enthusiastic thirst for detection of deception research such that it kept my spirits up and nose "to the grindstone".

TABLE OF CONTENTS

List of Tables

Chapter 1

Introduction

Throughout history, humans have attempted to deceive each other (Trovillo, 1939). Their deceptions have ranged from those involving trivial matters to lies with profound consequences. Human attempts to purposely deceive have provided motivation to develop ways to detect such attempts.

It has long been thought that physiological measures may be one method of exposing deception. A Hindu medical source, dating back to approximately 900 BC, described a situation in which alleged poisoners supposedly revealed their guilty status by the physiological response of blushing (Trovillo, 1939). Eristratus, a third century physician to the crown prince of Syria, used the "tumultuous" rhythm of the heart to reveal the prince's true thoughts (Trovillo, 1939). During the Spanish Inquisition, individuals were forced to swallow a slice of bread together with cheese. A verdict of guilt was delivered if the food stuck in their throat (Kleinmuntz and Szucko, 1984).

From these modest folkloric beginnings, the groundwork was laid for attention, and ultimately experimentation, in the area of detecting deception. Developments were made by Lombroso who used blood volume together with vasomotor activity as part of an interrogation procedure to solve crime (as cited in Trovillo, 1939). In doing so, he helped establish a basis for modern polygraphy. Marston (1917), in a laboratory demonstration, reported a 96% accuracy rate in detecting lying by utilising blood volume. At about the same time, Benuossi (1914, as cited in Ben-Shakhar and Furedy, 1990) examined slight

changes in the respiration rate of criminal suspects as an indication of deception (Ben-Shakhar and Furedy, 1990, p.3). Soon after, Burtt used respiration to accurately detect subjects' lies in 73% of cases (Trovillo, 1939, p.870). Larson created the forerunner of the modern polygraph by constructing a device that measured blood volume, pulse rate, and respiration simultaneously. His associate, Leonarde Keeler, modified it into a portable "field" model similar to the modern models presently in use today (Kleinmuntz and Szucko, 1984).

Before an explanation of the modern polygraph device can be given, however, it is necessary to understand, in more detail, how an individual's physiology can be associated with the detection of deception.

Physiology Associated With the Theory of Detection of Deception

Human physiological response systems are ultimately controlled by the nervous system. The nervous system is divided into two major components: the central nervous system, or CNS, which consists of the brain and spinal cord, and the peripheral nervous system, or PNS. The PNS consists of the somatic division, made up of sensory and motor neurons, and the autonomic division, made up of the sympathetic and parasympathetic nervous systems. The peripheral nervous system detects internal and external environmental changes and reports this information to the CNS, which in turn receives and processes this incoming information. The autonomic divisions of the peripheral nervous system reflect aspects of CNS activities and responses. The magnitudes of particular responses evidenced by the autonomic nervous system are based on novelty (Sokolov, 1966, as cited in Ben-Shakhar

and Furedy, 1990), meaningfulness, provocativeness, and the emotive properties (Martindale, 1981) of the stimuli. The autonomic nervous system transmits information to and from glands and smooth muscles, and generally governs these bodily processes over which individuals have little or no conscious control (Cacioppo and Petty, 1983).

The sympathetic component of the autonomic nervous system in the main serves to excite the organs and glands. The parasympathetic component operates in an almost opposite manner to quiet or calm down the internal environment. A variety of measures, including those mentioned earlier (i.e. electrodermal and cardiovascular) reflect autonomic nervous system activity. The process is complex because, as some systems related to fighting or fleeing are activated (such as heart rate, and insulin secretion), other systems related to long term maintenance (such as digestion and salivation) may be inhibited. In addition, these systems can trigger hormonal effects to further extend or inhibit various glandular effects. Overall, these two branches help the autonomic nervous system maintain homeostasis for the organism.

Since the struggle to maintain homeostasis is dynamic, rapid adjustments made in response to excitations are useful in polygraph investigations. Attempts to measure, record, and capture such rapid physiological adjustments for the purpose of aiding criminal investigations have resulted in the polygraph device.

The Polygraph Device

The polygraph is a multi-channel recorder that can monitor an individuals' physiological events at particular points in time and record them either on paper charts or in computer

files. It typically consists of the following components: 1) a transducer which receives physiological information from its sensing devices (such as a blood pressure cuff or skin conductance electrodes) and converts the information to an electrical signal; 2) a coupler that adapts the recorder circuitry in order to accept the transducer signal; 3) a preamplifier and power amplifier to initially boost the signal and then to increase the power to a sufficient level so as to operate the pen or tape, and finally; 4) either chart paper or computer files on which to record the output from the various physiological measurement devices (Stern, Ray, and Davis, 1980).

Whether responses are recorded in the traditional manner (on chart paper) or with the more technologically advanced method of capturing the data on computer files, the use of the polygraph in assisting in the detection of deception has become, and remains, widespread.

Prevalence in the Use of the Polygraph

According to Krapohl (1996), the polygraph has proven to be a powerful tool in searching for the truth. It is extensively applied with law enforcement and the U.S. government in security investigations. Approximately 60% of the large police departments in the U.S. use this technique in their pre-employment screening process (Kiang, 1996). In addition to the U.S., Kiang indicates that several other countries such as Canada, India, Israel, and Japan employ the use of the polygraph. The reliance on this technique continues and Malaysia has recently introduced the use of the polygraph by their country's police force (Kiang, 1996).

Surprisingly, reliance on polygraph testing has occurred in the absence of solid scientific evidence (Saxe, 1991). There has been much published research but it has remained difficult to resolve important issues. One focus of the research has been on physiological measures.

Measurement Examination

There has been a great deal of research devoted to discovering the most effective measures in detecting deception. Not only have a number of potential measures been considered (i.e. respiration, blood volume, skin resistance response), but a number of ways of assessing each measure, such as the time involved in the measure returning halfway to baseline and the area under the respiration curve, have been examined (Cutrow, Parks, Lucas, and Thomas, 1972).

Kircher and Raskin (1988) conducted one of the more ambitious studies. They utilised computer algorithms to process physiological reactions to a polygraph questioning technique (the Control Question Test) with data obtained from two mock crime experiments. Data from one hundred subjects (N=100), a standardisation sample, was used to develop a discriminant function that included weightings from electrodermal, cardiovascular, and respiration measures. Data was then collected from forty-eight additional subjects and used to cross validate the computer model.

Eighteen values for each response parameter were obtained from each subject for each of eighteen questions. They were measured by assessing chart tracings.

Highly redundant measures were eliminated and the set was reduced to twelve variables that provided relatively independent sources of information. An all-possible-subsets regression analysis was then completed to identify a subset that would best discriminate between groups, resulting in a subset of 5 variables. These variables were for skin conductance: amplitude, recovery time, and electrodermal burst frequency; for blood volume: amplitude, and for respiration: length. Skin conductance accounted for 61% of the predictable variance in the standardised sample and was the most useful measure discriminating between guilt and innocence.

In several reviews of experimental studies (Ben-Shakhar and Furedy, 1990; Kugelmass and Lieblich, 1968; Thackray and Orne, 1968b), the electrodermal measure (skin resistance response or "SRR") was superior to other measures. This apparently holds true in both laboratory and field studies (Raskin, Kircher, Horowitz, and Honts, 1989). There has been one study (Timm, 1982) in which the SRR has been inferior to another measure (respiration), but this result has not been replicated.

Cutrow et al. (1972) assessed a wide variety of physiological measures including: breathing amplitude, breathing cycle time, eye-blink rate, eye-blink latency, finger pulse volume, heart rate, palmar galvanic skin response, volar forearm galvanic skin response, and voice latency. Although all measures obtained significance between response rankings to guilty and innocent questions, the electrodermal measures remained superior. Recently, electro-enchephalographic measures have shown promise in lie detection (Honts, Raskin, and Kircher, 1987).

There have been instances in which certain physiological measures, other than the SRR, have produced better-than-chance detection accuracy, but the results have not been uniform. Respiration was useful in a study by Kugelmass and Lieblich (1968) as well as in Timms' investigation (1982) but did not perform at better-than-chance levels in Thackray and Orne's study (1968). Elaad (1987) did not find blood volume a useful measure but Kugelmass et. al., (1968) found that it performed at better than chance levels. Five studies used voice stress analysis, and found no evidence that it was a better-than-chance indicator of deception (Ben-Shakhar and Furedy, 1990, p. 89).

Questioning Formats

Regardless of the particular measures utilised, a variety of tests have been formulated to attempt to assess physiological activity associated with crime-related questioning concerning police investigations. This paper will explain two such tests. The first is referred to as the Control Question Test, or "CQT". The CQT is widely utilised in police work. This will be followed by discussion of the Guilty Knowledge Test, or "GKT". The GKT has been extensively tested in laboratory settings by academic psychologists (e.g. Lykken, 1959). The features of each test will be pointed out. A new question test, containing elements of both the CQT and the GKT, will be introduced later as well as an explanation of how it performed in the present research.

The Control Question Test

John Reid introduced the Control Question Test or "CQT" (Reid, 1947) and, according to Reid and Inbau (1977), it has become the main tool in criminal polygraph investigations. Central to this technique is the use of control questions as well as event-related questions. It was designed to evoke the attention of innocent suspects to non-crime-related items and the attention of guilty suspects to crime relevant information.

Prior to the polygraph examination, the typical CQT procedure includes a lengthy (up to 2 hours) interview involving the polygraph examiner and the suspect (participant) from which a list of both relevant and control questions are formulated. Control questions are devised such that the suspect can deny events, but not without some misgivings, doubts, or even lies. The examiner purposefully formulates the control questions so that the innocent participants will focus concern on the control questions. An example of a Control Question Test follows:

(1) Your first name is ------?

(2) Are you going to tell me the entire truth during this examination?

(3) Do you believe that I will only ask the questions that we have reviewed in this test?

(4) Have you ever stolen money from a friend?

(5) Did you stab a man at the bar last night?

(6) Have you ever cheated someone you know?

(7) Did you rob the man of his wallet?

(8) Is your last name ------?

(9) Did you ever wish to hurt a friend seriously?

(10) Did you hide the knife after your attack?

The first three questions, together with question 8, are not scored as they allow the subject to adapt to the novelty of being in a polygraph examination.

The crime relevant and control questions are organised in 3 pairs such that one member of each pair is a control question and the other is crime-relevant. Typically, the control questions are numbers 4, 6, and 9. They refer to issues that are purposely unsettling to the suspect because they are both incriminating and ambiguous. It is quite often the case that the suspects are somewhat uncertain of the truthfulness of their answers. They may actually have behaved in such a fashion over the course of their life.

It is reasoned that if the control questions are sufficiently incriminating, somewhat vague, and evocative of a lie, they will raise the emotional level among the innocent because of their intense desire and necessity of appearing truthful in the testing circumstance. An important assumption here is that the wording of the control questions could cause uncertainty and stress, with a concomitant physiological reaction. In contrast, innocent suspects are unambiguously truthful on crime relevant questions. Therefore, responsivity should be greatest to control questions.

Questions 5, 7, and 10 typically are relevant to the crime under investigation. Guilty suspects focus on the crime-relevant questions because a) they are unambiguously lying on those questions and b) if their lies are discovered they will face the consequences of their crime. Therefore, guilty suspects should produce stronger physiological reactions to crime relevant questions than control questions.

The responses to the control question and crime relevant questions are compared in order to determine guilt or innocence. If a response to a particular control question is greater than that to the appropriate crime relevant question, a positive score is assigned. If the opposite occurs, a negative score is given. The scores vary in a range that depends upon the number of presentation sets and the number of physiological channels recorded. In general, however, the more negative the score, the more likely the individual will be considered guilty whereas a positive score is considered indicative of telling the truth. Scores near zero are considered inconclusive.

Shortcomings of the Control Question Test

Widespread reliance on polygraph testing has occurred in the absence of solid scientific evidence (Saxe, 1991). From a theoretical viewpoint many important issues are still unresolved. For example, issues pertaining to the standardisation of interviews, criticisms of ambiguity and vagueness of crime-relevant questions, and the transparency and comparability of question pairings are not resolved. There is also the concern that different evoked emotions may produce similar physiological responses. The latter may also effect the rate of false positives.

One problem involves the nonstandardization of questioning techniques. Standardisation is required for any psychological test (Anastasi, 1988). Standardisation demands a uniformity of procedure in the administration and scoring of the test. This is necessary so that scores obtained by various persons may be comparable. Issues for both crime relevant and control questions are explored in a pre-test interview. That influence is

intended to carry over to the actual administration of the test. While there is uniformity in the goals of the pre-interrogation interview, it is difficult to standardise goals. An attempt to match control and crime relevant questions in terms of severity of actions is made but there are such a variety of crimes that each interrogation is somewhat unique and therefore not fully standardised. Barland and Raskin (1973) suggested that the control questions were not scientific in the normal sense, but rather designed to provide a type of emotional standard. Nevertheless, no matter how much on the surface two CQT's may look the same, the rational conclusion is that the interviews were not standardised. Consequently, the CQT itself is non-standardised.

Another problem with the CQT is that the questions may be of uneven quality. For example, in some cases the question pairings are such that the control question is ambiguous ("Did you ever steal something valuable?"; "Did you ever cheat in any way?"). The issue is that the control questions are not true scientific controls. "Comparison questions", as implied by Barland and Raskin (1973), may be a better term. Some researchers may contend that if only one variable is changed in an experimental condition, the experimenter may not know what was specifically included in the control condition. That condition, nevertheless, may still act as a control. Other researchers, however, would contend that in science, if you cannot specify exactly what you are controlling for, it is difficult to pinpoint exactly why different questions evoke differential responses. Is the inference of deception completely justified or could it be something else? Of course, this is what makes the area controversial in that empirically lie detection is effective but errors occur.

Ben-Shakhar and Furedy (1990) argue that not only are crime relevant and control question pairings not comparable but also that the CQT is transparent. Both innocent and deceptive subjects know that the crime relevant questions are the most important questions to "pass successfully" to avoid a judgement of guilt. Due to reactions to experiencing emotions such as stress, worry, and fear, an innocent subject may produce similar physiological responses as a guilty subject. Therefore, several investigators believe that the test would find an unacceptable number of innocent subjects as guilty, and they cite laboratory evidence to back this view (Ben-Shakhar and Furedy, 1990; Lykken, 1981; Saxe, Dougherty, and Cross, 1985).

Even with this common complaint of finding guilt in an unacceptable number of innocent subjects, Iacono and Patrick (1988) believe that the laboratory setting overestimates the accuracy of the CQT. From their perspective, they argue that the CQT works well in the lab for reasons not applicable to field situations. They believe that students guilty of mock crimes in a laboratory setting are more likely to be found guilty since they possess very little emotional reason to try to "beat" the test (no real consequence to determination of guilt). Conversely, innocent subjects are more reactive to the incriminating control questions than to questions about a pretend crime in which they were not involved.

Lykken (1981) has focussed on the high false positive rate (misclassification of innocent subjects as guilty) found in laboratory studies. He believes that these rates, which can be as high as 49% (Szucko and Kleinmuntz, 1981) occur in field situations. From his perspective, the fact that suspects are being interrogated for a real crime together with their

fear of consequences of appearing deceptive, makes them more physiologically reactive to crime relevant questions than to control questions. As a result, an unacceptably high "false positive" rate is produced (Lykken, 1981). For example, a direct question of asking the suspect if he shot a man potentially carries greater consequences than asking him if he ever cheated somebody. Anyone, even innocent individuals could be nervous and responsive, knowing a verdict of guilt may lead to a term of imprisonment.

Iacono and Patrick (1988) suggested that criteria should be developed and applied to the validation of polygraph testing. These criteria would be similar to those involved in the evaluation of psychological and medical diagnostic tests. They suggested that experimental and control groups be employed and that the procedures should: a) utilise real life cases rather than simulations; b) allow the researchers to be able to determine "ground truth" (actual guilt or innocence); c) allow for "blind analysis" of the polygraph charts (without access to the case facts and any other information about the subjects). Even though Iacono and Patrick recommend field studies, past experience with field studies has led to ambiguous results. The problems are numerous and it is difficult to know what is absolute truth or guilt. A "confession" may be forced or supplied to protect the real culprit. A court decision of "guilt" is only based on the available evidence and may, in fact, be in error. Investigators in the field do not necessarily conduct their review of polygraph charts "blindly". They are usually aware of key information about the subjects they have questioned.

The complexity of the issues raised by the CQT format has made alternative approaches that circumvent the problem seem attractive.

The Guilty Knowledge Test

The Guilty Knowledge Test rests on readily identifiable sound scientific principles. The assumptions underlying the GKT are plausible, compatible with psychological theory, and supported by research (Ben-Shakhar and Furedy, 1990). Further, each test administration can be designed in a standard and objective fashion.

The GKT consists of several multiple-choice items, only one of which is relevant to the crime and known only by those familiar with the crime. In this manner, the GKT presents unequivocal items. The other choices are unrelated to the crime under investigation but in all other respects appear to be equivalent to the relevant choice. An example of such an item would be:

(1) The victim who was murdered was -----

(a) shot to death.

(b) strangled to death.

(c) beaten with a candlestick to death.

(d) stabbed to death.

(e) suffocated to death.

In the above example, only those guilty or in the possession of guilty knowledge (spouse, accomplice, witness, etc.) would have known what the relevant answer was and, in turn, would have shown differential responsivity to the item. If, over the course of several items, the suspect consistently exhibited differential responsivity to the relevant choices, it would then be judged that the suspect had guilty knowledge concerning the crime under

investigation. Only 4 of the 5 items in each question are scored, as it is believed that a large Orientation Response (OR), explained later, would occur for the first item.

Shortcomings of the Guilty Knowledge Test

Although the GKT possesses promise, a recent field study did not yield results that readily compare to those found in laboratory studies. Elaad, (1990), utilised the GKT in the field for real criminal investigations. Although this method (when used with the skin resistance response measure) resulted in a high accuracy rate of 97.9% for the innocent subjects, the detection rate of 42% among the guilty subjects was much lower than expected. For example, the results of 8 laboratory studies reported by Lykken (1988) suggested a detection rate of 88% among the guilty.

Elaad, Ginton, and Jungman (1992) point out that the differences in the detection rates could be explained by the possibility that, in the field, the suspects may not remember all the necessary information whereas in the laboratory the likelihood of guilty subjects being aware of all of the necessary information is high. M.T. Bradley (personal communication, August, 1999) pointed out that laboratory studies often require participants to read the information. The participants are students and the GKT exam is similar to a multiple-choice test. Students are well trained to pick up the information that the experimenter intends them to notice. This is in contrast to real life, where certain information about the crime could be overlooked and lost in the excitement of the event. The suspects are unlikely to be students or to study their crime. The multiple-choice GKT is not as likely to be familiar to the suspects. Further, most academic research involving the

polygraph utilises a set of procedures that include examining the subjects immediately after the commission of the crime whereas in real life it may be days, weeks, and possibly even months later.

A final explanation was offered by Elaad (1990) who suggested that only a single measure was used to determine accuracy, thus leading to the assumption that the use of multiple measures might increase detection rates. Elaad, Ginton, and Jungman (1992) found that utilising a combination of respiration line length and skin resistance response significantly increased the accuracy of detection among the guilty subjects. Elaad (1994) found similar results for combining those two measures. From his perspective, the integration of two efficient measures can increase the likelihood of detection of deception.

Regardless of the number of measures issue, Elaad's (1990) inability to effectively detect guilty subjects with the GKT procedure may explain why most "field" users of the polygraph (consisting mostly of investigative agencies such as the police) continue to prefer to use the Control Question Test (CQT) technique.

Presently we know that physiological changes rapidly occur in response to an external stimulus (for example, a Control Question Test item). Further, the technology exists to record and measure those responses via the polygraph. A theoretical context is still necessary in order to interpret the meaning of those responses. The next section will explore the theories underlying lie detection.

