OCULAR-MOTOR METHODS FOR DETECTING DECEPTION: DIRECT VERSUS INDIRECT INTERROGATION

by

Pooja Patnaik

`

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Educational Psychology

The University of Utah

August 2013

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Pooja Patnaik**

has been approved by the following supervisory committee members:

John C. Kircher	, Chair	**4/17/2013**
		Date Approved
Anne E. Cook	, Member	**4/17/2013**
		Date Approved
Dan J. Woltz	, Member	**4/17/2013**
		Date Approved

and by **Elaine Clark** , Chair of

the Department of **Educational Psychology**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

The present mock crime study investigated whether the accuracy of an ocular-motor deception test (ODT) that directly asks if the person committed illicit acts differs from the accuracy of an ODT that indirectly asks if the person provided false answers on a questionnaire about those illicit activities. Guilt, statement type, relevant issue, and completion of a pre-ODT questionnaire were manipulated in the present study to assess their effects on ocular-motor and behavioral measures of deception. Half the subjects were guilty of taking $20 from a secretary's wallet, and the other half were innocent. All subjects were told that some subjects took an exam from a professor's office, but in actuality, no one committed that crime. Three-fourths of the subjects completed a pre-ODT questionnaire that asked about their involvement in the crimes. Subjects answered 48 true/false items five times while their eye movements and pupil diameters were recorded. Half of the guilty and innocent subjects answered test items that directly asked if they committed the thefts. The remaining subjects were asked if they falsified information about the crimes on the pre-ODT questionnaire.

Guilty subjects showed the largest pupil diameter while reading the cash items. For direct items, a discriminant function of four ocular-motor measures correctly classified 95% of innocent subjects and 83% of guilty subjects. For indirect items, the discriminant analysis of three ocular-motor measures correctly classified 79% of innocent

subjects and 58% of guilty subjects. Results suggest that indirect test items are less effective than direct ones.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

ACKNOWLEDGMENTS

This project would not have been possible without the support and encouragement of several people. I owe my deepest gratitude to my advisor, John Kircher, for his invaluable support and guidance. John spent countless patient hours with me discussing the project design, procedures, results, and many drafts of my thesis. I also am extremely grateful to John for providing all of the financial support for this project. I feel very lucky that I have the opportunity to work with John.

I also would like to thank my committee members for their insights and support throughout the many stages of this project.

Special thanks to those who played the role of the secretary in my project; thank you Wei Wei for your numerous hours, Ali Pappas, Sarah Goldman, Jackie Barco, and Susan Stephenson.

Finally, I would like to thank my family and friends for supporting me through all my highs and lows of this project; I could not have completed this thesis without their love and support.

CHAPTER I

INTRODUCTION

Polygraph tests are widely used in law enforcement, courts, pre-employment screening, and the national security systems of some countries. The National Research Council published a report (NRC, 2003) that questioned the validity of the polygraph for use as a security screening tool. The National Research Council called for alternatives to the polygraph because it relies on emotional responses to test stimuli, and emotional reactions are not specific to deception. In some cases, emotional responses may be incorrectly interpreted as instances of deception.

In response, Cook et al. (2012) developed a cognition-based deception test that may be used in screening contexts. In the first experiment described by Cook et al., subjects committed one of two mock-crimes or were innocent of both crimes. One group of guilty subjects stole $20 from a secretary's purse. Another group of guilty subjects downloaded credit card information from a graduate student's computer. All subjects were fitted with an eye tracker and answered true/false items on a computer screen. There were three categories of items: items that pertained to the theft of the $20, items that pertained to the theft of the credit card information, and neutral items. Cook et al. included questions about two crimes to simulate a security screening situation because screening tests often cover multiple issues, and a person may or may not be deceptive

about one or more issues on the test. Dependent measures included number of fixations made on an item, first pass duration (time spent reading an item), second pass duration (time spent rereading an item), and pupil diameter (PD). Subjects showed larger PD on items pertaining to the crime they committed, suggesting that they engaged in effortful processing of those items. However, for number of fixations, first pass duration, and second pass duration, guilty subjects made more fixations and spent more time reading and rereading items that pertained to the crime they did *not* commit. Interestingly, the same pattern was seen for both groups of guilty subjects.

The second experiment described by Cook et al. (2012) was Andrea Webb's dissertation experiment (Webb, 2008). Half of her subjects stole $20 from a secretary's purse and the other half were innocent and did not steal anything. All subjects were told that some subjects had to download an exam from the professor's computer, but in actuality, no one committed that crime. Subjects answered 48 true/false items while their eye movements and pupil diameter were recorded. One third of the questions pertained to the theft of the $20, one third pertained to the theft of an exam, and the remaining items were neutral. Guilty subjects showed the largest PD change to the items concerning a crime they had committed, followed by items about a crime they did not commit, and then neutral items. Guilty subjects also took less time to respond, made fewer fixations, and did less reading and rereading of items concerning the crime they committed than to items concerning the crime they did not commit and neutral items. Innocent subjects showed greater PD change to the crime-related items than to the neutral items and tended to show less difference in ocular-motor and behavioral responses to the three item types than did guilty subjects.

Most of the ocular-motor and behavioral measures discriminated between guilty and innocent groups. Accuracy of classifications exceeded 80% for both guilty and innocent groups. The findings in the experiments discussed by Cook et al. (2012) were consistent and supported the idea that ocular-motor measures from a reading task can be used to distinguish between guilty and innocent subjects.

Pupil Diameter

Pupil diameter (PD) is of interest in the present study. Research has shown that changes in PD are reliable and valid indicators of cognitive effort and emotional arousal (Loewenfeld, 1999), and most theories of deception detection posit that deception is cognitively more demanding than telling the truth (Johnson, Barnhardt, & Zhu, 2005; Kircher, 1981; Steller, 1989; Vrij, Fisher, Mann, & Leal, 2006). Lying can be more cognitively demanding for several reasons. First, creating a convincing lie itself may be cognitively demanding. Liars need to fabricate a story and keep track of it in order to maintain consistency. Deception is cognitively challenging because it requires two processes; first, subjects must inhibit the truthful response, and second, they must formulate a deceptive response. In the context of a polygraph examination, Kircher (1981) and Stellar (1989) suggested that deceptive individuals attempt to monitor their internal physiological responses to test items. Monitoring internal states is a cognitive process that demands resources and produces autonomic and somatic changes that are characteristic of deceptive individuals.