Theories in the Detection of Deception

Theoretical considerations in the detection of deception have been formulated in concert with the testing techniques developed. In a generic fashion, the Control Question Test and the Guilty Knowledge Test each represent one of the two major approaches in detection techniques. Control Question Tests rely on the assumption that suspects will have large responses to questions on which they are lying whereas GKT's are created on the assumption that suspects will respond more to items of information they recognise. Both tests involve emotion and cognition. The Control Question Test, because it relies on responses to lies and negative life events, is usually considered to be based more on emotion. Cognitions play a significant role but theoretical considerations of that role have not been pursued as vigorously as with the GKT.

The Guilty Knowledge Test specifically focuses on cognitive aspects, such as the knowledge of specific items. Emotions may play a role in relevant physiological responses, but that aspect has been less explored with this test.

Although emotions and cognitions are interrelated, theories focusing on each concept are considered separately.

A) Emotions

The ideal explanatory potential for emotion-based theories could be considered in two steps. First, specific emotions would have to be related to differentiated physiological responses. Second, particular emotions would have to be regularly and reliably associated with lying.

There have been many attempts to discover physiological and behaviour patterns linked with specific emotions. For example Levenson, Ekman, and Friesen (1990), reported differential physiological responding associated with different emotional expressions. According to Ortony and Turner (1990), however, most of the many attempts to discover specific physiological patterns of responses and neurological underpinnings of such patterns have not been very successful. Theorising is not entirely back to the notion of general arousal underlying all emotions (Cannon, 1927), but potential differences in physiological responses are small and it is not known what role differential emotions play in lying. More specifically, Lykken (1981) notes that there is no characteristic lie response. Rather, the polygraph probably records autonomic changes reflecting a general arousal.

Heslegrave (1981) found that increased skin conductance resulted from conflict involved in deception. Instead of increases of sympathetic activity resulting in excitatory arousal however he found, from heart rate activity measures, evidence of increases in parasympathetic activity reflecting a type of inhibitory arousal.

Regardless of issues concerning patterns of responses or general arousal, Davis (1961) formulated three explanations as to how emotions could possibly influence the physiological detection of deception. These explanations were centered on fear of punishment; conflicting response tendencies; and conditioning.

Punishment theory focused on the potential negative consequences of the polygraph examination. The theory rests on the assumption that enhanced reactions to relevant questions are the result of fear of being judged either deceptive or guilty (Ben-Shakhar & Furedy, 1990). If the severity of punishment for a judgement of guilt were increased, guilty

subjects' responses to relevant questions would be "enhanced". Theoretically then, their guilt would be more detectable.

Kugelmass and Lieblich (1966) tested the effects of fear by manipulating the threat of punishment. They set up two groups: in one, members of a control group were told they were being subjected to a lie detection examination to determine if the equipment was operating in a proper fashion. Members of the second group were told they were being tested on their ability to control their emotions. Further, only those individuals who successfully exhibited such control would be viewed as appropriate candidates for a police force. Despite the threat, there was no significant difference in responsivity between the two groups. Their results did not support the punishment theory.

Similar results were found by Bradley and Janisse (1981) who threatened subjects with an electric shock if they were judged to be deceptive. The threat of shock did not affect detection rates. This result may be due to the fact that it was a laboratory-based study. In accordance with at least one of the standards suggested by Iacono and Patrick (1988), such subjects do possess reasons to "beat" the test even though the studies were laboratory based. The threat of punishment made no difference.

The conflict approach examines how a subject responds to the questions while they are being tested. If faced with a relevant question, they must either lie or tell the truth (Ben-Shakhar & Furedy, 1990). An emotional reaction to this conflict problem (telling the truth or lying) triggers a physiological response that can be recorded and subsequently compared to responses from non-crime relevant and therefore non-conflicting questions.

The most directly relevant study provides support for the conflict approach (Forman and McCauley, 1986). They utilised the Positive Control Test. The subject is required to tell both the truth as well as a lie for each question in the interrogation. For example ... "Tell me the truth, did you steal the twenty dollars? Now tell me a lie, did you steal the twenty dollars?"

It was assumed that guilty subjects would be less aroused by telling the instructed lie, which was the truth for a guilty person, than by telling the instructed truth, which was a lie (or conflict) for the guilty person. Conversely, innocent subjects should have been more aroused by telling the instructed lie, which was a lie for them, than by telling the instructed truth, which was the actual truth for them. Subjects (innocent or guilty) were more responsive to lying and this result supported the conflict theory.

Studies on lie detection tests for knowledge of specific information have also manipulated response conflict and have produced mixed results. In these studies, subjects were asked to respond to a series of questions (i.e. "Did you steal the 20 dollars?") either with a "no" (presenting a conflict to the guilty subjects on the key item), or "yes" (presenting no conflict on the key item). Kugelmass, Lieblich, and Bergman (1967) results supported the conflict effect. Elaad and Ben-Shakhar (1989), Furedy and Ben-Shakhar (1991), and Horneman and O'Gorman (1985) did not find such an effect.

Conditioned response theories rest on the assumption that an individual reacts to particular questions because they have been conditioned by their past experiences. The more serious the experience, the stronger will be the reactions evoked by cues relevant to

that experience (Ben-Shakhar and Furedy, 1990). Disgust, fear, shame, or anger may be evoked if the crimes were particularly graphic and violent.

Bradley and colleagues conducted a series of studies (Bradley and Cullen, 1993, Bradley, Cullen, and Carle, 1993, 1996 and Carle, 1996) that examined the possibility that the emotions associated with experiences, including crimes, could influence the accuracy rates of the detection of deception. The predictions were that the emotions created by experiences should increase the subject's physiological reactivity to the point where they should be more detectable through a polygraph examination than subjects who did not share the same experience. This series of research studies produced mixed results.

Bradley and Cullen (1993) examined questions that reminded subjects of a real event. They asked university students to write about an embarrassing incident from their own experience, which had a significant emotional impact on them. These experiences were of such a nature that the subjects preferred that no one knew about them and they were motivated to deny their involvement. The subjects were then examined using the CQT on two stories, one in which they were the principal actors and another in which they played no part.

It was hypothesised that the students lying about their own experiences or telling the truth about their lack of involvement in somebody else's embarrassing incident could be classified correctly, on the basis of their differential physiological responses. The interrogator, blind to the knowledge of who was the source of the reported embarrassing stories found that subjects could be correctly classified as deceptive or non-deceptive.

Through this study Bradley and Cullen (1993) attempted to link a "real" event from the field, that had personal relevance to the subject, as well as emotional content associated to that personally relevant field event, and to study it in a controlled laboratory study. The finding that subjects could be correctly classified based on the emotion of embarrassment indicated that emotions likely do influence the accuracy rates of detection of deception.

Bradley, Cullen, and Carle (1993) assessed the relative levels of emotionality in embarrassing stories and compared this assessment to the levels of emotionality found in regular laboratory studies. Generally negative emotions, such as embarrassment and anxiety, were associated more strongly with the real "embarrassing" experiences than to laboratory mock crimes. In spite of the prevalence of more negative emotions associated with real events, Bradley, Cullen, and Carle (1996), in a follow up study, failed to find a difference in detection rates between real events and laboratory mock crimes. Interestingly, the emotion differences were substantial, not only in degree but also in kind. Mock crimes are interesting, exciting and fun whereas the real events were negative and upsetting.

Carle (1996) pursued this research in a test of conditioned response theory by having introductory psychology students generate pleasant, unpleasant, and emotionally neutral real life stories, and subsequently conducted polygraph examinations on some of the students. Positive and negative events could have strongly conditioned components associated with them whereas neutral events, by definition, should not be emotionally evocative. There was no support for differences found in detection accuracy across the three types of stories.

In summary, it is believed that polygraph techniques, such as the CQT, may be in part based on emotions elicited by questions surrounding a particular event. Research, however, has shown at best only mixed results on the various theories to account for the emotional influence on the physiological detection of deception.

B) Cognitions

The present discussion has been concerned mainly with the control question test in the detection of deception. The guilty knowledge test, however, has a role as a source of control questions for the intended modification of CQT to be empirically explored. Therefore, this discussion will include not only the CQT but also the GKT. As was pointed out, cognitive and emotional factors play a role in each but the emphasis, at least from the theories, is somewhat different. The CQT is arguably considered an emotion-based test with attention and memory factors enhancing responding. The GKT is considered a cognitive test with emotional factors perhaps amplifying responses.

Cognitive theories in detection focus on attention, categorisation, recognition, novelty, knowledge, mental effort, decision making processes, memory and orienting responses. Some of these cognitive processes may be involved when a suspect is answering interrogation questions and all may be accompanied by physiological responding (Beatty and Khaneman, 1966; Ben-Shakhar, 1977; Khaneman and Beatty, 1966; Lieblich, Kugelmass and Ben-Shakhar, 1970; Lykken, 1974; Simpson and Hale, 1969; Waid, Orne, Cook and Orne, 1978; Waid, Orne and Orne, 1981). Subjects recognise stimuli that are significant and this recognition, in turn, triggers physiological responding. If a person lies, the knowledge and effort involved with deception will be related to physiological

responding. It is believed that, under controlled conditions, polygraph examiners can reliably recognise these short-lived sympathetic responses that are related to cognitive factors in deception.

According to Ben-Shakhar and Furedy (1990), theories should include a cognitive component. A lack of consideration for cognition makes it difficult to explain why detection may occur under even mild conditions.

Mental Effort

Successful detection occurs even when subjects are explicitly motivated to not be deceptive (Elaad and Ben-Shakhar, 1989; Horvath, 1978). Another situation where detection can occur is when subjects do not even attempt to conceal relevant information (Janisse and Bradley, 1980). Detection can even occur when subjects are unaware that they are being "monitored" (Thackray and Orne, 1968a).

Attention and Memory

One cognitive process is attention. Waid et al. (1978) studied attention by having subjects learn a list of code words. In a GKT examination, they were presented with both the code and control words. After the examination, they were asked to recall the control words. Memory for control words could not occur if subjects did not attend to the test. There was a positive correlation between responses to code words and the number of recalled control words following the GKT. Waid et al. (1981) extended the 1978 study and included the CQT as well as the GKT procedure and found that ... "Both relevant and

control questions that were later recalled produced larger SCR amplitudes than non-recalled questions" (Ben-Shakhar and Furedy, 1990, pg. 109). Waid et al. (1981) believed that if a subject was actively attentive to an item, he or she would be more likely to recall it later. Therefore, as Ben-Shakhar and Furedy (1990, p. 110) explained ..."the individual differences in the physiological responsivity to the different questions are explained in terms of the amount of attention paid to them".

Arousal affects attention. Easterbrook (1959) suggested that increases in arousal result in a narrowing focus of attention. No one has systematically examined laboratory or field situations to discover when a narrowing of attention might be optimal or sub-optimal for accurate detection. Speculatively, it is possible to imagine a suspect so concerned with issues relevant to the crime that he or she genuinely fails to remember past life events probed for in control questions. If such a person were innocent, they might appear guilty because a control question, with such forgetting, could evoke only small responses.

Recognition, Memory, and Orientation Response

Memory and general cognitive process problems could differentially affect GKT accuracy. Lykken (1974) has assumed that guilty "suspects" have guilty knowledge and remember specific aspects of a crime. The retention and recognition of that knowledge creates enhanced physiological responsivity to the guilty items. His rationale lies with the physiological process known as the "Orientating Response", or "OR" (Sokolov, 1963).

Lykken (1974) reasoned that the guilty knowledge had an added "signal value", and thus would produce "stronger" ORs that would differentiate the guilty item from the other

choices. During a polygraph examination, any item presented to the subject would evoke an OR. Repeated presentations of the same or similar stimuli will cause habituation, or a decline in the magnitude of the physiological response.

In a similar manner to Lykken's guilty knowledge theory (1974), Lieblich, Kugelmass and Ben-Shakhar (1970) and Ben-Shakhar (1977) developed the dichotomization approach, based on the mechanisms of orientation and habituation. As in Lykken's approach, the information for interrogation purposes is based on sets of relevant and irrelevant stimuli. Unlike Lykken's approach, it is not the signal value of the relevant stimulus that results in large responses. It is the relatively rare frequency of occurrence of relevant stimuli that results in greater physiological responding. For example, if eight questions were related to irrelevant amounts of money, and the ninth was the amount from the crime, a larger response should occur to that item than if only three items were irrelevant and one was relevant. Habituation should occur at different rates for different sized sets of relevant and irrelevant information.

Ben-Shakhar (1977) found that if the irrelevant stimuli were rarely presented (compared to frequent "relevant" information presentations), the subjects would have greater responsivity to the irrelevant material.

An interplay between cognition and emotion could be evidenced with OR's and defensive responses, or "DR's". The OR represents a complex range of physiological reactions evoked by any change in stimuli (Stern, Ray, and Davis, 1980). There are increases in the sensitivity of sense organs, movement towards the stimulus, and decreases in heart rate. Of particular relevance to the field of lie detection, habituation to the OR is

rapid. The rapid habituation is useful in detection studies, especially if it is reflective of the resolution of an appraisal or examination process. For example, control questions could be considered as "non-threatening".

The OR differs from the "defensive response", or "DR", which protects individuals from dangers of intense stimulation (Stern, Ray, and Davis, 1980). In essence, the "DR" acts in an opposite manner to the "OR". There is a decrease in the sensitivity of sense organs, there is a movement away from the stimulus, and there are increases in heart rate. Although habituation to the OR is rapid, habituation to the DR it is very slow. This makes the DR relevant to the field of lie detection. The DR results in continued responding to key stimuli.

Knowledge

Knowledge alone is not enough to differentiate between relevant and irrelevant stimuli. Giesen and Rollison, (1980) and Stern, Breen, Watanabe and Perry, (1981) provided "relevant" information to innocent subjects. Under polygraph examination, the authors could still discriminate between guilty and innocent but knowledgeable subjects because those innocent subjects were not differentially responsive to key items.

Bradley and Warfield (1984) supplied members in three of four groups with guilty knowledge. They were either guilty of the crime, witnessed it, or were told crime relevant details. Members of the fourth group were innocent, and possessed no information concerning the crime. Results indicated significant differences in detection between groups with the guilty group more detectable than the informed groups.

Memory

Waid et al. (1978, 1981) studied the role of memory in lie detection. Guilty Knowledge Tests are essentially recognition memory tests. They found enhanced electrodermal responding to relevant words that subjects were subsequently able to recall. They argued that the more actively a subject attends to a question, presumably because of recognition memory, the larger a response it evokes.

In a level of information processing study, Craig and Lockhart (1972) have shown that memory depends on how information was processed or attended to at the time of learning. Ben-Shakhar and Furedy (1990) have indicated that even under mild conditions, such as a polygraph situation where no motivational instructions to beat the test and no verbal response is required to the questions, some physiological differentiation may still be expected as long as the participant recognises some of the relevant information. This may be important if one considers the goal of the criminal, and how that may differ from the goal of an investigator. For example, to what does a criminal attend during a crime, and does this contrast with what an investigator believes is relevant information?

Given the importance of memory in GKT tests it is perhaps no surprise that Elaad, Ginton and Jungman, (1992) found that information tests that worked so well in the laboratory yielded poor results when used in actual criminal investigations. Many GKT studies may not present very realistic conditions. Typically in laboratory studies students (not criminals) learn (not live) their role under relatively safe, benign conditions (M.T. Bradley, personal communication, August, 1999). Attention, regard for detail, and memory storage could be very different in the field. Speculatively, findings such as the narrowing of

attention reported by Easterbrook (1959) could mean that high levels of arousal associated with criminal activities and subsequent interrogations precludes efficient attention, and memory processing of the stimuli involved in the crime. An investigator in a methodical study of a crime scene may judge certain stimuli to be important for an investigation, but the criminal may not have noticed or remembered a variety of details.

There is more potential for state-dependent effects in the field than in laboratory studies. Crimes may involve drug and alcohol abuse and positive and negative emotions at the extreme. Measurement of emotions in the laboratory generally show that subjects are excited and interested in both the crime and the following interrogation. Speculation is that the predominant emotion for a field interrogation would be depression and possibly anger. These emotions may not match the emotions of a crime, which could range from elation to horror. There is a continuity of emotions in the laboratory setting compared to the potential extremes of discontinuity of emotional and sobriety states in the field. In the field there may be large time delays before testing and important, simultaneous events in the subject's life that occur. These may be potential reasons for the differing results between the laboratory and the field.

To illustrate the above point, consider the situation of a theft. In the midst of a purse snatching, the thief may not have noticed the colour of the dress of his victim, her height, the contents of her purse, or even where he "dumped" the evidence. In fact, if he had stolen 2 or more purses that day and "stuffed" the money into his pockets, he may not even know how much money he stole from the first victim, the second, and so on. Therefore, if the guilty subject did not feel that certain information was relevant, he may have not retained it,

or it may be inaccessible with the cues supplied during the course of an interrogation. Bradley, Carle, and Nagaragi (1998) reported that the majority of participants who were allowed to choose a weapon to strike a victim of a mock crime did not remember the shirt colour of the victim whereas those who were specifically instructed to use a particular weapon did remember the shirt colour. This difference occurred even though neither group was instructed about the colour.

In summarising the cognitive portion of this review of the theories of lie detection, there is evidence that mental processes are very much involved in the detection of deception. The better the memory for items, the better the rate of detection (Waid et al. 1978, 1981). Memory depends upon the type of processing (Craig and Lockhart, 1972) and the attentional resources employed during both the crime and interrogation (Waid et.al., 1978; Ginten and Jungman, 1992; Bradley, Carle, and Nagaragi, 1998). Once material is categorised as relevant or irrelevant the detectability is affected, not only by the signal value of relevant stimuli (Lykken, 1974), but also the relative proportion of irrelevant stimuli included in an interrogation (Lieblich, Kugelmass, and Ben-Shakhar, 1970; Ben-Shakhar, 1977). Although research has shown mixed results concerning the emotional influence on the physiological detection of deception, it is still believed by many that polygraph techniques, such as the CQT are, in part, based on emotions elicited by questioning. The following table (Table 1) summarizes the general body of data examinining this issue and it is clear that both cognitive and emotional factors must be considered in conceptualizing a theory of lie detection.

Table 1

Summary of Research Concerning the Theories Associated with Emotions and Cognitions

Theories associated with emotions

1) Physiological/behavioural Patterns linked with specific emotions

Study (Year)	Evidence or Significant Finding
Levenson, Ekman, and Friesen (1990)	Yes
Ortony and Turner (1990)	Yes
Cannon (1927)	General Arousal
Lykken (1981)	General Arousal
Heslegrave (1981)	Inhibitory Arousal

2) Punishment theory (Davis, 1961)

Study (Year)	Evidence or Significant Finding
Kugelmass and Lieblich (1966)	No
Bradley and Janisse (1981)	No

3) Conflict approach

Study (Year)	Evidence or Significant Finding
Forman and McCauley (1986)	Yes
Kugelmass, Lieblich, and Bergman	Yes
Elaad and Ben-Shakhar (1989)	No
Furedy and Ben-Shakhar (1991)	No
Horneman and O'Gorman (1985)	No

4) Conditioned Response

Study (Year)	Evidence or Significant Finding
Bradley and Cullen (1992;93)	Yes
Bradley, Cullen, and Carle (1993)	Yes
Bradley, Cullen, and Carle (1996)	No
Carle (1996)	No

Theories associated with cognitions

1) Mental effort

Study (Year)	Evidence or Significant Finding
Elaad and Ben-Shakhar (1989)	Detection (not motivated to deceive)
Horvath (1978)	Detection (not motivated to deceive)
Janisse and Bradley	No attempt to conceal relevant info
Thackray and Orne (1968)	Detected even when unaware of being monitored

(continued on next page)

Table 1 (continued)

2) Attention and memory

Study (Year)	Evidence or Significant Finding
Waid, Orne, Cook, and Orne (1978)	No memory without attention
Waid, Orne, and Orne (1981)	Recalled words had better Physiological responses
Esterbrook (1959)	Increase in arousal means a narrowing of focus

3) Recognition and novelty

Study (Year)	Evidence or Significant Finding
Lykken (1974)	Guilty knowledge theory- recognition of relevant knowledge creates enhanced responsivity to guilty items creating greater signal values and thus stronger OR's

4) Orientation

Study (Year)	Evidence or Significant Finding
Sokolov (1963)	Orientation Responses (OR's) and habituation
Lieblich, Kugelmass, and Ben-Shakar (1970)	Dichotomization approach frequency of occurance of relevant stimuli that results in greater physiological responding. Habituation should occur at different rates for different sized sets of relevant and irrelevant info.
Ben-Shakhar (1977) Stern, Ray, and Davis (1980)	Found supporting evidence. Habituation rates between DR's and OR's are different.

5) Knowledge

Study (Year)	Evidence or Significant Finding
Giessen and Rollison (1980)	Could still discriminate between guilty and innocent but knowledgeable subjects
Stern, Breen, Watanabe, and Perry (1981)	Same findings as above
Bradley and Warfield (1984)	Significant differences in detection rates between groups with the guilty group the most detectable

(continued on next page)

Table 1 (continued)

6) Memory

Study (Year)	Evidence or Significant Finding
Waid et al. (1978; 1981)	Enhanced electrodermal responding to relevant words that subjects were subsequently able to recall. The more actively a subject attends to a question (due to recognition memory) the larger the response it evokes.
Craig and Lockhart (1972)	Level of information processing study. Memory depends on how info was processed or alluded to at the time of learning.
Ben-Shakhar and Furedy (1990)	Differentiation may still be expected as long as the participant recognises some of the relevant info.
Elaad, Ginton, and Jungman (1992)	GKT had poor results when used in actual criminal investigations.
Bradley, Carle, and Nagaragi (1998)	When subjects are given a choice, particular relevant info was not remembered but those not given a choice remembered the relevant info.

The results of the entire review (table 1) indicate that a desirable interrogation procedure should contain elements that rely on both cognition and emotion. The next section discusses the development of a procedure, which draws from both the CQT and the GKT in order to tap both the cognitive and emotive domains.

The Development of a Control Question Test With Actual Truth Control Questions

The present study will focus on control questions in the CQT. As discussed earlier, a CQT pairs a "crime relevant" question with a "control" question (designed to evoke an unsettling emotion in the innocent subject). Following the polygraph session, the physiological responses for each pairing are compared, and the question from each pairing

that produces the larger response leads to the judgement of guilt or innocence by the examiner.

Typical "control" questions, as mentioned, have several qualities. These include ambiguity, lack of specificity, and the possibility that a suspect could be deceptive. Bradley, MacLaren, and Black (1996) eliminated these control question qualities. They created unambiguous, specific questions to which no participant would lie. Three groups of subjects were created: a group whose members were guilty of a mock crime; a group whose members were innocent but informed of the details of the crime; and a group whose members were innocent and uninformed of any relevant details.

The "innocent and uninformed" group had no basis for distinguishing between the control questions and the crime relevant questions; the "innocent yet informed" group knew all of the crime related details; the "guilty" group was required to be deceptive to the crime relevant questions. Examples of questions for the examination of a mock crime theft of 20 dollars follow: the crime relevant question was ..."Did you steal the 20 dollars?"; and the control question was ..."Did you steal the 15 dollars?". In this manner ambiguity was eliminated, specificity increased and the potential for lying avoided, and the paired questions were identical except for the key detail associated with the crime.

In almost every manner the "pairings" of questions were similar except that the innocent subjects were not lying to either question whereas the guilty individuals were lying in response to the crime relevant question. The results of the Bradley et al. study (1996) provided support for the modified questioning technique. Guilty subjects scored as

deceptive, whereas the innocent subjects scored as truthful (even when they knew details of the crime).

Bradley, MacLaren, and Black (1996) presented their test in a set order with the control question always prior to the question about the crime. This means that subjects may not have responded to the specific content of the question, but rather had an orienting response to a new topic.

Despite researchers such as Bradley et al. (1996), Furedy, Davis, and Gurevich (1988), and Furedy, Posner, and Vincent (1991), mentioning potential order effects, there is a lack of published research specifically examining this issue. Guilty Knowledge Test (GKT) studies never involve the guilty information in the first position of a series of items and they never include the first item in scoring the test. Researchers, such as Lykken, assume that the orienting response to the first item will be stronger than to any other item regardless of its relevance to a crime. Not all those who utilise the CQT technique, e.g. Elaad and Elaad (1994), ensure that control questions are always ahead of the crime relevant questions.

Ben-Shakhar and Lieblich (1982) manipulated the serial position of a relevant item in a GKT paradigm. They found that it was advantageous to present the relevant item at an early point in the list. This result, however, was not replicated by Ben-Shakhar, Asher, Poznansky-Levy, Asherwitz, and Lieblich (1989), who found similar detection rates for early as well as late presentation of the relevant item within a series.

Bradley, MacLaren, and Black (1996), assuming an order effect, suggested that placing the control question in the first position would provide protection for innocent

subjects. That is, if the first question of each pairing is a control question, the responding due to an OR will be enhanced. Since innocent subjects are truthful to both the control and crime relevant questions, and even though OR's habituate over subsequent presentations, the response to the initial question should remain larger because of its position. Guilty subjects, however, will produce a large response to the crime relevant item, even if it is in the second position. That response will be a combination of guilt, lying and possession of guilty knowledge.

The changes by Bradley et al. (1996) to a single GKT style of control question creates a variation on the CQT. This variation, which could be called the "actual truth control question test" (ATCQT), at face value resembles several other techniques. The apparent similarity is not fundamental. Bradley et al. (1996) analysed how the test differs from other tests and will be discussed in the next section.

Testing Techniques Related to the Control Question Test With Actual Truth Control Questions

The "Relevant-Irrelevant Test" (RIT), as described by Lykken (1981) seems similar to the ATCQT. The test compares subjects' responses evoked by crime "relevant" questions to neutral ("Is your shirt red?") or somewhat provocative ("Have you been drunk in the past year?") questions. As Bradley, MacLaren, and Black (1996) pointed out in the case of the RIT, however, there are no criteria to select the irrelevant "control" questions. The question pair could contain a crime relevant question and a control question that is anything from neutral to provocative and is, or is not, plausibly related to the crime. The ATCQT, on the

other hand, contains control questions that are plausibly related to the crime in the sense that a specific action, amount, person or item could have been involved in the crime under investigation but, in fact, was not.

The ATCQT is also somewhat similar to the Truth Control Question Test, or (TCT) (Lykken, 1981). Here the suspect is accused of both a real crime as well as a fictitious one that is similar in nature. It is assumed that innocent suspects would have large reactions to the fictitious crime, and judged to be nervous over being accused in general rather than focusing on the relevant crime questions. Bradley et al. (1996) contend that, although theoretically, the TCT has potentially true, scientific control questions, in reality the technique is complex and cumbersome to carry out. For example, time must be devoted to investigating the comings and goings of the subject over the last while to ensure that the time of the fictitious crime should not allow the suspect to have a ready alibi. Careful planning must go into both developing and delivering a plausible explanation to the subject as to why (especially for a serious crime) the crime received no media attention.

In contrast to the Truth Control Question Test, the ATCQT is relatively simple and easy to construct. There is no pretence of another crime. There are simply pairs of questions on which innocent suspects will be equally truthful and on which guilty suspects will be lying to one and truthful to the other. Interestingly, if innocent suspects were uninformed of the details of the crime, the ATCQT may, in fact, act as a TCT in that, from the suspect's perspective, either or both questions may concern a real crime.

The ATCQT is somewhat similar to the Guilty Knowledge Test (GKT) developed by Lykken (1959). Simply stated, guilty suspects possess guilty information while the

innocent suspects do not. Aside from the obvious difference that deception examinations with the ATCQT involve a choice between only 2 items (like the traditional CQT) rather than choosing between 5 items (as in the GKT), a more fundamental difference occurs. Bradley et al. (1996) indicate that the first item of each GKT question is not scored because the orienting response may be large enough to mask the response to the critical crime question. The ATCQT takes advantage of using the control question in the first position for protection purposes. That is, Bradley et al. (1996) found both suspects who are innocent and uninformed as well as those who are innocent and informed were judged innocent. The OR was large enough to offset any response to the crime relevant response, even when the information is known. The researchers found, however, that the orienting responses of the guilty suspects to the control item were smaller than the responses to the subsequent crime relevant questions on which they lied. The fundamental difference between the two tests, therefore, is that the GKT avoids the orienting response whereas the ATCQT incorporates it into the test to provide some measure of protection for the innocent suspects. Table 2 summarises some of these points.

Table 2

Comparisons of Different Questioning Techniques

CQT	GKT	TCT	ATCQT
1) Question type a) Control emotional control	4 to 5 similar foils	1 false crime foil	1 similar foil
b) Crime Relevant 1 crime relevant	1 crime fact	1 crime fact	1 crime fact
2) Order important	unimportant after first question	unimportant	important
3) Concerning OR untested	unimportant after first question	untested	important

Hypothesis

It was hypothesized that accurate classification of participants based on physiological responses would depend jointly upon their actual guilt condition and the serial position of the crime relevant items. Magnitudes of physiological responses depend on the position of items in a sequence and deception. Greater responding occurs to questions that are early in a sequence and to those that are answered with a lie. Therefore if the crime relevant question is first in a control crime relevant question pair, guilty participants should be classed accurately whereas innocent suspects will be classed inaccurately. When the crime relevant question is first, guilty suspects have both factors, serial position and deception, evoking relatively large responses to initial questions. Innocent participants have only the position effects since they are not deceptive but this

should still be enough to result in their inaccurate classifications. When crime relevant questions are second, guilty participants will be accurately classed as guilty and innocent participants will be accurately classed as innocent. That is, acts of deception will enhance responses enough such that they will be larger than the reactions to initial questions. Because innocent participants are not deceptive to the crime relevant questions, their reaction to initial control questions will be larger when compared to the crime relevant material.

Chi square analysis should show support for the following results. There should be 1) a condition effect showing more participants classed as guilty in the guilty condition than in the innocent conditions; 2) an order effect should also be evident with more participants classed as guilty when the crime relevant question is in the first position; and 3) an interaction should be found in a guilt condition by question position analysis showing proportionately more participants classed correctly in their respective condition when crime relevant questions are in the second position than when they are in the first position.

Secondary Analysis and Predictions

Secondary analyses involved examining the derived scores, upon which the Chi Square analyses were based, to understand which physiological measures provided underlying support for the Chi Square results. It was hypothesised using a MANOVA involving all measures, a condition, order, and a condition by order interaction effect would be found.

A condition effect would result in more guilty subjects scoring as more guilty than the innocent subjects. An order effect would be evidenced if crime relevant items in the first position would result in subjects scoring as more guilty than subjects who received the information in the second position. Support for the predicted interaction would be found if scores for participants, regardless of condition, would be towards guilt when the crime relevant question was in the first position. Additionally, when the crime relevant question was in the second position, only the guilty participants should score in the guilty direction.

A third analysis involves an examination of the raw scores. Whatever support is found for the Chi Square hypothesis, in terms of a condition, order, or condition by order interaction effect, should also be found in this analysis. In addition, it provides an opportunity to explore habituation.

The habituation process should be reflected in large responses occurring early and diminishing later. Orienting responses are particularly large responses to new stimuli, and these should diminish over repetitions of the same stimuli. In addition, there should be a new topic OR such that when a topic is introduced, the response is large to the initial question and smaller to subsequent questions on that topic. Analysis of the raw scores (actual physiological measures) in the form of ANOVA's was utilised to test this assumption.

The last analysis involves memory. The prediction was that the better the memory for the crime relevant items (involving both the innocent and informed and the guilty participants), the more detectable the participant would be on the polygraph examination. A

MANOVA and subsequent ANOVA's over conditions, order and gender involving recall and recognition scores were used to test this assumption.

The current examination was conducted in the laboratory. The major reasons for this decision were: (1) there are many difficulties and factors to control in the field situation; (2) that it is difficult to convince polygraph operators to introduce a new, modified, and untried technique in the field; (3) traditionally, either introducing a new test (or modifying an old technique) should be undertaken in the most controlled conditions possible (i.e. a controlled laboratory situation); (4) new methods may not work or cause some degree of harm and therefore a controlled environment is needed to monitor whether the guilty and innocent suspects are being properly classified.

Although the majority of "testing" for deception in the field is done with males, there was no reason to exclude females in this study. Ben-Shakhar and Furedy (1990) reported no gender differences. The basic nature of OR's suggests that responses should have worked equally well with both genders.

Chapter 2

Method

Participants

One hundred and twenty Introductory Psychology student volunteers, 60 male and 60 female, took part in the study. Participants received a bonus point to add to their Psychology grade.

Apparatus

A Lafayette 750-566 field polygraph was used to record skin resistance responses (SRR), blood volume (BV), thoracic respiration responses (THR), and abdominal responses (AB). Skin resistance responses were measured by Zinc-zinc chloride electrodes attached to the medial phalanges of the first and third fingers of the participant's right hand. Respiration was measured by two pneumatic tubes positioned around the thoracic area and the abdomen. Cardiovascular activity (a combination of heart rate and blood volume) was measured with a photoplethysmograph meter attached to the participant's second finger on the right hand during the interrogation.

Interrogators

Two graduate students (an M.A. candidate and the author) alternated between serving as the lab assistant and the interrogator. Both of the graduate students had been trained by the same university professor (Dr. M. Bradley, who had been researching in this area for over two decades). Further, both students had ran subjects for at least three different published articles.

Procedure

A consent form was given to all students who were interested in participating in the study (appendix A). It contained information describing the study, as well as information on the potential risks and benefits of participation. An important feature of the consent form was the clear indication that participants could withdraw from the study at any point without penalty.

Individual participants were asked to go to the experimental area at an agreed upon time and report to a laboratory assistant. That assistant randomly assigned a "file" package to the subject. It contained an assignment to one of three conditions (guilty, innocent, or innocent and informed). The "file" package also determined the question order in which the examiner conducted the interrogation (order 1 presented the crime relevant question first among all pairings of questions and order 2 presented the control question first among the pairings). Lastly, the package contained the actual questions that would be asked during the polygraph interrogation. The last two aforementioned items, the order presentation and the actual interrogation questions, appeared on a worksheet (appendix B).

If participants were in the "guilty" condition, they read and carried out a set of instructions (appendix A) requiring that he or she had to go to a specific Professors' office and: a) enter without knocking; b) remove twenty dollars out of a wallet located in a sportcoat hanging over a chair; c) stash the stolen money in their footwear (left foot); d) place the wallet back into the jacket and; e) report back to the laboratory assistant.

If the participant was in the "innocent" uninformed condition (appendix A), the laboratory assistant asked him or her to go into the hall and read the instructions, which

contained no information concerning the crime, and report back to the laboratory assistant. The "innocent and informed" participants read material describing the crime with the same information given to the guilty subjects, but were instructed not to do the crime (see appendix A). They waited and then reported back to the laboratory assistant. Once they had returned to the assistant, regardless of their condition of guilt or innocence, the subjects were instructed to act as if they were innocent by co-operating with the interrogator but to deny any questions relevant to the theft. It was stressed that the polygraph interrogator was unaware ("blind") of their actual condition so that judgement rested solely on their performance on the polygraph test. This was actually the case inasmuch as only the participant actually knew what condition he or she was in prior to and during the interview. The assistant and the interrogator were only aware of which order the questions should be administered during the examination. The pre-test interview was standardized, with the lab assistant reading the questions to the participant and indicating what the answers should be, and lasted approximately four minutes. There was neither a discussion concerning the accuracy of the polygraph nor any type of accuracy demonstration.

In the test room, the polygraph interrogator went over the questions and briefly showed the instrumentation to the participant. Once the physiological measuring instruments were attached, one of a possible two polygraph examinations, order one or order two, was administered with three repetitions (appendix B). One polygraph examination was referred to as "order 1" because each of the three questions with information relevant to the crime was asked first in each question pair. In this study 20 dollars was stolen by guilty participants so..."Did you steal 20 dollars?" was followed by

..."Did you steal 30 dollars?". The other possible polygraph examination was referred to as "order 2" because each of the three crime relevant questions followed the paired control question (that is, the question of theft of the 20 dollars followed the question of theft of 30 dollars). The interrogations used in this study were the same length as is commonly used in the field investigations.

After the examination the assistant met with the subjects and had them perform both recall and recognition tests concerning the theft (appendix C). They received 50 cents per item recalled or recognised. They were assured that the interrogator would have no knowledge of any participant's condition prior to marking the polygraph examination. They were also told that, following the entire data collection, copies of the educational component of the study could be obtained from the main Psychology Office (appendix D).

Data Analysis

Respiration scores, thoracic (THR) and abdominal (AB), were measured with the use of an Alvin 1112 contour map wheel; SRR amplitudes and BV variations were assessed with the use of a ruler. The scores for each physiological measure were submitted to different levels of statistical analyses.

Chi Square Analyses

In the first analyses, classification data was examined through Chi Square calculations. In practical instances of lie detection, the interest is in the optimal accuracy of the test in discriminating categorically between guilty and innocent subjects. The

classification process discriminates between the categories. Selected chi-square analyses illustrate findings at this level.

Summing the derived scores by a process described in a later section, from each of the 4 measures (SRR, THR, AB, and BV) to obtain a total made a determination of "guilt" or "innocence" for that measure. Those scores for each measure could range from -9 to +9. Participants with scores of +2 or above were judged as innocent and those with a score of -2 or less were judged as guilty. Those scores equal to or between -1 and +1 were considered inconclusive.

An overall score was calculated by summing over the four measures. Scores here could range from -36 to +36. Scores between +2 to +36 inclusive were judged as innocent; scores between -2 and -36 were judged guilty; scores falling between -2 and +2 were judged inconclusive.

Secondary Analyses

Derived Score Analysis

The level of data analysis involved "derived" scores based on the comparative differences between physiological responses to crime relevant and control questions. Much of the literature provides analysis at this level. This level of analysis simplifies results by assigning only one score per physiological measure for each participant. Therefore, potential interactions involving within-subject factors are eliminated.

The first step in obtaining a derived score is to take the raw score (previously explained for each measure) and then examine relevant pairs of questions. For each pair of questions (crime relevant and control), a score was assigned. If the amplitude of a skin

resistance response (in millimetres), within 10 seconds after the start of a question, was larger following a relevant question, a -1 was awarded; if the appropriate control question was larger, a +1 was given. If both scores were equivalent, a zero was awarded. The linear distance of a respiration response was measured for the same 10-second period. If the linear distance was smaller on a crime relevant question as compared to a control question, it received a value of -1; if the response on the crime relevant question was larger, a value of +1 was awarded. If both questions were precisely the same in response magnitude, a value of zero was awarded.

A value for the cardiovascular activity (BV) was obtained by measuring a 10-second block of time starting with the question presentation. Measures of the vertical distance within that timeframe were in millimetres beginning at the "elbow", or low point, of the first negative slope of blood volume until the end of an ascending recovery point. If an ascent took place first, the highest point was measured against the lowest point within the timeframe. If the measured response to a relevant question was larger than the response to the control question, a value of -1 was assigned; if the control question's response was greater, a +1 was awarded.

These four derived scores were analysed with a MANOVA. This approach combines the measures. Individual ANOVA's would be used to follow up on significant reaults.

Raw Score Analysis

Raw scores computed from each physiological measure were analyzed directly. Theoretically this is the most powerful analysis to look for differences in detail amongst all factors in this experiment and pinpoint exactly where they happened.

The largest skin resistance response amplitude, in millimetres, (SRR) was measured for a 10-second period starting with the beginning of a question presentation was recorded.

The respiration scores were measured with the use of a contour map wheel. Starting with the beginning of question presentation, the linear distance of each tracing was measured in mm for a 10-second period.