Research on PD and lie detection generally has found that deception is associated with greater increases in pupil size than telling the truth. Cook et al. (2012) found PD to

be a reliable indicator of deception, which is consistent with the idea that guilty subjects exerted more cognitive effort when they lied than did truthful subjects. Dionisio, Granholm, Hillix, and Perrine (2001) measured PD while subjects made truthful and deceptive responses and the largest increase in PD was found when subjects were deceptive. Bradley and Janisse (1979) and Janisse and Bradley (1980) measured PD as subjects answered truthfully or deceptively to questions regarding a numbered card they had chosen. PD discriminated between the truthful and deceptive groups. Subsequently, Bradley and Janisse (1981) conducted a mock-crime experiment in which guilty subjects stole a dollar and hid it. Innocent subjects did not steal anything. Subjects were given two polygraph tests: a concealed information test and a comparison question test. PD discriminated between the guilty and innocent subjects for the concealed information test but not for the comparison question test. In contrast to the latter result, Webb, Honts, Kircher, Bernhardt, and Cook (2008) administered a comparison question test and found that PD discriminated between guilty and innocent subjects. PD discriminated as well as skin conductance and better than cardiovascular and respiration measures. Lubow and Fein (1996) also conducted a mock-crime experiment and monitored PD while subjects completed a concealed information test. Stimuli in the concealed information test were pictures rather than the auditory questions used by Bradley and Janisse. As with previous work, PD discriminated between the guilty and innocent subjects.

The reading behaviors observed by Cook et al. (2012) were not consistent with basic research on reading. In the psychology of reading literature, increases in PD, frequent fixations, and long reading times are viewed as indications that subjects had difficulty processing those items (Rayner, 1998; Rayner, Chace, Slattery, & Ashby,

2006). If deception is more difficult than being truthful, then it should be associated with increased PD and longer reading times. As expected, deception was associated with the greatest increases in PD. However, in comparison to truthful answers, deception also was characterized by fewer fixations and shorter reading and rereading times. Although the pupil data from the Cook et al. (2012) experiments were consistent with the reading literature, the fixation and response time measures were not.

Recent work suggests that the observed effects on pupil size are mediated by both mental effort and emotional arousal. The emotional component may contribute to discrimination despite evidence of habituation (Kuhlman et al., 2011). In response to the National Research Council's call for the development of new security screening techniques, PD seems to be a promising measure for detecting deception because it has been shown to discriminate between guilty and innocent subjects and may not rely exclusively on emotional responses.

Motivation for the present study comes from an attempt to replicate and extend the results of Webb's (2008) dissertation in a subsequent lab study (Hacker, Cook, & Kircher, 2010). In that study, some subjects lied about changing grades on their academic transcript, some subjects lied about their driver's license, some subjects lied about both issues, and some subjects were truthful to both issues. In contrast to Webb (2008), subjects in the Hacker et al. study were not randomly assigned to guilty and innocent treatment conditions. Rather, the subjects chose whether to be guilty or innocent of the mock crime(s). Prior to the ODT, subjects completed a pre-ODT simulated application for a scholarship that asked for the date of their driver's license and their grades. Some of the statements on the ODT then asked if the subject had falsified

the date of his or her driver's license. Other statements on the ODT asked if the subject had falsified his or her grades on the scholarship application form. The ODT statements did not ask directly if the subject had changed his or her grades. Overall, only 60% of the guilty and 78% of the innocent subjects were classified correctly. The relatively low accuracy in the Hacker et al. study, especially for guilty subjects, may have been due to chance. Hacker et al. attempted to classify subjects into four groups rather than only two or three. Therefore, the chance probability of a correct classification was lower in the Hacker et al. study than in the prior two experiments. The low accuracy in the Hacker et al. study may have been due to the nonrandom procedures for assigning subjects to groups. Subjects more likely to defeat the ODT may have self-selected into the guilty treatment condition. Alternatively, the difference in accuracy rates may have been due to the nature of the illicit activities, since the Webb experiment stole $20 from a secretary, and subjects in the Hacker et al. study used a secretary's computer to change their grades. Finally, the difference in accuracy rates for the two studies may have been due to the use of indirect statements concerning subjects' answers on a pre-ODT questionnaire by Hacker et al. rather than direct questions about subjects' illicit behavior in Cook et al. The present study will explore the latter possibility. The present study is designed to test if high accuracy depends on the use of direct questions about the crime rather than questions about the subject's answers on a questionnaire. Outcomes obtained from subjects who responded to direct statements on the ODT about illicit activities were compared to outcomes from subjects who responded to indirect statements about their falsification of answers on a pre-ODT questionnaire about those illicit activities.

The results from the proposed study have significant implications for use in field settings. For example, in an actual employment screening situation, applicants are asked to complete an application or questionnaire that asks about relevant work experience, if they have been convicted of a crime, and other types of questions. If high levels of accuracy can be achieved with indirect questions that ask if the examinee falsified information on a pre-ODT questionnaire, then a standardized ODT could be developed for many different applications. The questions on the ODT would test if the subject provided false information on the pre-ODT questionnaire, and only the pre-ODT questionnaire would change from one application to the next. On the other hand, if the accuracy of the ODT depends on the use of direct questions about specific illicit behaviors, each new application would require a new set of ODT items, and it would be more difficult and costly to implement this type of ODT in the field.