A 6 factor analysis of variance was performed for each of the skin resistance response (SRR), thoracic (THR), and abdominal (AB) measures. The results of these analyses are presented in appendices J, L, and N respectively. Three factors were based on between subject factors. They were gender (male or female), three types of conditions (guilt, innocent, and innocent and informed), and order (two types of order of presentation questions). Three factors were based on within subject factors. They were 3 blocks (groupings of 10 item test repetitions), 3 pairs (of questions per test), and 2 question positions (actual crime information in the first or second position). Some data sets were incomplete. This was usually due to pen ink clotting or jamming (no ink comming out of the pen) for a few values. In these cases, an average was taken of the other available scores for that participants particular physiological measure and was substituted for the missing value.

Additional Analyses

A MANOVA was conducted with the recall and recognition scores as dependent variables and Gender, Condition, and Order between-subject factors as independent variables.

Chapter 3

Results

Statistical Analysis on Classification Data

Numeric classifications of Guilt, Innocence, and Inconclusive on which the analyses were based for the totals of each order and for each measure are shown in table 3.

Table 3

Numeric Classification of Guilt, Innocence, and Inconclusive Judgements

Classification Accuracy (Chi Square Analysis)

		Actual Guilty			Innocent Informed			Innocent		
Measure	Order	G	I	Inc.	G	I	Inc.	G	I	Inc.
Total	1	16	2	2	14	1	5	13	4	3
Total	2	11	4	5	2	15	3	2	12	6
Ab	1	8	5	7	9	5	6	6	6	8
Ab	2	7	3	10	3	8	9	4	8	8
Thor	1	4	3	13	7	5	8	8	2	10
Thor	2	7	6	7	5	4	11	4	0	16
BV	1	14	2	4	7	2	11	6	1	13
BV	2	8	6	6	2	14	4	2	10	8
SRR	1	16	1	3	12	2	6	10	3	7
SRR	2	8	3	9	1	12	7	2	11	7

Note. G =Guilty, I = Innocent, Inc. = Inconclusive

Chi Square Analyses for a condition effect showing more participants classed as guilty in the guilty condition than in the innocent conditions were significant for Total Scores (χ^2= 4.6, df=2, p= 0.00) and for BV (χ^2=9.4, df=2, p=0.00) and SRR (χ^2=5.4, df=2, p=0.00) measures. Refer to table 4.

Table 4

Subjects Judged as Guilty in Each Condition

	Guilty	Innocent Informed	Innocent	Totals
Total Scores				
Observed	27	16	15	58
Expected	19.3	19.3	19.3	58
BV Scores				
Observed	22	9	8	39
Expected	13.	13	13	39
SRR Scores				
Observed	24	13	12	49
Expected	16.3	16.3	16.3	49

Chi Square Analyses to examine for an order effect to determine if more participants would be classed as guilty when the crime relevant question was in the first position were significant for Total Scores (χ^2= 13.5, p= 0.00), BV (χ^2=5.8, df=1, p=0.00) and SRR (χ^2= 14.9, df=1, p=0.00) measures. Refer to table 5.

Table 5

Guilt Findings

	Order 1	Order 2	Total
Total Scores			
Observed	43	15	58
Expected	29	29	58
BV Scores			
Observed	27	12	39
Expected	19.5	19.5	39
SRR Scores			
Observed	38	11	49
Expected	24.5	24.5	49

Chi Square Analyses on correct judgements conducted to examine for a hypothesised interaction concerning a guilt condition by question position were significant for Total Scores (χ^2= 13.50, df=2, p= 0.00) and SRR (χ^2=13.08, df=2, p=0.00) and BV (χ^2=14.9, df=2, p=0.00) measures. Refer to table 6.

Table 6

Overall Scores of Participants Judged Correctly

Total Scores	Correct Guilty	Correct Innocent-Inf	Correct Innocent	Totals
Observed Order 1	16.0	1.0	4.0	21
Expected Order 1	9.6	5.7	5.7	21
Observed Order 2	11.0	15.0	12.0	38
Expected Order 2	17.4	10.3	10.3	38
BV Measure	**Correct Guilty**	**Correct Innocent-Inf**	**Correct Innocent**	**Totals**
Observed Order 1	14.0	2.0	1.0	17
Expected Order 1	7.6	5.6	3.8	17
Observed Order 2	8.0	14.0	10.0	32
Expected Order 2	14.4	10.4	7.2	32
SRR Measure	**Correct Guilty**	**Correct Innocent-Inf**	**Correct Innocent**	**Totals**
Observed Order 1	16.0	2.0	3.0	21
Expected Order 1	9.7	5.7	5.7	21
Observed Order 2	8.0	12.0	11.0	31
Expected Order 2	14.3	8.3	8.3	31

Appendix E shows the accuracy in percentage form of judgements made for each order and all of the measures, and includes Chi Squares showing greater classification accuracy in Order two for all measures but Thoracic respiration.

Statistical Analysis on Derived Scores

In this analysis "derived" scores based on the comparative differences between physiological responses to crime relevant and control questions were examined. As described in the data analysis section, each of the nine pairs of critical and control questions

for each measure were assigned a plus or minus one. These values were then summed to give an overall score per subject for each measure.

Multivariate Analyses of Variance were conducted using conditions (3 levels) by Gender (2 levels) by Order (2 levels) as between subject factors and the derived scores from skin resistance response, blood volume, abdominal and thoracic respiration scores as the dependent variables (see appendix F).

A multivariate Condition effect was found (F (8, 212) =4.2). Univariate analyses on the separate scores found that the multivariate Condition effect occurred because of differences in SRR's, (F(2, 108) =8.2) and BV (F(2, 108) =6.1). The relevant results are presented in table 7.

Table 7

Means of SRR and BV Across Conditions

	Condition		
	Innocent	Innocent-Inf	Guilty
SRR	.25	-.48	-2.48
BV	.27	.73	-1.15

The univariate F tests for these analyses are presented in appendix G. No differences were found for responses from the respiration scores.

A Duncan's Multiple Range Test on SRR means found that innocent-uninformed participants (M =0.25) score more towards innocence than innocent yet informed participants (M =-0.48) and guilty participants (M =-2.48). Also the innocent participants

who had information did not score as guilty as those who have information and are guilty (see appendix H).

A Duncan's Multiple Range Test on BV means found that both the innocent-uninformed (M =0.27) and innocent and informed participants (M =0.73) score more towards innocence than the guilty and lying participants (M =-1.15). Refer to appendix I for the relevant results.

A multivariate Order effect was found (F(4, 105) =18.2). Univariate tests (appendix G) showed that the effect was due to differences in order for SRR scores (F(1,108)=41.7 and BV scores (F(1,108) =38.6). To review the relevant means for this analysis, refer to table 8.

Table 8

Means of SRR and BV Across Order.

	Order	
	Crime relevant 1st	Crime relevant 2nd
SRR	-2.73	.93
BV	-1.47	1.37

With both measures, when subjects were presented with question pairs in which the critical crime information was in the first position, their scores were more in the guilty direction than when the critical crime information was presented in the second position. No differences were found for the responses from respiration scores.

Although no multivariate gender effect was found, a univariate test found a significant BV difference for gender (F(1,108) =7.7).

Statistical Analyses on Raw Scores

Results of the Skin Resistance Response Analysis

Between Subjects Effects

The analysis of SRR Scores found a gender effect ($F(1, 108) = 6.86$), such that females had larger responses ($M = 12.81$) than males ($M = 8.74$).

There was a condition by order effect ($F(2,108) = 3.79$. The means for this interaction are shown in table 9.

Table 9

Means for the Condition by Order Interaction for SRR

Condition	Order	Mean	Std. Error
innocent	crime rel 1st	15.13	1.90
innocent	crime rel 2nd	9.45	1.90
innocent informed	crime rel 1st	8.68	1.90
innocent informed	crime rel 2nd	9.58	1.90
guilty	crime rel 1st	8.57	1.90
guilty	crime rel 2nd	13.24	1.90

Using Duncan's Multiple Range Test it was found that responses to questions in Order 1 for the innocent uninformed condition were significantly larger ($M = 15.13$) than responses to questions in Order 1 for both the innocent and informed condition ($M = 8.68$, act. dif.=6.45 > crit. diff.=5.87) and the guilty condition ($M = 8.57$, act. dif.=6.56 > crit. diff.=5.99). There were no other significant differences (see appendix K).

Within Subjects Effects

SRR response magnitudes differed amongst blocks of questions ($F(2, 216) = 34.27$). Duncan's Multiple Range Test analysis showed that responses in Block 1

(M =13.9) were larger than responses from Block 2 (M =9.30, act. diff.=4.6 > crit. diff.=1.27) and Block 3 (M =9.12, act. diff.=4.78 > crit. diff.=1.34). Blocks 2 and 3 did not differ from each other (see appendix K).

A significant effect (F(2, 216)=8.59) occurred for Pairs. Duncan's Multiple Range Test analysis showed that SRR responses in the first (M =11.75)and second (M =10.96) pairs were significantly larger than responses from the third pair (M =9.60, act. diff.=2.15 > crit. diff.=1.08) and (act. diff.= 1.36 > crit. diff.=1.03), respectively. Pair 1 and Pair 2 were not significantly different from each other (see appendix K).

SRR responses differed significantly (F(1, 108)=14.34) between question positions, such that responses to question position number one (M =11.42) were larger than those to question position number two (M =10.12).

An order by question position interaction (F(1, 108)=5.08) occurred and the relevant means are presented in table 10.

Table 10

Means for the Order by Question Position Interaction

Order	Question Position	Mean	Std. Error
Crime rel 1st	1	11.83	1.08
Crime rel 1st	2	9.75	1.16
Crime rel 2nd	1	11.02	1.08
Crime rel 2nd	2	10.49	1.16

Duncan's Multiple Range Test analysis (see appendix K) showed that SRR responses to critical questions in the first position (M =11.83) were larger than responses

to critical questions in the second position (M =10.49, act. diff.=1.34 > crit. diff.=.99) and larger than responses to control questions in the second position (M =9.75, act. diff.=2.08 > crit. diff.=1.03). Responses to control questions in the first position (M =11.02) were also larger than responses to control questions in the second position (M = 9.75, act. diff.=1.27 > crit. diff.=.99).

A question position by condition by order interaction occurred (F (2, 108)=2.99). Refer to table 11 for the means of this analysis.

Table 11

Means for the Question Position by Condition by Order Interaction

Condition	Order	Mean QP 1	Mean QP 2
innocent	crime rel 1st	15.89	14.36
innocent	crime rel 2nd	10.26	8.64
innoc inf	crime rel 1st	9.69	7.67
innoc inf	crime rel 2nd	10.15	9.01
guilty	crime rel 1st	9.90	7.23
guilty	crime rel 2nd	12.64	13.83

Note. QP = question position.

A Duncan Multiple Range Test analysis (appendix K) found that, in the innocent condition, responses to critical questions in the first position (M = 15.89), while not significantly different than those to control questions in that condition (M =14.36), were larger than other responses to all questions in any of the other conditions. The responses to control questions in the innocent condition (M =14.36) were not significantly different from the responses to control questions in the guilty condition in the first position (M =12.64). Guilty questions in the second position (M =13.83) were larger than any other

responses. Responses to the critical question in the first position, in the innocent and informed condition (M =9.69), were larger than responses to its control questions (M =7.67, act. diff.=2.02 > crit. diff.=1.82). In the guilty condition, when the critical question was in the first position (M =9.9), the average response was larger than to that of its control question (M =7.23, act. diff.=2.67 > crit. diff.=1.90). The average responses to both the critical (M =13.83) and control (M =12.64) questions in the second order in the guilty condition were larger than the responses to questions in either position in either order in condition 2 and to those in the second order of condition 1 and the first order of condition 3.

A block by pairs interaction (F(4, 432)=9.41) was found (see means in table 12).

Table 12

Means for the Blocks by Pairs Interaction

Block	Pair	Mean	Std. Error
1	1	17.04	1.14
1	2	13.25	.93
1	3	11.41	1.12
2	1	9.59	.92
2	2	10.29	1.06
2	3	8.03	.94
3	1	8.63	.84
3	2	9.36	.92
3	3	9.36	1.03

Duncan's Multiple Range Test analysis found that the pairs in the first block were larger than any pairs in other blocks except for the second pair in block 2 (see appendix K). That is, the third pair in the first block (M =11.41) did not differ from the second pair

in block 2. The second pair in the second block was significantly larger than the third pair in the second block (M =8.03). No other pairs were different from each other.

A block by question position interaction was found (F(2, 216)=7.91). The relevant means are presented in table 13.

Table 13

Means for the Blocks by Question Position Interaction

Block	Question Position	Mean	Std. Error
1	1	15.23	1.00
1	2	12.57	1.02
2	1	9.43	.77
2	2	9.18	.84
3	1	9.62	.84
3	2	8.61	.90

Note. QP = question position

Duncan Multiple Range Test analysis found that block 1, when the guilty question was in the first position (M =15.23), produced the largest significant response (appendix K). Following this, when the control question was in the first position in block 1 (M =12.57) the response was significantly larger than any others. When the critical question was in the first position for the third block (M =9.62), the response was significantly larger than its control question (M =8.61).

A block by question position by condition by gender interaction occurred (F(4, 216)=2.98). The relevant means for this interaction are presented in table 14.

Table 14

Means for the Block by Question Position by Condition by Gender Interaction

Block	QP	Condition	Gender	Mean	Std. Error
1	1	innocent	male	15.30	2.45
1	1	innocent	female	18.88	2.45
1	1	innoc inf	male	12.28	2.45
1	1	innoc inf	female	14.36	2.45
1	1	guilty	male	13.72	2.45
1	1	guilty	female	16.82	2.45
1	2	innocent	male	9.59	2.49
1	2	innocent	female	17.39	2.49
1	2	innoc inf	male	9.06	2.49
1	2	innoc inf	female	11.45	2.49
1	2	guilty	male	12.22	2.49
1	2	guilty	female	15.73	2.49
2	1	innocent	male	7.61	1.88
2	1	innocent	female	14.51	1.88
2	1	innoc inf	male	5.86	1.88
2	1	innoc inf	female	8.94	1.88
2	1	guilty	male	8.08	1.88
2	1	guilty	female	11.56	1.88
2	2	innocent	male	8.21	2.06
2	2	innocent	female	12.94	2.06
2	2	innoc inf	male	5.19	2.06
2	2	innoc inf	female	9.00	2.06
2	2	guilty	male	7.19	2.06
2	2	guilty	female	12.54	2.06
3	1	innocent	male	8.75	2.06
3	1	innocent	female	13.40	2.06
3	1	innoc inf	male	7.35	2.06
3	1	innoc inf	female	10.76	2.06
3	1	guilty	male	6.79	2.06
3	1	guilty	female	10.68	2.06
3	2	innocent	male	7.04	2.19
3	2	innocent	female	13.83	2.19
3	2	innoc inf	male	5.86	2.19
3	2	innoc inf	female	9.46	2.19
3	2	guilty	male	7.21	2.19
3	2	guilty	female	8.29	2.19

Note. QP = question position

The Duncan Multiple Range Test was conducted. Refer to appendix K for the results of the analysis. The major finding was that females in the innocent uninformed condition in the first block of questions produced the largest responses. Significant differences occurred for both the first (M =18.9) and second (M =17.4) question positions. Females in the guilty condition, in the first block, also had large responses (M =16.8 and M =15.7) for both question positions. A block by pairs by question position interaction was found (F(4, 432) =3.69). Table 15 presents the means for this analysis.

Table 15

Means for the Blocks by Pairs by Question Position Interaction

Blocks	Pairs	QP	Mean	Std. Error
1	1	1	19.66	1.23
1	1	2	14.42	1.27
1	2	1	13.62	.96
1	2	2	12.88	1.09
1	3	1	12.40	1.19
1	3	2	10.42	1.16
2	1	1	9.32	.98
2	1	2	9.85	.99
2	2	1	10.72	1.05
2	2	2	9.86	1.25
2	3	1	8.24	.89
2	3	2	7.82	1.11
3	1	1	9.00	.91
3	1	2	8.26	.97
3	2	1	9.66	1.05
3	2	2	9.06	1.00
3	3	1	10.20	1.06
3	3	2	8.53	1.17

Note. QP = question position

The results of the Duncan Multiple Range Test (see appendix K) was that the first 2 pairs of questions in block 1 produced responses that were larger than virtually any

other responses across the three conditions. Questions in the first position in block 1 pair 1 (M =19.7) produced responses larger than any others. The responses to questions in the second position (M =14.4) in Block 1 Pair 1 were larger than those in Block 1 Pair 3 and all other Pairs in Blocks 2 and 3. The second pair of questions in either position in block 1 were significantly larger than any responses in Blocks 2 or 3, regardless of Pairs. Although the difference in each pair was not always significant, the general trend was that the questions in the first position were larger than when the critical question was placed in the second position.

Results of the Thoracic Analysis

Between Subjects Effects

The analysis of Thoracic scores found a gender effect (F(1, 108)=17.57). Male respiratory responses were significantly larger (M =1.96, Std.Error=0.07) than those of females (M =1.55, Std.Error=0.07).

Within Subjects Effects

Scores differed amongst the blocks (F(2, 216)=12.72). The Duncan Multiple Range Test analysis (appendix M) found that respiratory scores in block 1 (M =1.83) were significantly larger than in block 2 (M =1.75, act. diff.=.08 > crit. diff.=.05) and responses to block 3 (M =1.70, act. diff.=.13 > crit. diff.=.053). Blocks 2 and 3 were not significantly different from one another.

A condition by order by blocks interaction (F(4, 216)=2.54) occurred. The relevant means of this interaction are presented in table 16.

Table 16

Means for the Condition by Order by Blocks Interaction

Condition	Order	Block	Mean	Std. Error
Innocent	crime rel 1st	1	1.96	.13
Innocent	crime rel 1st	2	1.92	.12
Innocent	crime rel 1st	3	1.84	.12
Innocent	crime rel 2nd	1	1.88	.13
Innocent	crime rel 2nd	2	1.75	.12
Innocent	crime rel 2nd	3	1.69	.12
Innocent Inf	crime rel 1st	1	1.51	.13
Innocent Inf	crime rel 1st	2	1.40	.12
Innocent Inf	crime rel 1st	3	1.42	.12
Innocent Inf	crime rel 2nd	1	1.85	.13
Innocent Inf	crime rel 2nd	2	1.87	.12
Innocent Inf	crime rel 2nd	3	1.79	.12
Guilty	crime rel 1st	1	1.98	.13
Guilty	crime rel 1st	2	1.74	.12
Guilty	crime rel 1st	3	1.73	.12
Guilty	crime rel 2nd	1	1.78	.13
Guilty	crime rel 2nd	2	1.79	.12
Guilty	crime rel 2nd	3	1.76	.12

The results of the Duncan Multiple Range Test analysis (appendix M) found that the largest response score, indicative of no respiratory suppression, was obtained when the critical question was in the first position in the first block of questions in the guilty condition ($\underline{M}$ =1.98). Although not significantly different from the largest response, almost equal responses occurred in the innocent uninformed condition in the first 2 blocks ($\underline{M}$ =1.96 and $\underline{M}$ =1.92), and these responses were larger than the majority of other responses. The next largest response was obtained from the innocent uninformed group ($\underline{M}$ =1.88) when the critical question was in the second position. The greatest respiratory

suppression came from those who were in the innocent-yet informed condition (M =1.40), who were given the critical stimulus in the first position.

A Condition by Gender by Pairs by Question Position interaction (F(4, 216)=3.18) occurred and the relevant means are presented in table 17.