In addition to the manipulation of guilt and question type (indirect or direct), a questionnaire was administered to a portion of the subjects to determine if pretest questioning moderates differences between guilty and innocent subjects on ocular-motor measures. The administration of a pre-ODT questionnaire about the illicit activities subsequently covered on the ODT could affect the diagnostic validity of ocular-motor measures. For example, guilty subjects who completed a pre-ODT questionnaire may become habituated to questions about the crimes and be less affected by those questions on the subsequent ODT. Alternatively, guilty subjects who completed a pre-ODT questionnaire may become sensitized to relevant issues and react more strongly during the ODT than if they had not been asked about the crimes before the ODT. Prior exposure to the relevant issues could habituate or sensitize innocent subjects as well. The

present study compared outcomes obtained from subjects who did or did not complete a pre-ODT questionnaire that asked about the subject's involvement in the crimes under investigation.

Finally, a Behavioral Approach System (BAS) and Behavioral Inhibition System (BIS) questionnaire was administered. The BAS is believed to mediate appetitive motives, where the goal is to move towards something that is desired. The BIS is said to mediate aversive motives, where the goal is to move away from something unpleasant (Carver & White, 1994). The BAS/BIS questionnaire has three BAS scales: BAS Drive, BAS Fun Seeking, and BAS Reward Responsiveness. Typically, the scales are not combined because they focus on different aspects of incentive sensitivity. People with high BIS sensitivity should be especially responsive to punishment cues and should experience greater anxiety in situations with cues of impending punishment compared to people with lower BIS sensitivity. Carver and White (1994) concluded that their research provides support for the idea that the BAS/BIS scales reflect individual differences in the sensitivity of the presumed underlying neurophysiological regulatory systems. Regulatory processes that modulate reactivity include selective attention and processing of cues to reward and punishment. Research has shown that negative affectivity may contribute to guilt by providing strong internal cues of discomfort, increasing the likelihood that the cause of these feelings will be attributed to an internal conscience rather than external punishment or coercion (Ross, Millis, Bonebright, & Bailley, 2002). The proposed experiment investigated the possibility that there is a relationship between the ODT and the BAS/BIS scales. Specifically, the study tested the hypothesis that guilty subjects with high BIS scores will show greater diagnostic changes in PD and reading

measures because they will be concerned that their deception will be detected. Conversely, innocent subjects with a high BAS Reward Responsiveness score are expected to show less change in PD and smaller differences in reading across item types because they believe they have a greater chance of receiving the bonus.

In summary, the present study attempted to answer the following questions:

1. Does the accuracy of an ODT that asks directly if the person committed illicit acts differ from the accuracy of an ODT that indirectly asks if the person provided false answers on a questionnaire about those illicit activities?

2. Does the administration of a pre-ODT questionnaire about illicit activities covered by the ODT affect the accuracy of the subsequent ODT?

3. Are guilty subjects with relatively high BIS scores more likely to fail the ODT than guilty subjects with low relatively BIS scores?

4. Are innocent subjects with relatively high BAS scores more likely to pass the ODT than innocent subjects with relatively low BAS scores?

In the present study, some subjects were given a pre-ODT questionnaire that asked if they took an exam from a professor's office and if they took $20 from a secretary. After they completed the questionnaire, the subjects were given an ODT. Half of those subjects were asked directly if they committed each of the two crimes. The remaining subjects were asked if they falsified information on the pre-ODT questionnaire.

CHAPTER II

METHOD

Subjects and Design

One hundred nine subjects were recruited from the general University of Utah campus population. Recruitment flyers were posted on campus that advertised an opportunity to earn $30 and a possible bonus of $30 (for a total of $60) for participation in a psychological experiment. Subjects who spoke fluent English, were over the age of 18, could read a computer screen without glasses, and could read were scheduled for a session. Of these 109 subjects, 5 chose not to participate after learning their experimental condition, 6 did not follow instructions, and 2 had poor or incomplete data. This resulted in a sample size of 96 subjects. The sample sizes for the groups into which subjects were randomly assigned to are shown in Table 1. Demographic information obtained from subjects is presented in Tables 2 and 3.

Table 1
Sample sizes for cells of the design matrix

	Indirect Statements; Pre-ODT Questionnaire	Direct Statements; Pre-ODT Questionnaire	Direct Statements; No Pre-ODT Questionnaire
Innocent	24	12	12
Guilty	24	12	12

Table 2
Means, Standard Deviations, and Ranges for Age, BIS, BAS Drive, BAS Fun, and BAS reward responsiveness

Variable	*M*	*SD*	Range
Age	23.79	8.88	18 to 68
BIS	20.08	3.67	8 to 28
BAS Drive	11.29	2.34	4 to 16
BAS Fun	12.33	2.91	4 to 20
BAS Reward Responsiveness	16.94	3.35	5 to 24

Table 3
Frequencies and Percentages for Categorical Demographic Questions

Variable	Category	Frequency	%
Marital Status	Single	77	80.2
	Married	15	15.6
	Divorced	3	3.1
	Separated	1	1.0
Ethnicity	African American	1	1.0
	Asian	11	11.5
	South Pacific Islander	3	3.1
	Latino/a	13	13.5
	American Indian	1	1.0
	Middle Eastern	3	3.1
	Caucasian	64	66.7
Status	Student	88	91.7
	Staff	8	8.3
Class Standing	Freshman	20	22.7
	Sophomore	14	15.9
	Junior	27	30.7
	Senior	19	21.6
	Graduate	8	9.1
Enrollment Status	Full-Time	73	83.0
	Part-Time	15	15.6
Primary Language English	Yes	83	86.5
	No	13	13.5
Vision Correction	Glasses	12	12.5
	Contacts	31	32.3
	Neither	53	55.2

Overview of Design and Procedure

The design was a 2 x 3 x (3 x 5) mixed design with two between-subjects factors and two within-subjects factors. The between-subjects factors were guilt with two levels (guilty or innocent) and protocol with three levels: indirect ODT statements with pre-ODT questionnaire, direct ODT statements with pre-ODT questionnaire, and direct ODT statements with no questionnaire. The two within-subject factors were statement type (neutral, cash, and exam) and repetition (5 repetitions of the ODT test items). Time with 40 levels (10 Hz samples x 4 seconds) was also included as a within-subjects variable for the PD analyses.