Table 17

Means for the Condition by Gender by Pairs by Question Position Interaction

Gender	Pair	QP	Condition Innocent	Innocent-Inf	Guilty
Male	1	1	2.12	1.79	2.05
Male	1	2	2.11	1.84	1.92
Male	2	1	2.08	1.82	1.99
Male	2	2	2.09	1.83	2.01
Male	3	1	2.03	1.84	1.95
Male	3	2	2.12	1.82	1.93
Female	1	1	1.59	1.48	1.60
Female	1	2	1.57	1.44	1.65
Female	2	1	1.56	1.47	1.66
Female	2	2	1.59	1.46	1.63
Female	3	1	1.63	1.44	1.56
Female	3	2	1.59	1.44	1.60

Note. QP = question position

The results of the Duncan Multiple Range Test analysis (appendix M) found that responses from males showed lesser suppression of respiration than responses from females, regardless of question position. Scores from innocent uninformed males (means of 2.12, 2.11, 2.08, 2.09, 2.03, and 2.12 respectively) showed the least suppression of Thoracic Respiration.

Each gender shows differences amongst their respective groups but the female scores within each of their groups do differ between pair scores and question position whereas the male scores in the guilty condition in the first pair differ between question position. This is the essence of the interaction.

Results of the Abdominal Respiration Analysis

Between Subject Effects

The analysis of abdominal scores found a gender effect ($F(1, 108)=33.64$). Responses from males were significantly larger (M =2.34, Std. error =0.105) than those from females (M =1.48, Std. error =0.105).

Within Subjects Effects

Scores amongst blocks differed significantly ($F(2, 216)=4.79$). A Duncan's Multiple Range Test analysis (see appendix O) found that responses in the first block were significantly larger (M =2.02) than responses in block 2 (M = 1.86, act. diff.=.16 > crit. diff.=.11) and block 3 (M = 1.85, act. diff.=.17 > crit. diff.=.12). Responses from block 2 and block 3 were not significantly different from each other.

Scores differed between pairs ($F(2, 216)=4.05$). Duncan Multiple Range Test (see appendix O) found that the responses in the first pair were significantly larger (M =1.95) than responses in pair 2 (M =1.89, act. diff.=.06 > crit. diff=.055) and pair 3 (M =1.88, act. diff.=.07> crit. diff.=.058). There were no significant differences between the responses of the other pairs.

A significant interaction ($F(2, 216)=3.04$) occurred between gender and pairs (refer to Table 18).

Table 18

Means for the Gender by Pairs Interaction

Gender	Pair	Mean	Std. Error
Male	1	2.42	.11
Male	2	2.30	.10
Male	3	2.30	.11
Female	1	1.48	.11
Female	2	1.49	.10
Female	3	1.46	.11

Duncan's Multiple Range Test analysis (appendix O) found that the responses obtained from the first pair for males (M =2.42) were significantly larger than responses from any other pair for either males or females. The responses from the second and third pairs for males were of equal value (M =2.30) and significantly larger than responses from any of the pairs for females. There were no differences amongst the responses from female pairs.

Statistical Analysis on Recall and Recognition Scores

A MANOVA was conducted with the recall and recognition scores as dependent variables and Gender, Condition, and Order between-subject factors as independent variables. A condition effect was found for both recall (F (2, 120) = 70.12, p =0.00) and recognition (F (2, 120) = 78.21, p =0.00). No significance was found for either Gender or Order. Subsequent post hoc analyses using the Tukey HSD method found that innocent participants given no guilty information (M recall = 1.13, M recognition = 1.30) showed

chance levels of memory whereas participants given information in the innocent informed condition (M recall = 2.75, M recognition = 2.83) and guilty condition (M recall = 2.90, M recognition = 3.00) remembered the information.

Chapter 4

Discussion

Support for the main hypothesis was found. It was predicted that the relative magnitude of responding to questions would be jointly determined by position and relevance. Support was from classifications showing that the majority of guilty participants were judged correctly regardless of whether crime relevant questions were in the first or second position, whereas the majority of innocent participants were judged correctly only when the crime relevant question was in the second position. These results indicated that large responses occurred to questions by virtue of position such that when the crime relevant question was first participants, regardless of their crime condition, responded as if they were guilty. When, however, the crime relevant question was second, relevance or salience became important. The result was that guilty participants had relatively large reactions to the crime relevant questions in the second position but the innocent participants did not. Classifications based on the composite of all scores, SRR scores and blood volume scores were the source of these results.

Guilty participants reacted because of the unique importance of the crime relevant question to them. That is, they had committed the actions involved with the questions and had to lie to those questions. The crime relevant information could be considered important for innocent informed participants because they have the knowledge. It turned out, empirically, they were not relatively more responsive. The suggestion is that mere knowledge of the events is not as important as lying about the behaviour.

Some additional predicted findings were supported. Guilty participants were, in general, more likely to be classed as guilty than innocent participants. Most participants were classed as guilty when the crime relevant question was first than when it was second. These results were found with total, blood volume, and SRR scores.

There are two ways that clear support for the main hypothesis could be found in the derived score analysis. It was expected from the hypothesis that a condition by order interaction would occur. That is, large responses would occur to crime relevant questions in the first position for all participants, regardless of condition, but would occur to crime relevant questions only for those in the guilty condition when the crime relevant questions were in the second position.

The other way that support could be found was not anticipated but was the way that it empirically happened. That is, when the crime relevant questions were in the first position, all participants scored as guilty but the guilty scored as most guilty (most negative mean score). In the second position scores became, in general, less negative. They only became actually positive in the innocent conditions. Thus, when the crime relevant question was in the second position, innocent participants were innocent and the guilty were judged guilty. These results held for the SRR and blood volume scores as the main contributors in the multivariate analysis. Results similar to the classification analysis were found with innocent participants and order effects. That is, innocent participants did not differ from each other. Again, this confirms that knowledge in this context does not result in an

appearance of guilt. The order effect confirmed that participants tended to score as guilty when the crime relevant question was first.

The desired interaction to support the main hypothesis did occur in the raw score analysis. Unfortunately, exploration with the post hoc tests indicated that the interaction was not for the hypothesised reasons. As it turned out, responses from the innocent uninformed participants, when the crime relevant question was in the first position, were so large that they exceeded responses to every other question but their own control. The responses to control questions in that position were larger than nearly every other question. These exceptional differences may have masked the potential difference between the control question in the first position and the critical question in the second position in the guilty group.

Why innocent uninformed participants were so responsive to crime relevant information that they are unaware of poses a mystery. It is possible that participants accused of a crime, but unaware of the relevance of any information, may simply be reacting to the uncertainty of the entire interrogation process. That is, for the innocent participants, question relevance is viewed differently from the innocent informed and the guilty groups. The latter two groups, by virtue of their knowledge of the crime relevant questions can also identify control questions. Innocent uninformed participants are ignorant of which questions are which during the polygraph examination. It is possible that relevant knowledge can provide some focus of attention or direction and, without such knowledge, individuals are uncertain and over-react, perhaps experiencing and exhibiting confusion.

As plausible as it sounds, even this explanation is problematic. The innocent uninformed, when the crime relevant question was in the second position, had average responses for both the crime relevant and control questions. Since they have no information, the scores should have been the same across both positions.

In spite of the lack of support for the main hypothesis, the raw score analyses still provide information on the process of habituation. With SRR scores, habituation is reflected in the diminishment of large amplitude responses over repeated stimulus presentations.

The raw score analysis supported SRR habituation through a block effect such that responses in the first block were significantly larger than those in blocks 2 and 3 (which were not significantly different from each other). Additionally, responses to initial pairs of questions, presented in the first block for SRR measures were larger than responses to other pairs in that block and to other pairs in successive blocks.

Further support comes from the block by pairs interaction. Responses in the first 3 pairs in the first block, in descending order, were the largest responses. This pattern of responding was essentially repeated in the third block. The interaction came about because of an anomaly of large responses to questions in the third pair in the second block.

There was a block by pairs by question position interaction such that almost all the habituation took place in the first block. This was particularly so for the first question position. This provides some support for habituation. That which is presented first is large, regardless of question content. Also, the results showed that habituation occurs rapidly.

Respiration results are more complex. The measure, the linear length of waves of respiration, is affected by both the volume of each breath and the respiration rate. Reactivity

is reflected in respiratory suppression. Habituation specific to stimuli should be reflected in an increase in amplitude but a decrease in rate. Simultaneously, it is possible for participants to become more relaxed in the detection situation and this could result in an overall reduction in respiration. Therefore it is difficult to gauge exactly which process is dominant. That is, is respiration diminished because of reactivity or because of relaxation? There are four results that pertain to the habituation question that reflect this ambiguity. Both abdominal and thoracic responses showed a decrease from block 1 to block 3. Additionally, Abdominal respiration showed a decrease over the 3 pairs of responses, regardless of blocks.

A block by condition by order interaction on the thoracic measure resulted from innocent participants being the least responsive, progressing through blocks 1-3, with scores significantly lower in the third block.

Lastly, the results from the thoracic pairs by question position by condition by gender interaction found that both guilty males and females, when the critical question was in the first position, evidenced reduced reactivity in their third block responses compared to their earlier responses.

In summary, this level of analysis found support from the block, pairs, and question position factors concerning the responses to the SRR measure to demonstrate the habituation effect. It was unclear from the respiration response measures which process (reactivity or relaxation) was dominant during the polygraph examination.

Analyses for memory scores showed a condition effect for both recall and recognition. The subsequent analyses indicated that the groups with information (innocent-

informed and the guilty groups) both had almost complete recall and recognition memory of the crime material. While not different from each other, they differed from the innocent uninformed group who had low memory scores that were at chance levels. Even though the guilty and innocent informed groups had similar scores, they differed in detection rates. This finding suggests that knowledge, while a necessary condition for detection, is not necessarily a factor in the accurate detection of deception. This finding is counter to a memory explanation for detection. Even though innocent-informed participants had knowledge, that knowledge did not result in detection.

Such a finding extends work by Waid et al. (1978), in using the GKT, who found that recalled words were more likely to evoke an electrodermal response compared to non-recalled words. Waid, Orne, and Orne (1981), subsequently examined recall using both the GKT and the CQT and found that both correctly recalled relevant and control questions produced larger skin conductance responses when compared to non-recalled words. Iacono, Boisvenu, and Fleming (1984) found a positive correlation of .53 between recalled items viewed on a videotaped crime and skin conductance responsivity to critical items.

These studies, however, all presented correlational data. Ben-Shakhar and Furedy (1990) pointed out the relationship between psychophysiological responding and memory may not be a causal one. The present study affirms that other factors (lying or actions) affect responses to the various questions.

There is an issue regarding the study of memory in detection that may change how the area is viewed. This study, together with many of the other studies in the literature (e.g. Bradley, MacLaren, and Black, 1996), required subjects to remember what we as

researchers and investigators believe is pertinent crime relevant information that the subject must be able to remember. In fact, subjects in laboratory studies typically read over and study the important material (e.g. Waid et al., 1981). In real life, however, materials selected by investigators may not actually pertain to the types of information that suspects remember.

A case in point is a purse-snatcher. He may steal 10 purses in one evening while in a busy downtown market area. He may dispose of the purses in a number of ways, not look for any identification (just the cash), and may not notice the colour of any particular purse. He views it as simply a sack with money in it.

If the polygrapher chooses crime relevant information garnered from only one of the thief's ten victims, the questions devised may be too specific and, in fact, be unknown to the thief. For example, questions relating to the amount of money contained in a specific purse, the owner of the purse, the exact location of the theft, the colour of the purse, etc. will likely meet with only marginal success at best. Add to this the possibility that the thief may have been under the influences of substances or "hurting" for a "fix", and the perpetrator's focus of attention, and what information he might retain, is up for debate. Perhaps this should be the case. The example of the purse-snatcher suggests that memory may be state dependent or lie within a context effect. Future research may benefit from more investigation of what various types of offenders look for, attend to, and remember during the planning of, commission of, or events following a crime.

Gender effects were found in the present study. These effects were found in the raw score analyses. Overall, females showed significantly larger responses on SRR scores (eta =.24), and significantly smaller responses on Abdominal (eta =.49) and Thoracic (eta =.37) measures. Typically, researchers interpret larger SRR and smaller respiration responses with deceptive or guilty subjects. In the present study respiration responses with females were in these incriminating directions, regardless of their guilt or innocence. Therefore, their responses do not appear to be due purely to guilt. A more likely interpretation is that the lower abdominal and thoracic scores probably reflect the smaller physical size of the females.

The gender differences are not evident in the derived score analysis. This is likely due to the fact that the one summary score that is supplied for each physiological measure for each participant is based on subjects being their own control. Therefore, potential size differences cannot enter into the analyses. In contrast to this, the raw score analysis examined eighteen data points per measure per participant. It is likely that this large data set per measure allowed the differences to become evident.

The present study did include gender as a factor but there was no strong reason from the literature to predict differences. A sampling of 45 studies (Bradley, 1998, unpublished) found gender mentioned in the subject section 20 times. Nine of these studies explicitly mention testing for gender differences. Two studies, Bradley and Cullen (1993) and Honts, Hodes, and Raskin (1985) found gender differences. The other studies did not. It was concluded by Ben-Shakhar and Furedy (1990) that potential gender effects may be small.

The overall results of this study may have both practical and theoretical significance if they can be generalised to field situations. Examiners could severely bias a test towards classifying innocent suspects as guilty by placing crime relevant questions in the first position in CQT pairs. Alternatively, placing the crime relevant questions in the second position allows more accurate classification of innocent suspects without a significant sacrifice of accuracy with guilty suspects.

In regards to theory and the structure of tests, the necessity of certain attributes for control questions is called into question. Ambiguity, incrimination, and lying apparently are not necessary for ample relative responding if the control question is in the first position. An orienting response to a clear, unambiguous question in the initial position may be sufficient to protect innocent suspects from false judgements of guilt. An explanation of the alleged success of the irrelevant / relevant technique could be formulated on this basis. The irrelevant question in the first position evokes a full OR whereas the relevant question for the innocent suspects evokes a somewhat habituated OR. One qualifying contrast for the irrelevant / relevant test and the actual truth control question test is that the control question in the actual truth control question test is related through topic and plausibleness to the crime relevant item. In this circumstance, it should be more effective to habituate the response to the subsequent crime relevant question.

One problem that appears to be very complicated with the current findings is the question of whether standard control questions test attributes augment responding beyond the expected initial position OR response value. The raw score analysis did not support the idea that lying by the guilty participants resulted in additional increments to responses when

crime relevant questions were in the initial position. By this token, standard control question attributes, when they are in the initial position, may not augment responsiveness. If they do not augment responsiveness, the contention over the effectiveness of these attributes (Lykken, 1981) appears misguided since the real issue is over OR evocation. On the other hand, in the second position, in a fashion analogous to the large responsiveness to crime relevant questions by guilty participants, control question attributes could theoretically allow control question questions to be placed in the second position and still result in an effective test. This remains an open-ended question.

The conundrum is presented by the current data. Lying did not augment the initial question OR's but lying to the second question resulted in responses that exceeded those resulting from initial position OR's. It is difficult to speculate on a mechanism that would accommodate this finding. The OR process would have to be conceived of as pre-eminent and other simultaneous processes (lying) are not expressed until the process has habituated.

Summary

The results of this study found that the order of presentation of information has a significant effect on the accuracy of detection. Crime relevant information in the second position results in greater test accuracy for both groups of innocent subjects without significantly sacrificing accuracy in the guilty condition.

Questions in the first position evoke orienting responses, regardless of knowledge or lying. When the questions in the first position were relevant to the crime many innocent subjects were misclassified.

To compare this form of the CQT to other studies, the crime relevant information has to be in the second position. The results of the present study found that 73% of the guilty and 87% of the innocent subjects were correctly classified. These numbers are similar to those reported in a review by Ben-Shakhar and Furedy (1990). In that review of 9 studies they found 80% of the guilty and 63% of the innocent subjects were classified correctly. They reviewed an additional 9 studies to evaluate the validity of the field studies of the CQT and found 84% and 72% correct classifications for the guilty and innocent subjects respectively.

Habituation was evident with SRR responses and occurred in the first block over the first two pairs. It was not clear with respiration as relaxation and reactivity were potentially confounded. There was a gender effect on the physiological data and it was probably due to females being physically smaller than the males. Rapid habituation for SRR responses was evident for all three groups of subjects.

An examination of memory indicated that the groups with information did not differ in recall and recognition scores but did differ in terms of detection rates, suggesting that knowledge alone is not sufficient for accurate detection. Of course, in considering application of the present findings, it must be remembered that the study was conducted in the laboratory.

Practical Implications and Future Direction

The control questions in the "Actual Truth Control Question Test", used in this study, are unlike normal Control Questions in that there is no "emotional" content (Bradley, MacLaren, and Black, 1996). That is, there is no pairing of a crime relevant question such

as "Did you steal the money from the drawer?" with an emotionally laden control question such as "Did you ever steal anything of value?". The present test is more like the Guilty Knowledge Test, but with only one buffer question, such as "Did you steal $20?", or "Did you steal $30?".

The results of this study support the contention by Bradley, MacLaren, and Black (1996) that control questions do not have to be emotionally evocative. They must, however, in the same manner as distractor items in the Guilty Knowledge Test, be plausibly related to the crime. This requirement may even be questionable, since informed participants were aware that the control questions were not related to the crime but still responded to them in a guilty manner if they were presented in the first order. This was likely due to the OR effect.

The results with the Control Question Test, with actual controls for truth in this study, replicated the results of Bradley, MacLaren and Black's (1996) study, and indicated better than chance accuracy in classification rates. Not only was this found with Innocent and Guilty participants, but also with Innocent and Informed participants.

The CQT with Actual Truth Control Questions addresses several issues surrounding the CQT. Firstly, the questions involved in the test may be inherently standardised. That is, the crime relevant questions would be on an item related to the crime and, in the same format and style; the control questions would be on a matched item but that just so happens not to be involved in the crime.

Secondly, at the present time the CQT interview is not standardised although such standardisation is required for any psychological test (Anastasi, 1988). The pre-

interrogation interviews for the Actual Truth Control Question Test may be standardised since there would be no need to draw particular attention to the control question features. In fact, there is no need for an extensive interview. A rather brief interview can accomplish the same result. The only feature that might need to be explained is that innocent suspects are telling the truth to both of the questions whereas guilty suspects are lying on one.

Thirdly, the transparency problem (Ben-Shakhar and Furedy, 1990) is dealt with in a direct manner. As previously discussed, the problem arises because both innocent and guilty subjects know that the crime relevant questions are the most important to "pass successfully" to avoid a judgement of guilt. Due to various emotional reactions, innocent suspects may produce similar responses as a guilty suspect on crime relevant questions. In the Actual Truth Control Question Test, if innocent suspects are uninformed to crime details then neither of the questions in each pairing are transparent to the innocent subjects. As a result, those who are innocent and uninformed truthfully answer "no" to both questions with no knowledge of which is which whereas the guilty will answer truthfully to one question and lie in responding to the other item.

The present study supplies two observations for the transparency problem. Empirically knowledgeable participants did not respond as guilty simply on the basis of recognising the crime relevant question. The other observation is that the Actual Truth Control Question Test could be administered like a Guilty Knowledge Test with the suspect completely unaware of the information.

More realistically, at this stage, this study has raised intriguing possibilities for future empirical work. That is, with the Actual Truth Control Question Test containing the

elements of a standard control question, researchers could examine specific aspects of the interrogation process. We have already manipulated position and information elements. Future work could examine lying (i.e. a suspect could be interrogated for crime "A" but had also committed crime "B"), and ambiguity (i.e. many suspects would be unsure of how to respond to a particular control question).

Further investigation in this area is warranted, as it offers both a new direction as well as a possible method that may move us closer towards achieving more accuracy in detecting deceit.

References

Anastasi, A. (1988). Psychological Testing (sixth edition). New York: Macmillan Publishing Company.

Barland, G.H., and Raskin, D.C. (1973). The use of electrodermal activity in the detection of deception. In W.F. Prokasy and D.C. Raskin (Eds.), Electrodermal Activity in Psychological Research. New York: Academic Press.

Beatty, J., and Khaneman, D. (1966). Pupillary changes in two memory tasks. Psychonomic Science, 5, 371-372.

Ben-Shakhar, G. (1977). A further study of the dichotomization theory in detection of information. Psychphysiology, 14, 408-413.