Apparatus

An Arrington ViewPoint Eye Tracker was used to record eye movements and pupil diameter. The eye tracker was affixed to a pair of lensless plastic goggles. Viewing was binocular, but eye movement and pupil diameter was recorded only from the right eye. Data were collected at 30 Hz. Eyelab 3.48 (Kircher, Webb, & Cook, 2011) presented stimuli to the subject, and collected, edited, and analyzed the ocular-motor data. Eyelab communicated with the Arrington ViewPoint Eye Tracker software via functions in Arrington's software development kit (SDK). The 30 Hz PD data were imported into CPSLAB 11 (Scientific Assessment Technologies, Inc, Salt Lake City, UT), a general-purpose computer program for psychophysiological research. Stimuli were presented to the subject on a 19-inch Dell flat screen monitor. The monitor was positioned approximately 50 centimeters from the subject's eyes.

Ocular-motor Deception Test

Instructions and practice items were presented to the subject in black font with a pale grey background. Subjects answered test items after answering 15 practice items. There were 48 test items, and these same 48 items were presented five times in different orders. Sixteen items pertained to the theft of the $20, 16 pertained to the theft of the exam, and 16 were neutral items. The items were arranged such that no two items from the same category appear in succession. Statements were presented one at a time half way between the top and bottom of the screen starting on the left side. The screen width was 141 characters and the screen height was 51 lines. A T/F appeared to the right of the statement to remind subjects of their answer choices. Subjects answered by pressing 1 (True) or 3 (False) on the keypad. The correct (nonincriminating) answer was true for 8 of 16 items in a category and false for the remaining 8 items in each category. The test items are presented in Appendix A.

Procedures

Subjects reported alone to a room in a building on campus. Instructions in an envelope taped to the door instructed the subject to enter the room and read and sign the consent form, fill out the questionnaires in order, take the consent form and questionnaires with them when they left, and give the materials to the experimenter. Subjects completed the Behavioral Activation (BAS) and Behavioral Inhibition (BIS) questionnaire (Appendix C) and a demographic questionnaire (Appendix D). The subject then listened to a recording that gave their instructions for the study. A hard copy of the

recorded instructions was included. A phone number was provided for subjects to call if they did not wish to participate.

Half of the subjects were in the guilty condition. Guilty subjects were instructed to go to a secretary's office and ask the secretary where Dr. Mitchell's office is located. The secretary informed the subject that there was no Dr. Mitchell in the building, and the subject left. The subject was told to wait inconspicuously for the secretary to leave her office unattended, then enter her office, find her purse, remove $20 from a wallet in the purse, and concealed the money on their person. Subjects were told to prepare an alibi in case they were caught and not to leave fingerprints. They were informed that they had no more than 20 minutes to commit the crime and report to the experimenter.

Half of the subjects were in the innocent condition and did not steal anything. They were told that some subjects had to steal money from a secretary, but that they were innocent subjects and should not steal anything. Innocent subjects were instructed to wait approximately 20 minutes before reporting to the experimenter.

All subjects also were informed that there was another crime in which some subjects had to download an exam from a professor's computer onto a disk, but in actuality, no one committed that crime.

Subjects reported to the experimenter after committing their crime or after an appropriate waiting period. Subjects assigned to a pre-ODT questionnaire condition completed the questionnaire that asked if they took the exam or took the money. The subjects then sat at a computer, put on the Arrington eye tracker, and were tested about their possible commission of the two crimes or whether they lied on the questionnaire.

Subjects also completed an intervening task. The intervening task consisted of 18 T/F general world knowledge questions. The purpose of the intervening task was to minimize retention of the test items and answers. Subjects completed 5 repetitions of the test items and 4 sets of intervening task items. Intervening task items were not repeated across repetitions and were not used to make decisions about the subject's veracity.

After completing the tasks, subjects were paid and debriefed. Subjects were given an additional bonus if the computer determined they were innocent ($30 base pay plus $30 bonus). After the debriefing, subjects were asked not to discuss details of the study with others and released.

Dependent Measures

Behavioral Outcome Measures

Response time (RT). RT was the time in seconds from the appearance of the item on the screen to a button press response from the subject.

Proportion wrong. Proportion wrong for a particular statement type (neutral, cash, exam) was the number of incorrect responses divided by the number of items (16).

Ocular-motor Outcome Measures

An area of interest (AOI) was defined for each T/F test item. The AOI began with the first character of the item and ended at the period at the end of the statement. Ocular-motor reading measures were computed for the fixations in each AOI. Fixations were determined from the data files produced by the Arrington eye tracker by identifying a sequence of samples in which the eye shows little movement for at least 100 ms.

Fixations longer than 1000 ms were considered artifacts and were discarded (Rayner, 1998).

Number of fixations. Number of fixations was the number of fixations detected in an area of interest.

First pass duration. First pass duration was the sum of all fixation durations in an AOI before the eye fixated outside the area of interest.

Second pass duration. Second pass duration was the sum of all fixation durations in an AOI after the first time the eye fixated outside the area of interest.

Reread duration. Reread duration was the sum of all leftward eye movement fixation durations in the AOI. This measure assessed rereading, whether or not the eye fixated outside the AOI.

Peak amplitude of pupil response. Peak amplitude was obtained from a pupil response curve. The response curve began the moment the test statement appeared on the computer screen and ended 4 seconds later. The computer identified high and low points in the response curve and computed the difference between each low point and every succeeding high point. Peak amplitude was the greatest observed difference.

Area under the pupil response curve. Area under the curve was the area under the response curve from response onset to the point at which the response returned to the initial level or to the end of the 4-second sampling interval, whichever occurred first. Response onset was defined at the low point in the response curve from which peak amplitude was measured.

Level at T/F response. The PD waveform was standardized within subjects. The standard scores were averaged for a period of +/-1 second relative to the T/F response.