Ben-Shakhar, G,, Asher, T., Poznansky-Levy, A. Asherwitz, R., and Lieblich, I. (1989). Stimulus novelty and significance as determinants of electrodermal responsivity: the serial position effect. Psychophysiology, 26, 29-38.

Ben-Shakhar, G., and Furedy, J.J. (1990). Theories and Applications in the Detection of Deception. New York: Springer-Verlag.

Ben-Shakhar, G. and Lieblich, I. (1982). The dichotomization theory for differential autonomic responsivity reconsidered. Psychophysiology, 19, 277-281.

Bradley, M.T. (1998). Selected survey of 45 studies to examine gender selection practices. Unpublished raw data.

Bradley, M.T., Carle, S.C., and Nagaragi, P. (1998, August). Choice, instructions, and responding on a concealed information test, Paper presented at the International Congress of Applied Psychology, San Francisco, Calif., U.S.A.

Bradley, M.T., and Cullen, M.C. (1993). Polygraph lie detection on real events in a laboratory setting. Perceptual and Motor Skills, 76, 1051-1058.

Bradley, M.T., Cullen, M.C., and Carle, S.B. (1993, May). Ratings of emotionality, embarrassing stories and the detection of deception. Poster presentation at the Canadian Psychological Associations' annual conference (Montreal, 1993).

Bradley, M.T., Cullen, M.C., and Carle, S.B. (1996). Control question tests by police and laboratory polygraph operators on a mock crime and real events. Polygraph, 25, 1-14.

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

Bradley, M.T., MacLaren, V.V., and Black, M.E. (1996). A True Control Polygraph Control Question Test. Perceptual and Motor Skills, 83, 755-762.

Bradley, M.T., and Rettinger, J. (1992). Awareness of crime relevant information and the guilty knowledge test. Journal of Applied Psychology, 77, 55-59.

Bradley, M.T., and Warfield, J.F. (1984). Innocence, information, and the guilty knowledge test in the detection of deception. Psychophysiology, 21, 683-689.

Cacioppo, J.T., and Petty, R.E. in Social Psychophysiology- a Sourcebook, Cacioppo and Petty(Eds). 1983, The Guildford Press, N.Y., N.Y.,10003.

Cannon, W.B.(1927). The James-Lange theory of emotions: A critical examination and an alternative theory. American Journal of Psychology, 39,106-124.

Carle, S.B. (1996). A Crime-oriented Approach to Emotions and Lie Detection. Unpublished Masters of Arts Thesis (still in process). University of New Brunswick in Saint John.

Craig, F.I.M., and Lockhart, R.S. (1972). Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Cutrow, R.J.; Parks, A., Lucas, N., and Thomas, K. (1972). The objective use of multiple physiological indices in the detection of deception. Psychophysiology. 9, 578-588.

Davis, R.C. (1961). Physiological responses as a means of evaluating information. In A.D. Biderman and H. Zimmer (Eds.), The Manipulation of Human Behavior (142-169). New York: Wiley.

Easterbrook, J.A. (1959). The effect of emotion on cue utilisation and the organization of behavior. Psychological Review, 66, 183-201.

Elaad, E.(1987). Psychophysiological Detection in the Guilty Knowledge Test. Unpublished doctoral dissertation, the Hebrew University of Jerusalem, Israel.

Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. Journal of Applied Psychology, 75, 521-529.

Elaad, E. (1994). The accuracy of human decisions and objective measurements in psychophysiological detection of knowledge. Journal of Psychology, 128, 267-280.

Elaad, E., and Ben-Shakhar, G.(1989). Effects of motivation and verbal response type on psychophysiological detection in the Guilty Knowledge Test. Psychophysiology, 26, 442-451.

Elaad, E., and Elaad, I.(1994). The control question technique in vocational search. Perceptual and Motor Skills, 79, 825-826.

Elaad, E., Ginton, A., and Jungman, N. (1992). Detection measures in real-life criminal guilty Knowledge tests. Journal of Applied Psychology, 77, 757-767.

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

Furedy, J.J., Davis, C., and Gurevich, M. (1988). Differentiation of deception as a psychological process: a psychophysiological approach. Psychophysiology, 25, 683-688.

Furedy, J.J., Posner, R.T., and Vincent, A. (1991). Electrodermal differentiation of deception: perceived accuracy and perceived memorial content manipulations. International Journal of Psychophysiology, 11, 91-97.

Geisen, M., and Rollinson, M.A., (1980). Guilty knowledge versus innocent associations: effects, of trait anxiety and stimulus context on skin conductance. Journal of Research in Personality, 14, 1-11.

Heslegrave, R.J. (1981). A psychophysiological analysis of the detection of deception: The role of information, retreival, novelty and conflict mechanisms., Unpublished Doctoral dissertation, University of Toronto, Ontario, Canada.

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

Honts, C.R., Raskin, D.C., and Kircher, J.C. (1987). Effects of physical countermeasures and their electromyographic detection during polygraph tests for deception. Journal of Psychophysiology, 1, 241- 247.

Horneman, C.J., and O'Gorman, J.G. (1985). Detectability in the card test as a function of the subject's verbal response. Psychophysiology, 22, 330-333.

Horvath, F.S., (1978). An experimental comparison of the psychological stress evaluator and the galvanic response in detection of deception. Journal of Applied Psychology, 63, 338-344.

Iacono, W.G., Boisvenu, G.A., and Fleming, J.A. (1984). Effects of diazepam and methylphenidate on the electrodermal detection of guilty knowledge. Journal of Applied Psychology, 69, 289-299.

Iacono, W.G. and Patrick, C.J. (1988). In Rogers, R. (Ed.), Clinical Assessment of Malingering and Deception (pp. 205- 233). New York: The Guilford Press.

Janisse, M.P., and Bradley, M.T. (1980). Deception, information, and the pupillary response. Perceptual and Motor Skills, 50, 748-750.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. Science, 145, 1583-1585.

Kiang, L.E. (1996). Polygraph testing procedures in Malaysia. Polygraph, 25, 59-72.

Kircher, J.C., and Raskin, D.C. (1988). Human versus computerized evaluations of polygraph data in laboratory settings. Journal of Applied Psychology, 73, 291-302.

Kleinmuntz, B., and Szucko, J.J. (1984). Lie detection in ancient and modern times: a call or contemporary scientific study. American Psychologist, 39, 766-776.

Krapohl, D.J. (1996). A taxonomy of polygraph countermeasures. Polygraph, 25, 35-56.

Kugelmass, S., and Lieblich, I. (1966). The effects of realistic stress and procedural interference in experimental lie detection. Journal of Applie Psychology, 50, 211-216.

Kugelmass, S., and Lieblich, I. (1968). An analysis of mechanisms underlying psychophysiological detection. Proceedings of the 15th International Congress of Applied Psychology, Amsterdam. 509-512.

Kugelmass, S., Lieblich, I., and Bergman, Z. (1967). The role of "lying" in psychophysiological detection. Psychophysiology, 3, 312-315.

Kugelmass, S., Lieblich, I., Ben-Ishai, A., Opatowski, A., and Kaplan, M. (1968). Experimental evaluation of galvanic skin response and blood volume change indices during criminal interrogation. The Journal of Criminal Law, Criminology, and Police Science, 59, 632-635.

Levenson, R.W., Ekman, P., and Friesen, W.V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. Psychophysiology, 27, 363-384.

Lieblich, I., Kugelmass, S., and Ben-Shakhar, G.(1970). Efficiency of SRR detection of information as a function of stimulus set size. Psychophysiology, 6, 601-608.

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

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

Lykken, D.T. (1981). A Tremor in the Blood: Uses and Abuses of the Lie Detector. New York: McGraw-Hill.

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

Marston, W.M., (1917). Systolic blood volume changes in deception. Journal of Experimental Psychology, 2, 143-163.

Martindale, C.(1981). Cognition and Consciousness. The Dorsey Press: Homewood, Illinois.

Ortony, A. and Turner, T.J. (1990). What's basic about basic emotions? Psychological Review, 97, 315-331.

Raskin, D.C., Kircher, J.C., Horowitz, S.W., and Honts, C.R. (1989). In Yuille, J.C. (Ed.), Credibility Assessment (pp. 1-24). Dordrecht, The Netherlands: Kluwer Academic Press.

Reid, J.E. (1947). A revised questioning technique in lie-detection tests. Journal of Criminal Law and Criminology, 37, 542-547.

Reid, J.E. and Inbau, F.E. (1977). Truth and Detection, the Polygraph ("Lie Detector") Technique. (2nd ed.). Baltimore, MD: Williams and Wilkins.

Saxe, L, (1991). Science and the CQT polygraph: a theoretical critique. Integrative Physiological and Behavioral Science, 26, 223-231.

Saxe, L., Dougherty, D., and Cross, T.P. (1985). The validity of polygraph testing: Scientific analysis and public controversy. American Psychologist, 40, 355-366.

Simpson, H.M., and Hale, S.M. (1969). Pupillary changes during a decision-making task. Perceptual and Motor Skills, 29, 495-498.

Stern, R.M., Breen, J.P., Watanabe, T., and Perry, B.S. (1981). Effects of feedback of physiological information on responses to innocent associations and guilty knowledge. Journal of Applied Psychology, 66, 677-681.

Stern, R.M.; Ray, W.J.; and Davis, C.M. (1980). Psychophysiological Recording. New York: Oxford University Press Inc.

Szucko, J.J., and Kleinmuntz, B. (1981). Statistical versus clinical lie detection. American Psychologist, 36, 488- 496.

Thackray, R.I., and Orne, M.T. (1968a). A comparison of physiological indices in detection of deception. Psychophysiology, 4, 329-339.

Thackray, R.I., and Orne, M.T. (1968b). Effects of the type of stimulus employed and the level of subjectawareness on the detection of deception. Journal of Applied Psychology, 52, 234-239.

Timm, H.W. (1982). Analyzing deception from respiration patterns. Journal of Police Science and Administration, 10, 47-51.

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

Waid, W.M., Orne, E.C., Cook, M.R., and Orne, M.T. (1978). Effects of attention, as indexed by subsequent memory, on electrodermal detection of information. Journal of Applied Psychology, 63, 728-733.

Waid, W.M., Orne, E.C., and Orne, M.T., (1981). Selective memory for social information, alertness, and physiological arousal in the detection of deception. Journal of Applied Psychology, 66, 224-232.

Appendix A

Information Given to Participants

Consent Form

"Polygraph Examinations Using the CQT with Actual Truth Control Questions"

A Psychological Research Study

University of New Brunswick (Saint John Campus)

Researcher: Murray C. Cullen, Psychology Department

Advisor: Dr. M.T. Bradley, Hazen Hall (648-5658)

Purpose- The present research involves a polygraph examination. The polygraph, or lie detector, measures heart rate, respiration, cardiac changes, and sweating reactions while responding to questions in which individuals may attempt to be deceptive. The polygraph should assist the examiner in determining whether or not the individual is being truthful or attempting to deceive. This particular research concerns a "mock" crime" that you may, or may not, be guilty of committing. This would depend upon which experimental condition you will be assigned.

Procedure- The study will take place in the psychology office area and Dr. Bradley's lab. If you agree to participate in this study, you may be asked to perform a "mock crime", after which you will be instructed (during a polygraph examination) to claim that you are innocent. Even if you do "steal" in the pre-arranged scenario, you will still be asked to claim your innocence during a polygraph examination.

Following the examination you will be asked to fill out brief memory, recall and recognition tests which may include a small monetary renumeration. Further, if you are found innocent, a sum of 5 dollars will be awarded to you. A written "educational"

component regarding the research questions being investigated will be given to you when the study is completed.

Benefits- The benefits of participation in this study are that it allows you an opportunity to experience, first hand, research in-the-making. You will be contributing to the field of Psychology by being part of this study. Additionally, you will receive a bonus point towards your final course mark.

Risks/Costs- It will take approximately one hour of your time. Appointment times will be made throughout each of five days per week. You may experience some discomfort because you will have physiological measurement devices attached to your body (specifically the arm, hands, and chest) in the standard polygraph configuration. Even though this study involves a "mock" crime, you may feel temporarily "ill at ease" if you have committed the "crime". Additionally, you may experience some nervousness during the 30 minute polygraph examination during which you are claiming innocence, and may be "lying" if you are guilty.

I, _____, have read and understand this information/consent form. By signing this, I agree to participate in this study. I understand that I am free to withdraw from this study at any time without penalty. I understand that there will be a debriefing concerning the study and that I can attend if I so choose.

Signature _____

Date _____

Phone _____ Student # _____

Researcher _____

If I have any questions or concerns, I am free to contact the researchers in the Psychology Department. The work number for Murray Cullen is 636-5957. The number for Dr. Bradley is 648-5658.

Prior to the actual participation in the study, I have answered all of the above noted participant's questions concerning the research to the best of my ability.

_____ _____ _____

Researcher Signature Date

_____ _____ _____

Participant Signature Date

Information concerning the crime:

Guilty Condition

You are going to commit a crime and, after you have committed it, you are going to be interrogated on the polygraph.

You are to perform the following acts:

Go to room 13, a Professor's office, and proceed directly into the room (without knocking). Take out the wallet from his jacket that is hanging over a chair, remove a $20 bill, and stash it in your footwear (left foot). At this point you are to place the wallet back into the jacket and return to meet with the research assistant.

You are to maintain that you are innocent with the polygrapher. The polygraph examiner does not know. Cooperate with the examiner and try to convince him that you are innocent.

Information Concerning the Crime:

Innocent and Informed Condition

You are Innocent. Please stand in the hall and read the following information concerning the crime.

Do not do the crime!!!

Please read the description carefully! You will be accused of the following crime:

The "crime" involves going into a professor's office (room 13) and taking out a wallet from his jacket that is hanging over a chair. The guilty party went directly into the office (without knocking), removed a $20 bill, and stashed it in their footwear (left foot). They then returned the wallet to the jacket and left the room.

Walk down to the main Psychology Office, turn around, and return to report to the research assistant. This should take approximately 2 minutes.

You are to maintain that you are innocent with the polygrapher. The polygraph examiner does not know. Cooperate with the examiner and try to convince him that you are innocent.

Information Concerning the Crime:

Innocent Condition

You are innocent but you will be accused of a crime and interrogated on the polygraph. Walk down to the main Psychology Office, turn around, and return to report to the research assistant. This should take approximately 2 minutes.

You are to maintain that you are innocent with the polygrapher. The polygraph examiner does not know. Cooperate with the examiner and try to convince him that you are innocent.

Appendix B

Orders of Polygraph Examination Questions

Polygraph Examination Questions (Order 1)

Examiner_____Subject____Sex___Con.____Result_____

Name Date

1: Is your last name ?

2: Are you afraid that I will ask you a question that was not reviewed with you?

3: Do you intend to answer each question truthfully?

4: Did you steal 20 dollars?

5: Did you steal 30 dollars?

6: Did you take the money out of a wallet?

7: Did you take the money out of a purse?

8: Is your first name ?

9: Did you stash the money in your footwear?

10: Did you stash the money in your pocket?

Remarks:

Polygraph Examination Questions (Order 2)

Examiner_____Subject____Sex___Con.____Result_____
Name Date

1: Is your last name ?
2: Are you afraid that I will ask you a question that was not reviewed with you?
3: Do you intend to answer each question truthfully?
4: Did you steal 30 dollars?
5: Did you steal 20 dollars?
6: Did you take the money out of a purse?
7: Did you take the money out of a wallet?
8: Is your first name ?
9: Did you stash the money in your pocket?
10: Did you stash the money in your footwear?
Remarks:

Appendix C

Recall and Recognition Tests

Memory Test #1

Name: _____ Date: _____

Instructions: In the spaces provided, please fill in the correct answer for each of the questions. If you cannot remember, guessing is permitted.

1) How much money was stolen?

_____

2) From where was the money stolen?

_____

3) Where was the money "stashed"?

_____

Memory Test #2

Instructions: Please circle the correct answer for each of the questions. If you cannot remember, guessing is permitted.

1) The amount of money stolen was:

a) $100.

b) $50.

c) $40.

d) $20.

e) $10.

2) The money was taken out of:

a) a pocket.

b) some type of foot wear.

c) a wallet.

d) a drawer.

e) a purse.

3) The money was stashed:

a) in your pocket.

b) in your foot wear.

c) in your wallet.

d) a drawer.

e) in a purse.

Appendix D

Educational Component

Educational Component:

Why We Are Conducting This Research

The polygraph, or "lie detector" has been an instrument utilised by criminal investigations for many years. The assumption is that physiological changes accompany an individuals attempt to purposely deceive others and that these changes are measurable. Research has indicated that the sweating of the fingers, slight changes in respiration, and cardiac responses may be measured and provide some degree of assistance in the investigators determination of guilt or innocence with regards to a particular issue.

Over the years many questioning techniques have been investigated, with mixed results. Some researchers believe that the "suspects" being interrogated need a few questions asked before the actual examination takes place. This is to allow for an "orientation" to occur with the suspects. If an individual is placed in a novel situation and supplied with novel stimuli (even if the stimuli is in the form of a question), there may be a larger than anticipated response to that question, simply on the basis that it is novel. During a lie detection examination, the novel response (collected and measured by the polygraph) may be misinterpreted as an attempt to deceive.

It may also be the case that the "orientation" response may assist the examiner in the final determination of guilt or innocence. The present study will study the order effects of question presentation to help determine if the "orientation" response could hinder or enhance the information the examiner uses in making his interpretation of the physiological data, and ultimately his determination of truth or deception. If you would like to receive a

brief summary of the major findings of this research at a later date, please indicate this to the principal investigators.

Appendix E

Additional Classification and Chi Square Analysis

Chi Square of Correct Percentages of Total Scores and Various Measures

	Total		Ab		Th		BV		SRR	
	Cor	Incor	Cor	Incor	Cor	Incor	Cor	Incor	Cor	Incor
Order 1										
Observed	42	58	49	51	36	64	60	40	48	52
Expected	62.5	37.5	59.5	40.5	39	61	68	32	66	34
Order 2										
Observed	83	17	70	30	42	58	76	24	84	16
Expected	62.5	37.5	59.5	40.5	39	61	68	32	66	34
χ^2 =	*35.8		*9.2		ns		*5.9		*27.8	

Note. All df=1; all * indicates p=0.00

Appendix F

MANOVA on Derived Scores

Multivariate Analysis of Variance on Derived Score Totals

Source	df	error df	F	Sig.
Condition	8	212	4.20	0.00
Gender	4	105	1.96	0.11
Order	4	105	18.24	0.00
Con*Gen	8	212	.76	.64
Con*Ord	8	212	.59	.78
Gen*Ord	4	105	.57	.69
Con*Gen*Ord	8	212	.37	.94

Note. Con = Condition; Gen = Gender; Ord = Order; * = by; SRR = Skin Resistance Responses; BV = Blood Volume; AB = Abdominal Respiration; THOR = Thoracic Respiration

Appendix G

ANOVA on Derived Scores

Univariate Analysis of Variance on Derived Score Totals

Source	Dependent Variable	Sum of Squares	df	Mean Square	F	Sig.
Condition	SRR	159.35	2	79.68	8.24	.00
	BV	76.65	2	38.33	6.14	.00
	AB	17.61	2	8.81	1.50	.23
	THOR	15.65	2	7.83	1.43	.25
Gender	SRR	.30	1	.30	.03	.86
	BV	48.13	1	48.13	7.71	.00
	AB	.21	1	.21	.04	.86
	THOR	.03	1	.03	.01	.94
Order	SRR	403.33	1	403.33	41.73	.00
	BV	240.83	1	240.83	38.59	.00
	AB	11.41	1	11.41	1.94	.17
	THOR	.53	1	.53	.10	.76
Con*Gen	SRR	10.85	2	5.43	.56	.57
	BV	10.62	2	5.31	.85	.43
	AB	23.22	2	11.61	1.97	.14
	THOR	3.62	2	1.81	.33	.72
Con*Ord	SRR	14.82	2	7.41	.77	.47
	BV	11.22	2	5.61	.90	.41
	AB	6.72	2	3.36	.57	.57
	THOR	1.72	2	.86	.16	.86
Gen*Ord	SRR	9.63	1	9.63	1.00	.32
	BV	.53	1	.53	.09	.77
	AB	4.41	1	4.41	.75	.39
	THOR	1.20	1	1.20	.22	.64
Con*Gen*Ord	SRR	6.72	2	3.36	.35	.71
	BV	1.72	2	.86	.14	.87
	AB	.72	2	.36	.06	.94
	THOR	5.56	2	2.78	.51	.61
Error	SRR	1043.80	108	9.67		
	BV	674.00	108	6.24		
	AB	635.70	108	5.89		
	THOR	593.00	108	5.49		
Total	SRR	1746.00	120			
	BV	1064.00	120			
	AB	701.00	120			
	THOR	636.00	120			
Corr. Total	SRR	1648.80	119			
	BV	1063.70	119			
	AB	699.99	119			
	THOR	621.30	119			

Note. Con = Condition; Gen = Gender; Ord = Order; * = by; SRR = Skin Resistance Responses; BV = Blood Volume; AB = Abdominal Respiration THOR = Thoracic Respiration

Appendix H

Duncan Multiple Range Test (DMR) on

Skin Resistance Response (SRR) Means

Duncan's Multiple Range Test on Means for SRR

	.25	-.48	-2.48
.25	---	*.73	*2.73
-.48		---	*2.00
-2.48			---

Note. **"*" indicates significance at the .05 level.**

Appendix I

Duncan Multiple Range Test (DMR) on Blood Volume (BV) Means

Duncan's Multiple Range Test on Means for BV

	.73	.27	-1.15
.73	---	.46	*1.88
.27		---	*1.42
-1.15			---

Note. "*" indicates significance at the .05 level.