Item blink rate and next item blink rate. Blink rate was the number of blinks per second. Blink rate was computed for each item (item blink rate) from 1.5 seconds before the T/F response to initiation of the response. Blink rate was also computed for the subsequent item (next item blink rate) from the initiation of the T/F response for a duration of 1.5 seconds. A decrease in item blink rate may be thought of as an indicator of cognitive load, whereas an increase in next item blink rate may be viewed as a measure of relief (Stern & Skelly, 1984).

CHAPTER III

RESULTS

Significance for tests involving a repeating factor (statement type, repetition, and time) used Huynh-Feldt corrections to degrees of freedom. Effects were significant at $p<.05$ unless otherwise noted.

Preliminary Test for Effects of the Pretest Questionnaire

The primary goal of the present study was to determine if the accuracy of an ODT that asks directly if the person committed illicit acts differs from the accuracy of an ODT that indirectly asks if the person provided false answers on a questionnaire about those illicit activities. The manner in which this question was answered depended on whether the administration of a pre-ODT questionnaire affected ODT outcomes for subjects who received direct questions. Among the groups that received direct statements about their involvement in the mock crimes, I compared groups that did or did not complete a pre-ODT questionnaire.

Completion of the pre-ODT questionnaire did not interact with guilt for any of the outcome measures. Therefore, the questionnaire/no questionnaire groups that received direct questions were combined, and the presence or absence of pre-ODT questionnaires for participants who received direct statements on the ODT was dropped as a factor in all

subsequent analyses of ocular-motor and behavioral measures. Combination of the questionnaire and no questionnaire groups balanced the cells for subsequent comparisons of the direct and indirect treatments.

Repeated measures analysis of variance (RMANOVA) was used to analyze each dependent variable. The between-subjects factors were guilt and relevant issue (direct versus indirect). The within-subjects factors were statement type and repetition. For PD, time was an additional within-subjects factor. The RMANOVA contained many sources of variance. To simplify presentation of the results, only main effects of guilt and guilt interactions are presented and discussed in the text. Effect sizes for all statistically significant main effects and interactions for each dependent variable are presented in Appendix E.

There was no significant difference in the proportion of non-English speakers in the guilty and innocent groups, $p>.05$. There was also no significant difference between guilty and innocent groups in the proportions of those who reported that they wore glasses (for distance), contacts, or did not wear any corrective lens.

Means and standard deviations for the 11 dependent variables are presented in Table 4. They are broken down by relevant issue (direct or indirect), statement type (neutral, cash and exam), and guilt.

ANOVA Results for Behavioral and Reading Measures

There was no main effect of guilt or interaction of guilt with statement type or relevant issue for response time, proportion wrong, number of fixations, first pass duration, second pass duration, or reread duration.

Table 4

Means and Standard Deviations for the Dependent Variables by Direct/Indirect Relevant Issue, and Statement Type for Innocent and Guilty Subjects

		Direct						Indirect					
Dependent		Neutral		Cash		Exam		Neutral		Cash		Exam	
Variable		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Response Time	Innocent	24.95	12.82	25.41	11.20	25.50	12.75	23.76	5.23	24.13	5.61	24.51	5.44
	Guilty	27.01	6.01	26.74	6.13	28.46	6.88	25.83	7.49	26.77	8.05	28.32	9.32
Proportion Wrong	Innocent	.048	.052	.030	.053	.033	.054	.033	.023	.022	.024	.022	.026
	Guilty	.043	.029	.025	.020	.032	.019	.048	.047	.045	.081	.034	.044
Number of Fixations	Innocent	.105	.023	.109	.035	.107	.029	.111	.024	.118	.029	.119	.030
	Guilty	.117	.021	.117	.024	.124	.026	.111	.022	.119	.030	.124	.036
First Pass Duration	Innocent	32.53	6.78	32.65	7.92	32.40	7.49	35.05	6.24	37.32	7.28	37.90	7.07
	Guilty	37.16	5.96	37.27	7.54	37.94	5.91	34.68	6.87	37.06	7.87	38.36	8.85
Second Pass Duration	Innocent	3.60	4.89	3.48	5.80	3.59	6.13	2.08	3.47	1.71	3.63	1.71	2.89
	Guilty	3.31	3.32	2.27	2.59	3.27	3.50	3.80	4.02	3.14	3.68	3.82	4.71
Reread Duration	Innocent	10.34	4.00	9.32	5.03	10.21	4.93	9.77	3.87	9.60	4.18	9.89	3.88
	Guilty	11.37	3.65	9.87	5.14	12.24	4.71	10.35	3.89	10.27	4.64	11.22	5.47
PD	Innocent	.043	.070	.073	.078	.088	.082	.037	.122	.059	.116	.051	.127
	Guilty	.031	.092	.113	.094	.078	.089	.005	.066	.058	.084	.045	.089
PD Amplitude	Innocent	.623	.219	.632	.222	.643	.230	.829	.501	.838	.481	.843	.502
	Guilty	.698	.275	.733	.274	.727	.280	.883	.578	.897	.544	.884	.532
PD Area	Innocent	.408	.159	.431	.156	.459	.175	.526	.322	.544	.308	.554	.330
	Guilty	.438	.164	.535	.188	.498	.187	.543	.314	.621	.319	.603	.309
PD Level	Innocent	-.132	.164	.007	.112	.064	.129	-.071	.145	-.009	.113	.023	.119
	Guilty	-.193	.137	.133	.116	-.001	.131	-.125	.129	.057	.135	.058	.132
Item Blink Rate	Innocent	.754	.284	.798	.274	.837	.290	.683	.289	.680	.269	.672	.235
	Guilty	.736	.247	.766	.257	.789	.234	.677	.236	.701	.274	.705	.273
Next Item Blink	Innocent	.473	.180	.538	.191	.520	.182	.417	.229	.466	.235	.472	.213
Rate	Guilty	.426	.240	.490	.233	.476	.243	.479	.191	.458	.176	.509	.164

Note. Response time, first pass duration, second pass duration and reread duration are ms per character. Number of fixations is number of fixations per character. Pupil diameter is change from baseline in mm. Item blink rate is number of blinks per second on each item. Next item blink rate is number of blinks per second on the item following neutral, cash, and exam items.