Appendix J

ANOVA on Skin Resistance Responses

Analysis of Variance for Skin Resistance Responses (SRR)

Tests of Between-Subjects Effects (SRR)

Source	df	Mean Square	F	Signif.
CON	2	1803.94	1.39	.26
GEN	1	8931.38	6.86	.01
ORD	1	.66	.00	.98
CON*GEN	2	383.17	.29	.75
CON*ORD	2	4936.91	3.79	.03
GEN*ORD	1	889.58	.68	.41
CON*GEN*ORD	2	1571.98	1.21	.30
Error	108	1301.76		

Tests of Within-Subjects Effects (SRR)

Source	df	Mean Square	F	Signif.
BL	2	5285.61	34.27	.00
BL*CON	4	163.66	1.06	.38
BL*GEN	2	33.42	.22	.81
BL*ORD	2	126.23	.82	.44
BL*CON*GEN	4	26.84	.17	.95
BL*CON*ORD	4	29.25	.19	.94
BL*GEN*ORD	2	171.07	1.11	.33
BL*CON*GEN*ORD	4	135.04	.88	.48
Error (BL)	216	154.26		
PR	2	853.53	8.59	.00
PR*CON	4	142.49	1.44	.22
PR*GEN	2	34.34	.35	.71
PR*ORD	2	24.20	.24	.78
PR*CON*GEN	4	106.38	1.07	.37
PR*CON*ORD	4	215.20	2.17	.07
PR*GEN*ORD	2	93.97	.95	.39
PR*CON*GEN*ORD	4	86.66	.87	.48
Error (PR)	216	99.31		
QP	1	915.46	14.34	.00
QP*CON	2	41.91	.66	.52
QP*GEN	1	39.31	.62	.43
QP*ORD	1	324.38	5.08	.03
QP*CON*GEN	2	28.64	.45	.64
QP*CON*ORD	2	191.01	2.99	.05
QP*GEN*ORD	1	3.89	.00	.99
QP*CON*GEN*ORD	2	142.49	2.23	.11
Error (QP)	108	63.85		
BL*PR	4	742.51	9.41	.00
BL*PR*CON	8	48.90	.62	.76

Conitnued on next page

Source	df	Mean Square	F	Signif.
BL*PR*GEN	4	15.86	.20	.94
BL*PR*ORD	4	42.79	.54	.71
BL*PR*CON*GEN	8	53.19	.67	.71
BL*PR*CON*ORD	8	34.67	.44	.90
BL*PR*GEN*ORD	4	54.46	.69	.60
BL*PR*CON* GEN*ORD	8	116.90	1.48	.16
Error (BL*PR)	432	78.91		
BL*QP	2	271.48	7.91	.00
BL*QP*CON	4	29.10	.85	.50
BL*QP*GEN	2	42.19	1.23	.29
BL*QP*ORD	2	3.48	.10	.90
BL*QP*CON*GEN	4	102.31	2.98	.02
BL*QP*CON*ORD	4	4.87	.14	.97
BL*QP*GEN*ORD	2	9.74	.28	.75
BL*QP*CON* GEN*ORD	4	1.35	.04	1.00
Error (BL*QP)	216	34.31		
PR*QP	2	53.27	1.27	.28
PR*QP*CON	4	28.06	.67	.62
PR*QP*GEN	2	10.75	.26	.78
PR*QP*ORD	2	77.35	1.84	.16
PR*QP*CON*GEN	4	49.16	1.17	.33
PR*QP*CON*ORD	4	50.06	1.19	.32
PR*QP*GEN*ORD	2	49.44	1.18	.31
PR*QP*CON* GEN*ORD	4	16.83	.40	.81
Error (PR*QP)	216	42.04		
BL*PR*QP	4	160.85	3.69	.01
BL*PR*QP*CON	8	73.34	1.68	.10
BL*PR*QP*GEN	4	55.30	1.27	.28
BL*PR*QP*ORD	4	30.70	.70	.59
BL*PR*QP* CON*GEN	8	53.63	1.23	.28
BL*PR*QP* CON*ORD	8	49.38	1.13	.34
BL*PR*QP* GEN*ORD	4	26.92	.62	.65
BL*PR*QP*CON *GEN*ORD	8	39.06	.90	.52
Error (BL* PR*QP)	432	43.60		

Note. CON = condition; Gen = gender; Ord = order; * = by (as in CON by Gen interaction); BL = block; PR = pair; QP = question position

Appendix K

Duncan Multiple Range (DMR) Tests on the Skin Resistance Response (SRR) Analysis

Duncan's Multiple Range Test for the Condition by Order Means

	15.13	13.24	9.58	9.45	8.68	8.57
15.13	---	1.89	5.55	5.68	*6.45	*6.56
13.24	---	---	3.66	3.79	4.56	4.67
9.58	---	---	---	.13	.9	1.01
9.45	---	---	---	---	.77	.88
8.68	---	---	---	---	---	.11
8.57	---	---	---	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks

	13.90	9.30	9.12
13.90	---	*4.60	*4.78
9.30	---	---	.18
9.12	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test for Pairs Means

	11.75	10.96	9.60
11.75	---	.79	*2.15
10.96	---	---	*1.36
9.60	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test for Means of Order by Question Position

	11.83	11.02	10.49	9.75
11.83	---	.81	*1.34	*2.08
11.02	---	---	.51	*1.27
10.49	---	---	---	.74
9.75	---	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on means for Question Position by Condition by Order

	15.89	14.36	13.83	12.64	10.26	10.15	9.90	9.69	9.01	8.64	7.67	7.23
15.89	---	1.53	*2.06	*3.25	*5.63	*5.74	*5.99	*6.20	*6.88	*7.25	*8.22	*8.66
14.36	---	---	.53	1.72	*4.10	*4.21	*4.46	*4.67	*5.35	*5.72	*6.69	*7.13
13.83	---	---	---	1.19	*3.57	*3.68	*3.93	*4.14	*4.82	*5.19	*6.16	*6.6
12.64	---	---	---	---	*2.38	*2.49	*2.74	*2.95	*3.63	*4.00	*4.97	*5.41
10.26	---	---	---	---	---	.11	.36	.57	1.25	1.62	*2.59	*3.03
10.15	---	---	---	---	---	---	.25	.46	1.14	1.51	*2.48	*2.92
9.90	---	---	---	---	---	---	---	.21	.89	1.26	*2.23	*2.67
9.69	---	---	---	---	---	---	---	---	.68	1.05	*2.02	*2.46
9.01	---	---	---	---	---	---	---	---	---	.37	1.34	1.78
8.64	---	---	---	---	---	---	---	---	---	---	.97	1.41
7.67	---	---	---	---	---	---	---	---	---	---	---	.44
7.23	---	---	---	---	---	---	---	---	---	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Pairs

	17.04	13.25	11.41	10.29	9.59	9.36	9.36	8.63	8.03
17.04	--	*3.79	*5.63	*6.75	*7.45	*7.68	*7.68	*8.41	*9.01
13.25		--	*1.84	*2.96	*3.66	*3.89	*3.89	*4.62	*5.22
11.41			--	1.12	*1.82	*2.05	*2.05	*2.78	*3.38
10.29				--	.7	.93	.93	1.66	*2.26
9.59					--	.23	.23	.96	1.56
9.36						--	--	.73	1.33
9.36							--	.73	1.33
8.63								--	.6
8.03									--

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Question Position

	15.23	12.57	9.62	9.43	9.18	8.61
15.23	--	*2.66	*5.61	*5.8	*6.05	*6.59
12.57	--	--	*2.95	*3.14	*3.39	*3.96
9.62	--	--	--	.19	.44	*1.01
9.43	--	--	--	--	.25	.82
9.18	--	--	--	--	--	.57
8.61	--	--	--	--	--	--

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender

	18.88	17.39	16.82	15.73	15.50	14.51	14.36	13.83	13.72	13.40
18.88	--	1.49	2.06	*3.15	*3.58	*4.37	*4.52	*5.05	*5.16	*5.48
17.39		--	.57	1.66	2.09	*2.88	*3.03	*3.56	*3.67	*3.99
16.82			--	1.09	1.52	*2.31	*2.46	*2.99	*3.1	*3.42
15.73				--	.43	1.22	1.37	1.9	2.01	2.33
15.30					--	.79	.94	1.47	1.58	1.9
14.51						--	.15	.68	.79	1.11
14.36							--	.53	.64	.96
13.83								--	.11	.43
13.72									--	.32
13.40										--
12.94										
12.54										
12.28										
12.22										
11.56										
11.45										
10.76										
10.68										
9.59										
9.46										
9.06										
9.0										
8.94										
8.75										
8.29										
8.21										
8.08										
7.61										
7.35										
7.21										
7.19										
7.04										
6.79										
5.86										
5.86										
5.19										

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

	12.94	12.54	12.28	12.22	11.56	11.45	10.76	10.68	9.59	9.46
18.88	*5.94	*6.34	*6.6	*6.66	*7.32	*7.43	*8.12	*8.2	*9.29	*9.42
17.39	*4.45	*4.85	*5.11	*5.17	*5.83	*5.94	*6.63	*6.71	*7.18	*7.93
16.82	*3.88	*4.28	*4.54	*4.6	*5.26	*5.37	*6.06	*6.14	*7.23	*7.36
15.73	*2.79	*3.19	*3.45	*3.51	*4.17	*4.28	*4.97	*5.05	*6.14	*6.27
15.30	2.36	2.76	3.02	3.08	3.74	*3.85	*4.54	*4.62	*5.71	*5.84
14.51	1.57	1.97	2.23	2.29	2.95	3.06	3.75	*3.83	*4.92	*5.05
14.36	1.42	1.82	2.08	2.14	2.8	2.91	3.6	*3.68	*4.77	*4.9
13.83	.89	1.29	1.55	1.61	2.27	2.38	3.07	3.15	4.24	*4.37
13.72	.78	1.18	1.44	1.5	2.16	2.27	2.96	3.04	4.13	*4.26
13.40	.46	.86	1.12	1.18	1.84	1.95	2.64	2.72	3.81	*3.94
12.94	--	.4	.66	.72	1.38	1.49	2.18	2.26	3.35	*3.48

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

12.54	--	.26	.32	.98	1.09	1.78	1.86	2.95	3.08
12.28		--	.06	.72	.83	1.52	1.6	2.69	2.82
12.22			--	.66	.77	1.46	1.54	2.63	2.76
11.56				--	.11	.8	.88	1.97	2.1
11.45					--	.69	.77	1.86	1.99
10.76						--	.08	1.17	1.3
10.68							--	1.09	1.22
9.59								--	.13
9.46									--
9.06									
9.0									
8.94									
8.75									
8.29									
8.21									
8.08									
7.61									
7.35									
7.21									
7.19									
7.04									
6.79									
5.86									
5.86									
5.19									

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

	9.06	9.00	8.94	8.75	8.29	8.21	8.08	7.61	7.35
18.88	*9.82	*9.88	*9.94	*10.13	*10.59	*10.67	*10.8	*11.27	*11.53
17.39	*8.33	*8.39	*8.45	*8.64	*9.1	*9.18	*9.31	*9.78	*10.04
16.82	*7.76	*7.82	*7.88	*8.07	*8.53	*8.61	*8.74	*9.21	*9.47
15.73	*6.67	*6.73	*6.79	*6.98	*7.44	*7.52	*7.65	*8.12	*8.38
15.30	*6.24	*6.3	*6.36	*6.55	*7.01	*7.09	*7.22	*7.69	*7.95
14.51	*5.45	*5.51	*5.57	*5.76	*6.22	*6.3	*6.43	*6.9	*7.16
14.36	*5.3	*5.36	*5.42	*5.61	*6.07	*6.15	*6.28	*6.75	*7.01
13.83	*4.77	*4.83	*4.89	*5.08	*5.54	*5.62	*5.75	*6.22	*6.48
13.72	*4.66	*4.72	*4.78	*4.97	*5.43	*5.51	*5.64	*6.11	*6.37
13.40	*4.34	*4.4	*4.46	*4.65	*5.11	*5.19	*5.32	*5.79	*6.05
12.94	*3.88	*3.94	*4.0	*4.19	*4.65	*4.73	*4.86	*5.33	*5.59
12.54	3.48	*3.54	*3.6	*3.79	*4.25	*4.33	*4.46	*4.93	*5.19
12.28	3.22	3.28	3.34	3.53	*3.99	*4.07	*4.2	*4.67	*4.93
12.22	3.16	3.22	3.28	3.47	*3.93	*4.01	*4.14	*4.61	*4.87
11.56	2.5	2.56	2.62	2.81	3.27	3.35	3.48	3.95	*4.21
11.45	2.39	2.45	2.51	2.7	3.16	3.24	3.37	3.84	*4.1
10.76	1.7	1.76	1.82	2.01	2.47	2.55	2.68	3.15	3.41
10.68	1.62	1.68	1.74	1.93	2.39	2.47	2.6	3.07	3.33
9.59	.53	.59	.65	.84	1.3	1.38	1.51	1.98	2.24
9.46	.4	.46	.52	.71	1.17	1.25	1.38	1.85	2.11
9.06	--	.06	.12	.31	.77	.85	.98	1.45	1.71
9.0		--	.06	.25	.71	.79	.92	1.39	1.65
8.94			--	.19	.65	.73	.86	1.33	1.59

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

8.75	--	.46	.54	.67	1.14	1.4
8.29		--	.08	.21	.68	.94
8.21			--	.13	.6	.86
8.08				--	.47	.73
7.61					--	.26
7.35						--
7.21						
7.19						
7.04						
6.79						
5.86						
5.86						
5.19						

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

	7.21	7.19	7.04	6.79	5.86	5.86	5.19
18.88	*11.67	*11.69	*11.84	*12.09	*13.02	*13.02	*13.69
17.39	*10.18	*10.20	*10.35	*10.6	*11.53	*11.53	*12.2
16.82	*9.61	*9.63	*9.78	*10.03	*10.96	*10.96	*11.63
15.73	*8.52	*8.54	*8.69	*8.94	*9.87	*9.87	*10.54
15.30	*8.09	*8.11	*8.26	*8.51	*9.44	*9.44	*10.01
14.51	*7.3	*7.32	*7.47	*7.72	*8.65	*8.65	*9.32
14.36	*7.15	*7.17	*7.32	*7.57	*8.5	*8.5	*9.17
13.83	*6.62	*6.64	*6.79	*7.04	*7.97	*7.97	*8.64
13.72	*6.51	*6.53	*6.68	*6.93	*7.86	*7.86	*8.53
13.40	*6.19	*6.21	*6.36	*6.61	*7.54	*7.54	*8.21
12.94	*5.73	*5.75	*5.9	*6.15	*7.08	*7.08	*7.75
12.54	*5.33	*5.35	*5.5	*5.75	*6.68	*6.68	*7.35
12.28	*5.07	*5.09	*5.24	*5.49	*6.42	*6.42	*7.09
12.22	*5.01	*5.03	*5.18	*5.43	*6.36	*6.36	*7.03
11.56	*4.35	*4.37	*4.52	*4.77	*5.7	*5.7	*6.37
11.45	*4.24	*4.26	*4.41	*4.66	*5.59	*5.59	*6.26
10.76	3.55	3.57	3.72	*3.97	*4.9	*4.9	*5.57
10.68	3.47	3.49	3.64	*3.89	*4.82	*4.82	*5.49
9.59	2.38	2.4	2.55	2.8	3.73	3.73	*4.4
9.46	2.25	2.27	2.42	2.67	3.6	3.6	*4.27
9.06	1.85	1.87	2.02	2.27	3.2	3.2	3.87
9.0	1.79	1.81	1.96	2.21	3.14	3.14	3.81
8.94	1.73	1.75	1.9	2.15	3.08	3.08	3.75
8.75	1.54	1.56	1.71	1.96	2.89	2.89	3.56
8.29	1.08	1.1	1.25	1.5	2.43	2.43	3.1
8.21	1.0	1.02	1.17	1.42	2.35	2.35	3.02
8.08	.87	.89	1.04	1.29	2.22	2.22	2.89
7.61	.4	.42	.57	.82	1.75	1.75	2.42
7.35	.14	.16	.31	.56	1.49	1.49	2.16
7.21	--	.02	.17	.42	1.35	1.35	2.02
7.19		--	.15	.4	1.33	1.33	2.0
7.04			--	.25	1.18	1.18	1.85
6.79				--	.93	.93	1.6
5.86					--	--	.67
5.86						--	.67
5.19							--

Duncan's Multiple Range Test on Means for Blocks by Question Position by Condition by Gender (continued)

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Pairs by Question Position

	19.66	14.42	13.62	12.88	12.40	10.72	10.42	10.02	9.86
19.66	--	*5.24	*6.04	*6.78	*7.26	*8.94	*9.24	*9.46	*9.8
14.42		--	.08	1.54	*2.02	*3.70	*4.00	*4.40	*4.56
13.62			--	.74	1.22	*2.90	*3.20	*3.60	*3.76
12.88				--	.48	*2.16	*2.46	*2.86	*3.02
12.40					--	*1.68	*1.98	*2.38	*2.54
10.72						--	.30	.70	.86
10.42							--	.40	.56
10.02								--	.16
9.86									--
9.85									
9.66									
9.32									
9.06									
9.00									
8.53									
8.26									
8.24									
7.82									

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Blocks by Pairs by Question Position (continued)

	9.85	9.66	9.32	9.06	9.00	8.53	8.26	8.24	7.82
19.66	*9.81	*10.00	*10.34	*10.60	*10.66	*11.13	*11.40	*11.42	*11.84
14.42	*4.57	*4.76	*5.10	*5.36	*5.42	*5.89	*6.16	*6.18	*6.60
13.62	*3.77	*3.96	*4.30	*4.56	*4.62	*5.09	*5.36	*5.38	*5.80
12.88	*3.03	*3.22	*3.56	*3.82	*3.88	*4.35	*4.62	*4.64	*5.06
12.40	*2.55	*2.74	*3.08	*3.34	*3.40	*3.87	*4.14	*4.16	*4.58
10.72	.87	1.06	1.40	1.66	1.72	*2.19	*2.46	*2.48	*2.90
10.42	.57	.76	1.10	1.36	1.42	1.89	*2.16	*2.18	*2.60
10.02	.17	.36	.70	.96	1.02	1.49	1.76	1.78	*2.20
9.86	.01	.20	.54	.80	.86	1.33	1.60	1.62	*2.04
9.85	--	.19	.53	.79	.85	1.32	1.59	1.61	*2.03
9.66		--	.34	.60	.66	1.13	1.40	1.42	1.84
9.32			--	.26	.32	.79	1.06	1.08	1.50
9.06				--	.06	.53	.80	.82	1.24
9.00					--	.47	.74	.76	1.18
8.53						--	.27	.29	.71
8.26							--	.02	.44
8.24								--	.42
7.82									--

Note. "*" indicates significance at the .05 level.