Pupil Diameter

PD was assessed by computing change from baseline. The first data point was subtracted from every subsequent data point in the response curve. A positive value indicated PD increased relative to baseline, and a negative value indicated PD decreased relative to baseline.

PD response curves for the Guilt X Statement type X Time interaction are presented in Figures 1a and 1b for innocent and guilty subjects, respectively. The Guilt X Statement type interaction was significant, $F(1.95, 179.62) = 11.12$, as was the Guilt X Statement type X Relevant issue, $F(1.95, 179.62) = 3.25$. The Guilt X Statement type X Time interaction was significant, $F(7.87, 724.34) = 3.38$ as was the Guilt X Statement type X Relevant issue X Time, $F(7.87, 724.34) = 2.35$.

Peak Amplitude

There was no main effect of guilt or interaction of guilt with statement type or relevant issue for PD peak amplitude.

Area Under Pupil Response

The Guilt X Statement type interaction was significant, $F(1.798, 168.999) = 17.72$ and is presented in Figure 2.

Level at Response Onset

The Guilt X Statement type interaction was significant, $F(1.84, 172.75) = 14.26$ and is presented in Figure 3.

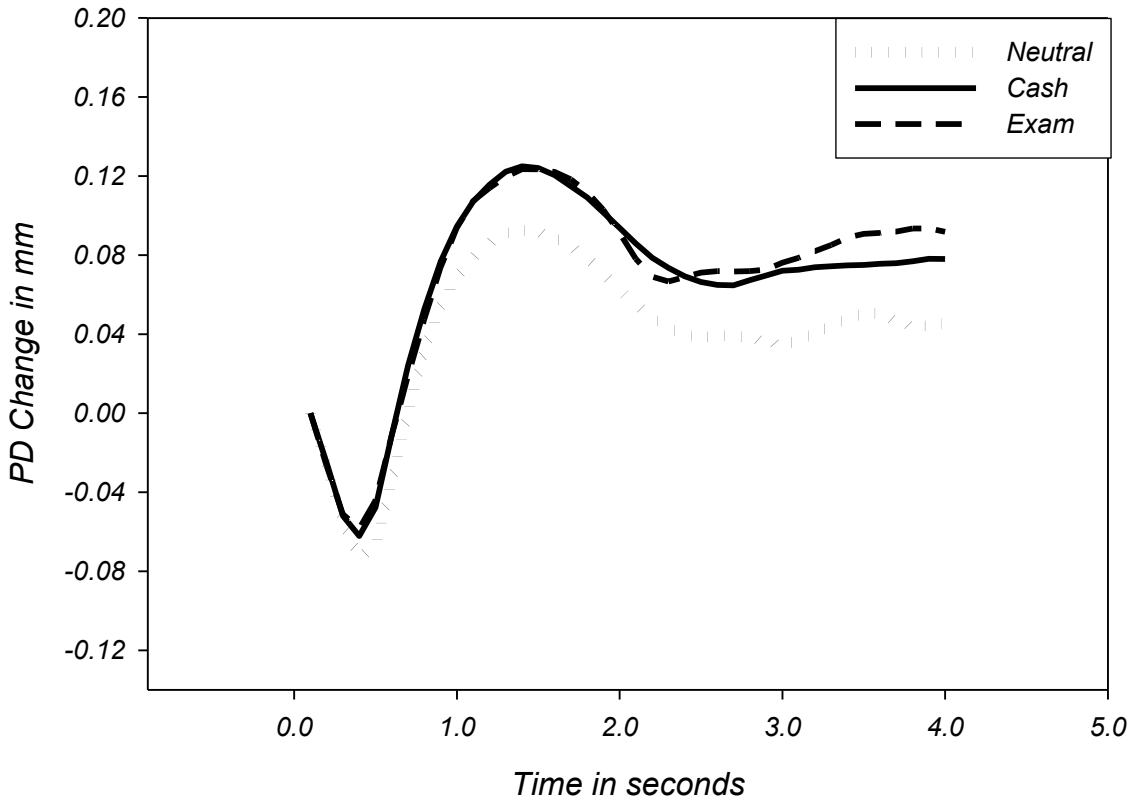

Figure 1. Pupil response to neutral, cash, and exam items. a) Innocent subjects. b) Guilty subjects.

b)

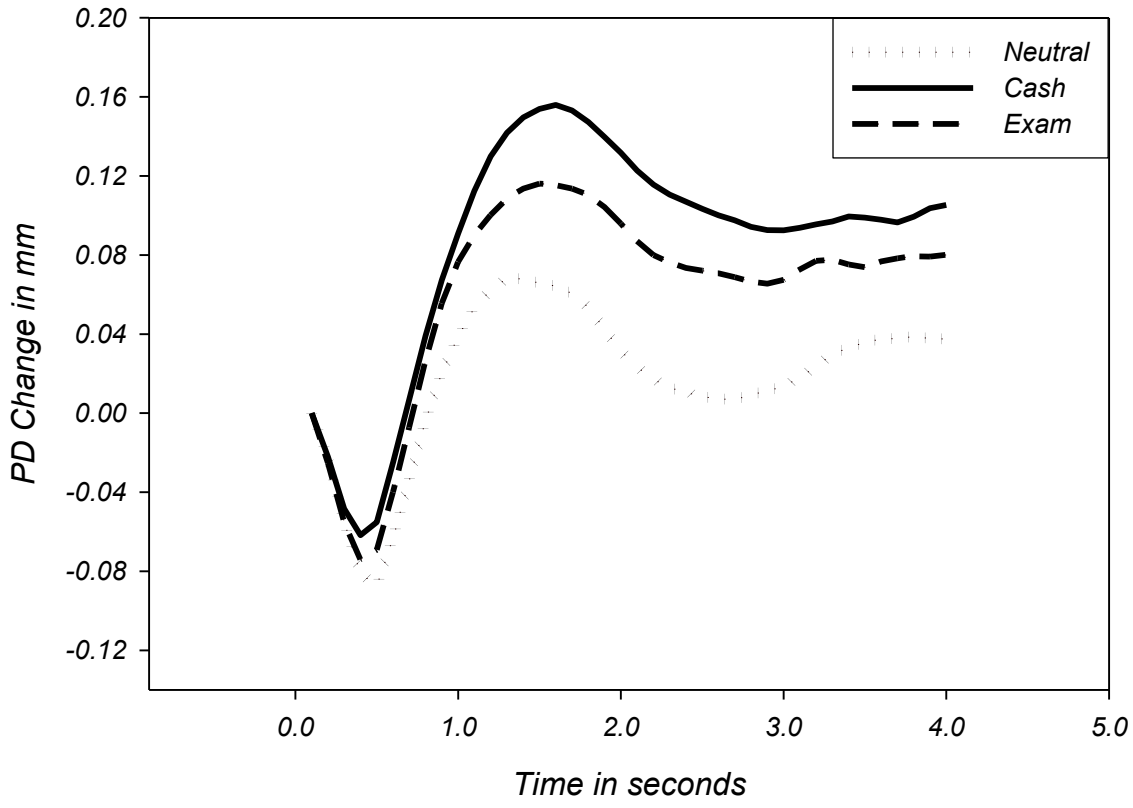

Figure 1. Continued

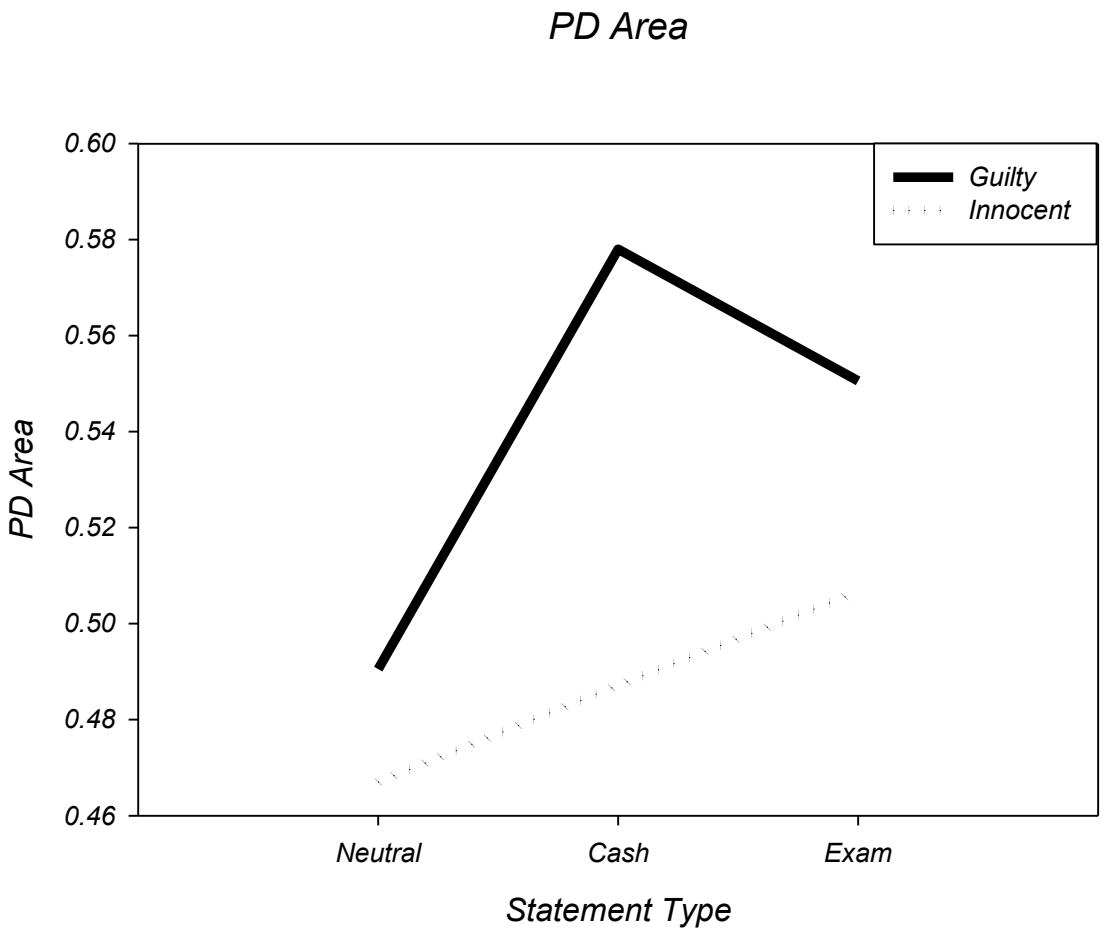

Figure 2. Guilt X Statement Type interaction for PD Area.