Appendix L

ANOVA on Thoracic Responses

Analysis of Variance for Thoracic Respiration

Tests of Between-Subjects Effects (THR)

Source	df	Mean Square	F	Signif.
CON	2	7.91	1.53	.22
GEN	1	90.88	17.57	.00
ORD	1	2.87	.56	.46
CON*GEN	2	1.14	.22	.80
CON*ORD	2	14.02	2.71	.07
GEN*ORD	1	11.39	2.20	.14
CON*GEN*ORD	2	2.70	.52	.60
Error	108	5.17		

Tests of Within-Subjects Effects (THR)

Source	df	Mean Square	F	Signif.
BL	2	2.86	12.72	.00
BL*ORD	2	.38	1.70	.19
BL*CON	4	.12	.52	.72
BL*GEN	2	.66	2.94	.06
BL*CON*ORD	4	.57	2.54	.04
BL*GEN*ORD	2	.23	1.04	.36
BL*CON*GEN	4	.02	.09	.99
BL*CON*GEN*ORD	4	.27	1.20	.31
Error (BL)	216	.23		
PR	2	.11	1.65	.20
PR*ORD	2	.03	.53	.59
PR*CON	4	.09	1.34	.26
PR*GEN	2	.00	.06	.94
PR*CON*ORD	4	.04	.68	.61
PR*CON*GEN	4	.05	.71	.58
PR*GEN*ORD	2	.04	.55	.58
PR*CON*GEN*ORD	4	.09	1.36	.25
Error (PR)	216	.06		
QP	1	.00	.00	.97
QP*CON	2	.02	.38	.69
QP*GEN	1	.00	.04	.85
QP*ORD	1	.04	.72	.40
QP*CON*GEN	2	.15	2.50	.09
QP*CON*ORD	2	.06	1.10	.34

Source	df	Mean Square	F	Signif.
QP*GEN*ORD	1	.00	.03	.86
QP*CON*GEN*ORD	2	.01	.16	.86

Error (QP)	108	.06		
BL*PR	4	.09	1.36	.25
BL*PR*CON	8	.09	1.37	.21
BL*PR*GEN	4	.03	.50	.74
BL*PR*ORD	4	.09	1.38	.24
BL*PR*CON*GEN	8	.08	1.19	.30
BL*PR*CON*ORD	8	.06	.88	.54
BL*PR*GEN*ORD	4	.02	.24	.92
BL*PR*CON* GEN*ORD	8	.01	1.47	.17
Error (BL*PR)	432	.07		
BL*QP	2	.02	.57	.57
BL*QP*CON	4	.04	1.26	.29
BL*QP*GEN	2	.02.	0.56	.57
BL*QP*ORD	2	.05	1.53	.22
BL*QP*CON*GEN	4	.03	.83	.51
BL*QP*CON*ORD	4	.05	1.54	.19
BL*QP*GEN*ORD	2	.01	.19	.83
BL*QP*CON* GEN*ORD	4	.04	1.02	.40
Error (BL*QP)	216	.04		
PR*QP	2	.03	.62	.54
PR*QP*CON	4	.02	.37	.83
PR*QP*GEN	2	.02	.40	.67
PR*QP*ORD	2	.04	.83	.44
PR*QP*CON*GEN	4	.17	3.18	.02
PR*QP*CON*ORD	4	.05	.87	.48
PR*QP*GEN*ORD	2	.06	1.10	.34
PR*QP*CON* GEN*ORD	4	.05	1.04	.39
Error (PR*QP)	216	.05		
BL*PR*QP	4	.04	1.05	.38
BL*PR*QP*CON	8	.03	.87	.54
BL*PR*QP*GEN	4	.02	.61	.66
BL*PR*QP*ORD	4	.06	1.52	.20
BL*PR*QP* CON*GEN	8	.06	1.60	.12
BL*PR*QP* CON*ORD	8	.03	.71	.68
BL*PR*QP* GEN*ORD	4	.02	.54	.71
BL*PR*QP*CON *GEN*ORD	8	.08	1.95	.05
Error (BL*PR*QP)	432	.04		

Note. CON = condition; Gen = gender; Ord = order; * = by (as in CON by Gen interaction); BL = block; PR = pair; QP = question position

Appendix M

Duncan Multiple Range (DMR) Tests on the Thoracic (TH) Analysis

Duncan's Multiple Range Test on Means for Blocks

	1.83	1.75	1.70
1.83	---	*.08	*.13
1.75	---	---	.05
1.70	---	---	---

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Condition by Order by Blocks

	1.98	1.96	1.92	1.88	1.87	1.85	1.84	1.79	1.79	1.78	1.76
1.98	--	.02	.06	.10	.11	*.13	*.14	*.19	*.19	*.20	*.22
1.96		--	.04	.08	.09	.11	.12	*.17	*.17	*.18	*.20
1.92			--	.04	.05	.07	.08	*.13	*.13	*.14	*.16
1.88				--	.01	.03	.04	.09	.09	.10	.12
1.87					--	.02	.03	.08	.08	.09	.11
1.85						--	.01	.06	.06	.07	.09
1.84							--	.05	.05	.06	.08
1.79								--	--	.01	.03
1.79									--	.01	.03
1.78										--	.02
1.76											--
1.75											
1.74											
1.73											
1.69											
1.51											
1.42											
1.40											

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Condition by Order by Blocks (continued)

	1.75	1.74	1.73	1.69	1.51	1.42	1.40
1.98	*.23	*.24	*.25	*.29	*.47	*.56	*.58
1.96	*.21	*.22	*.23	*.27	*.45	*.54	*.56
1.92	*.17	*.18	*.19	*.23	*.41	*.50	*.52
1.88	*.13	*.14	*.15	*.19	*.37	*.46	*.48
1.87	.12	*.13	*.14	*.18	*.36	*.45	*.47
1.85	.10	.11	.12	*.16	*.34	*.43	*.45
1.84	.09	.10	.11	*.15	*.33	*.42	*.44
1.79	.04	.05	.06	.10	*.28	*.37	*.39
1.79	.04	.05	.06	.10	*.28	*.37	*.39
1.78	.03	.04	.05	.09	*.27	*.36	*.38
1.76	.01	.02	.03	.07	*.25	*.34	*.36
1.75	--	.01	.02	.06	*.24	*.33	*.35
1.74		--	.01	.05	*.23	*.32	*.34
1.73			--	.04	*.22	*.31	*.33
1.69				--	*.18	*.27	*.29
1.51					--	.09	.11
1.42						--	.02
1.40							--

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Pairs by Question Position By Condition By Gender

	2.12	2.12	2.11	2.09	2.08	2.05	2.03	2.01	1.99	1.95	1.93	1.92
2.12	--	--	.01	.03	.04	.07	.09	*.11	*.13	*.17	*.19	*.20
2.12		--	.01	.03	.04	.07	.09	*.11	*.13	*.17	*.19	*.20
2.11			--	.02	.03	.06	.08	*.10	*.12	*.16	*.18	*.19
2.09				--	.01	.04	.06	.08	*.10	*.14	*.16	*.17
2.08					--	.03	.05	.07	.09	*.13	*.15	*.16
2.05						--	.02	.04	.06	*.10	*.12	*.13
2.03							--	.02	.04	.08	*.10	*.11
2.01								--	.02	.06	.08	.09
1.99									--	.04	.06	.07
1.95										--	.02	.03
1.93											--	.01
1.92												--
1.84												
1.84												
1.83												
1.82												
1.82												
1.79												
1.66												
1.65												
1.63												
1.63												
1.60												
1.60												
1.59												
1.59												
1.59												
1.57												
1.56												
1.56												
1.48												
1.47												
1.46												
1.44												
1.44												
1.44												

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Pairs by Question Position By Condition By Gender (continued)

	1.84	1.84	1.83	1.82	1.82	1.79	1.66	1.65	1.63	1.63	1.60	1.60
2.12	*.28	*.28	*.29	*.30	*.30	*.33	*.46	*.47	*.49	*.49	*.52	*.52
2.12	*.28	*.28	*.29	*.30	*.30	*.33	*.46	*.47	*.49	*.49	*.52	*.52
2.11	*.27	*.27	*.28	*.29	*.29	*.32	*.45	*.46	*.48	*.48	*.51	*.51
2.09	*.25	*.25	*.26	*.27	*.27	*.30	*.43	*.44	*.46	*.46	*.49	*.49
2.08	*.24	*.24	*.25	*.26	*.26	*.29	*.42	*.43	*.45	*.45	*.48	*.48
2.05	*.21	*.21	*.22	*.23	*.23	*.26	*.39	*.40	*.42	*.42	*.45	*.45
2.03	*.19	*.19	*.20	*.21	*.21	*.24	*.37	*.38	*.40	*.40	*.43	*.43
2.01	*.17	*.17	*.18	*.19	*.19	*.22	*.35	*.36	*.38	*.38	*.41	*.41
1.99	*.15	*.15	*.16	*.17	*.17	*.20	*.33	*.34	*.36	*.36	*.39	*.39
1.95	*.11	*.11	*.12	*.13	*.13	*.16	*.29	*.30	*.32	*.32	*.35	*.35
1.93	*.09	*.09	*.10	*.11	*.11	*.14	*.27	*.28	*.30	*.30	*.33	*.33
1.92	.08	.08	*.09	*.10	*.10	*.13	*.26	*.27	*.29	*.29	*.32	*.32
1.84	--	--	.01	.02	.02	.05	*.18	*.19	*.21	*.21	*.24	*.24

Duncan's Multiple Range Test on Means for Pairs by Question Position By Condition By Gender (continued)

1.84	--	.01	.02	.02	.05	*.18	*.19	*.21	*.21	*.24	*.24
1.83		--	.01	.01	.04	*.17	*.18	*.20	*.20	*.23	*.23
1.82			--	--	.03	*.16	*.17	*.19	*.19	*.22	*.22
1.82				--	.03	*.16	*.17	*.19	*.19	*.22	*.22
1.79					--	*.13	*.14	*.16	*.16	*.19	*.19
1.66						--	.01	.03	.03	.06	.06
1.65							--	.02	.02	.05	.05
1.63								--	--	.03	.03
1.63									--	.03	.03
1.60										--	--
1.60											--
1.59											
1.59											
1.59											
1.57											
1.56											
1.56											
1.48											
1.47											
1.46											
1.44											
1.44											
1.44											

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Pairs by Question Position By Condition By Gender (continued)

	1.59	1.59	1.59	1.57	1.56	1.56	1.48	1.47	1.46	1.44	1.44	1.44
2.12	*.53	*.53	*.53	*.55	*.56	*.56	*.64	*.65	*.66	*.68	*.68	*.68
2.12	*.53	*.53	*.53	*.55	*.56	*.56	*.64	*.65	*.66	*.68	*.68	*.68
2.11	*.52	*.52	*.52	*.54	*.55	*.55	*.63	*.64	*.65	*.67	*.67	*.67
2.09	*.50	*.50	*.50	*.52	*.53	*.53	*.61	*.62	*.63	*.65	*.65	*.65
2.08	*.49	*.49	*.49	*.51	*.52	*.52	*.60	*.61	*.62	*.64	*.64	*.64
2.05	*.46	*.46	*.46	*.48	*.49	*.49	*.57	*.58	*.59	*.61	*.61	*.61
2.03	*.44	*.44	*.44	*.46	*.47	*.47	*.55	*.56	*.57	*.59	*.59	*.59
2.01	*.42	*.42	*.42	*.44	*.45	*.45	*.53	*.54	*.55	*.57	*.57	*.57
1.99	*.40	*.40	*.40	*.42	*.43	*.43	*.51	*.52	*.53	*.55	*.55	*.55
1.95	*.36	*.36	*.36	*.38	*.39	*.39	*.47	*.48	*.49	*.51	*.51	*.51
1.93	*.34	*.34	*.34	*.36	*.37	*.37	*.45	*.46	*.47	*.49	*.49	*.49
1.92	*.33	*.33	*.33	*.35	*.36	*.36	*.44	*.45	*.46	*.48	*.48	*.48
1.84	*.25	*.25	*.25	*.27	*.28	*.28	*.36	*.37	*.38	*.40	*.40	*.40
1.84	*.25	*.25	*.25	*.27	*.28	*.28	*.36	*.37	*.38	*.40	*.40	*.40
1.83	*.24	*.24	*.24	*.26	*.27	*.27	*.35	*.36	*.37	*.39	*.39	*.39
1.82	*.23	*.23	*.23	*.25	*.26	*.26	*.34	*.35	*.36	*.38	*.38	*.38
1.82	*.23	*.23	*.23	*.25	*.26	*.26	*.34	*.35	*.36	*.38	*.38	*.38
1.79	*.20	*.20	*.20	*.22	*.23	*.23	*.31	*.32	*.33	*.35	*.35	*.35
1.66	.07	.07	.07	.09	*.10	*.10	*.18	*.19	*.20	*.22	*.22	*.22
1.65	.06	.06	.06	.08	.09	.09	*.17	*.18	*.19	*.21	*.21	*.21
1.63	.04	.04	.04	.06	.07	.07	*.15	*.16	*.17	*.19	*.19	*.19
1.63	.04	.04	.04	.06	.07	.07	*.15	*.16	*.17	*.19	*.19	*.19
1.60	.01	.01	.01	.03	.04	.04	*.12	*.13	*.14	*.16	*.16	*.16
1.60	.01	.01	.01	.03	.04	.04	*.12	*.13	*.14	*.16	*.16	*.16
1.59	--	--	--	.02	.03	.03	*.11	*.12	*.13	*.15	*.15	*.15
1.59		--	--	.02	.03	.03	*.11	*.12	*.13	*.15	*.15	*.15

Duncan's Multiple Range Test on Means for Pairs by Question Position By Condition By Gender (continued)

1.59	--	.02	.03	.03	*.11	*.12	*.13	*.15	*.15	*.15
1.57		--	.01	.01	*.09	*.10	*.11	*.13	*.13	*.13
1.56			--	--	.08	*.09	*.10	*.12	*.12	*.12
1.56				--	.08	*.09	*.10	*.12	*.12	*.12
1.48					--	.01	.02	.04	.04	.04
1.47						--	.01	.03	.03	.03
1.46							--	.02	.02	.02
1.44								--	--	--
1.44									--	--
1.44										--

Note. "*" indicates significance at the .05 level.

Appendix N

ANOVA on Abdominal Responses

Analysis of Variance for Abdominal Respiration

Tests of Between-Subjects Effects

Source	df	Mean Square	F	Signif.
CON	2	2.73	.23	.80
GEN	1	401.28	3.64	.00
ORD	1	1.12	.09	.76
CON*GEN	2	2.18	.18	.83
CON*ORD	2	4.02	.34	.72
GEN*ORD	1	12.45	1.04	.31
CON*GEN*ORD	2	5.25	.44	.65
Error	108	11.93		

Tests of Within-Subjects Effects (Abdominal)

Source	df	Mean Square	F	Signif.
CON	2	2.73	.23	.80
BL	2	6.09	4.79	.01
BL*CON	4	.62	.49	.75
BL*GEN	2	3.30	2.59	.08
BL*ORD	2	3.09	2.43	.09
BL*CON*GEN	4	.47	.37	.83
BL*CON*ORD	4	1.22	.96	.43
BL*GEN*ORD	2	1.98	1.56	.21
BL*CON*GEN*ORD	4	1.21	.95	.44
Error (BL)	216	1.27		
PR	2	1.06	4.05	.02
PR*CON	4	.44	1.68	.16
PR*GEN	2	.79	3.04	.05
PR*ORD	2	.36	1.38	.26
PR*CON*GEN	4	.34	1.29	.28
PR*CON*ORD	4	.14	.54	.71
PR*GEN*ORD	2	.43	1.62	.20
PR*CON*GEN*ORD	4	.49	1.87	.12
Error (PR)	216	.26		
QP	1	.48	.38	.54
QP*CON	2	1.12	.90	.41
QP*GEN	1	.23	.19	.67
QP*ORD	1	1.25	1.00	.32
QP*CON*GEN	2	1.38	1.11	.34
QP*CON*ORD	2	2.03	1.63	.20
QP*GEN*ORD	1	.72	.58	.45
QP*CON*GEN*ORD	2	1.88	1.51	.23
Error (QP)	108	1.25		

BL*PR	4	.63	1.29	.28
BL*PR*CON	8	.74	1.53	.14
BL*PR*GEN	4	.49	1.01	.40
BL*PR*ORD	4	.65	1.35	.25
BL*PR*CON*GEN	8	.85	1.75	.08
BL*PR*CON*ORD	8	.50	1.02	.42
BL*PR*GEN*ORD	4	.64	1.32	.26
BL*PR*CON* GEN*ORD	8	.49	1.00	.44
Error (BL*PR)	432	.49		
BL*QP	2	1.59	1.46	.24
BL*QP*CON	4	1.14	1.04	.39
BL*QP*GEN	2	1.86	1.71	.18
BL*QP*ORD	2	1.09	1.00	.37
BL*QP*CON*GEN	4	.84	.77	.54
BL*QP*CON*ORD	4	1.19	1.09	.36
BL*QP*GEN*ORD	2	1.19	1.09	.34
BL*QP*CON* GEN*ORD	4	.77	.71	.59
Error (BL*QP)	216	1.09		
PR*QP	2	.08	.61	.55
PR*QP*CON	4	.29	2.27	.06
PR*QP*GEN	2	.01	.07	.94
PR*QP*ORD	2	.02	.14	.87
PR*QP*CON*GEN	4	.31	2.39	.05
PR*QP*CON*ORD	4	.21	1.62	.17
PR*QP*GEN*ORD	2	.13	.99	.37
PR*QP*CON* GEN*ORD	4	.15	1.18	.32
Error (PR*QP)	216	.13		
BL*PR*QP	4	.71	1.64	.16
BL*PR*QP*CON	8	.42	.98	.45
BL*PR*QP*GEN	4	.41	.95	.44
BL*PR*QP*ORD	4	.33	.75	.56
BL*PR*QP* CON*GEN	8	.31	.71	.69
BL*PR*QP* CON*ORD	8	.24	.55	.82
BL*PR*QP* GEN*ORD	4	.55	1.27	.28
BL*PR*QP*CON *GEN*ORD	8	.37	.84	.56
Error(BL*PR*QP)	432	.43		

Note. CON = condition; Gen = gender; Ord = order; * = by (as in CON by Gen interaction); BL = block; PR = pair; QP = question position

Appendix O

Duncan Multiple Range (DMR) Tests on the Abdominal (Ab) Analysis

Duncan's Multiple Range Test on Means for Blocks

	2.02	1.86	1.85
2.02	--	*.16	*.17
1.86		--	.01
1.85			--

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Pairs

	1.95	1.89	1.88
1.95	--	*.06	*.07
1.89		--	.01
1.88			--

Note. "*" indicates significance at the .05 level.

Duncan's Multiple Range Test on Means for Gender by Pairs

	2.42	2.30	2.30	1.49	1.48	1.46
2.42	--	*.12	*.12	*.93	*.94	*.96
2.30		--	--	*.81	*.82	*.84
2.30			--	*.81	*.82	*.84
1.49				--	.01	.03
1.48					--	.02
1.46						--

Note. "*" indicates significance at the .05 level.