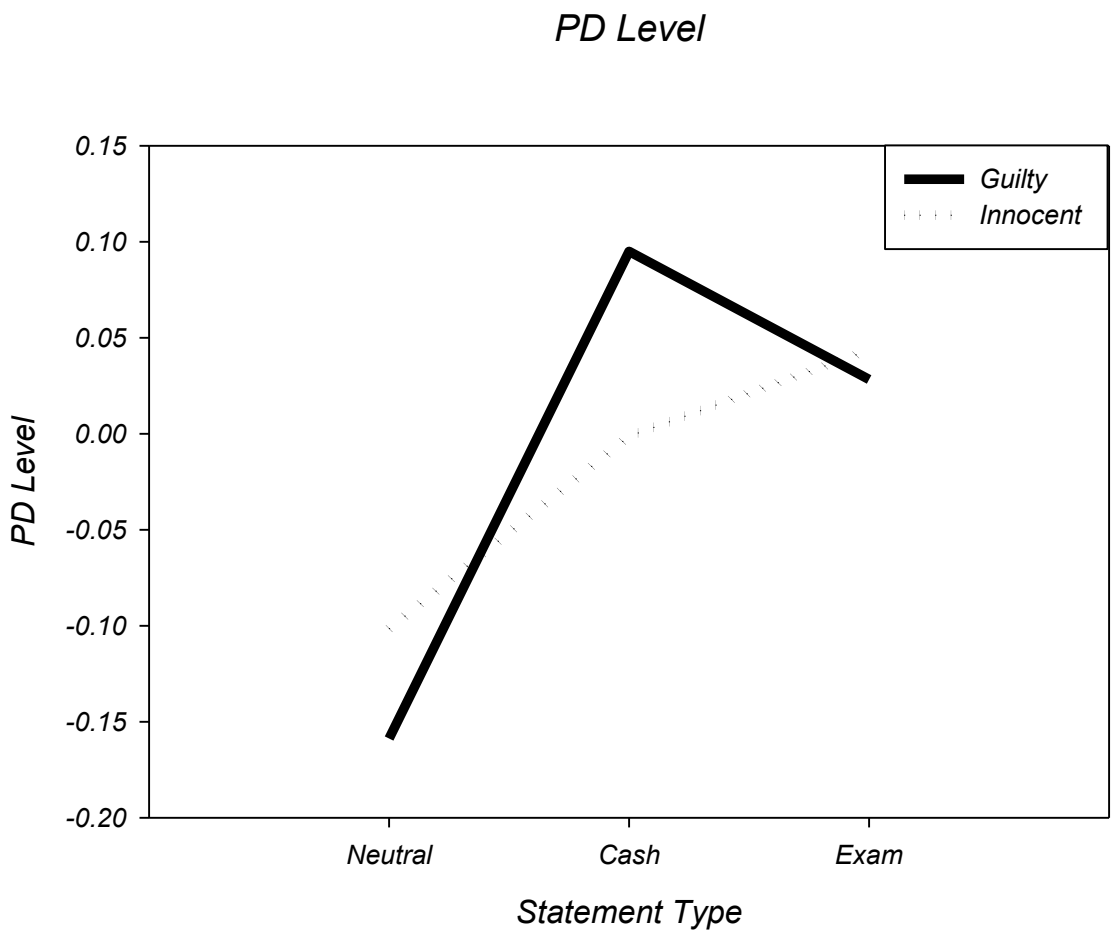

Figure 3. Guilt X Statement Type interaction for PD Level.

Item Blink Rate

There was no main effect of guilt or interaction of guilt with statement type or relevant issue for item blink rate.

Next Item Blink Rate

There was no main effect of guilt or interaction of guilt with statement type or relevant issue for next item blink rate.

Predictive Validity of Ocular-motor Measures

New dependent variables were generated from the ocular-motor measures to develop statistical classifiers. One dependent variable was the difference between the mean for crime-related items and the mean for neutral items (i.e., [R1+R2]/2-N]). This variable provided a measure of concern about the relevant issues in general. Another new dependent variable was the difference between the mean for cash items and the mean for exam items (i.e., R1-R2). This difference provided a measure of deception that controlled for the perceived relevance of test items. The third new dependent variable was the mean for the neutral items. This variable provided a general measure of cautiousness, trepidation, or general cognitive load. These measures were derived for each behavioral and ocular-motor variable.

To assess the diagnostic validity of a derived outcome measure, it was correlated with a dichotomous variable that distinguished between innocent (coded 1) and guilty (coded 2) subjects. This resulted in one mean for the neutral items, one mean for the cash items, and one mean for the exam items for each of the five repetitions.

The negative point-biserial correlations for RT, proportion wrong, and number of fixations between the relevant crimes means that guilty subjects took less time to respond, made fewer mistakes, and made fewer fixations on cash items than exam items. The positive point-biserial correlation for Neutral first pass duration (FirstPassNeutral) indicates that guilty subjects who received direct statements spent more time reading neutral items compared with innocent subjects. The negative reread duration correlation for Cash versus Exam items (r = -.364 for RereadCashExam) indicates that guilty subjects did less rereading of cash items than exam items. The correlations for the Crime versus Neutral items and Cash versus Exam items were positive for amplitude, area, and level, which indicates that guilty subjects showed greater increases in pupil size in response to relevant items than did innocent subjects. For innocent subjects, next item blink rate increased following Crime statements compared to Neutral statements. For guilty subjects, there was little difference between Crime and Neutral statements (see Table 4). The point-biserial correlations for each measure are presented in Table 5 separately for groups that answered direct or indirect statements.

Seven variables then were selected for possible inclusion in the discriminant function: NFixCashExam, FirstPassNeutral, RereadCashExam, PDAreaCrimeNeutral, PDAreaCashExam, PDLevelCrimeNeutral, and PDLevelCashExam. The variables were selected because they had significant point-biserial correlations of at least .33 in either the direct or indirect item groups. NextItemBlinkRateCrimeNeutral and DifferencebetweenblinksCrimeNeutral had statistically significant correlations with group membership, but they were omitted from the discriminant analyses because the observed differences among the means for the Guilt X Statement Type interactions

Table 5
Point-Biserial Correlations for Direct and Indirect Relevant Issues

Outcome Measure	Direct Correlations	Indirect Correlations
RTNeutral	.105	.161
RTCrimeNeutral	.010	.144
RTCashExam	-.311*	-.281
PropWrongNeutral	-.063	.198
PropWrongCrimeNeutral	.043	.041
PropWrongCashExam	-.311*	-.281
NFixNeutral	.269	-.001
NfixCrimeNeutral	.008	.099
NfixCashExam	-.402**	-.212
FirstPassNeutral	.348*	-.029
FirstPassCrimeNeutral	.070	.057
FirstPassCashExam	-.160	-.115
SecondPassNeutral	-.035	.227
SecondPassCrimeNeutral	-.135	.019
SecondPassCashExam	-.228	.-214
RereadNeutral	.137	.076
RereadCrimeNeutral	.051	.095
RereadCashExam	-.364*	-.177
PDAmplitudeNeutral	.152	.051
PDAmplitudeCrimeNeutral	.318*	-.040
PDAmplitudeCashExam	.325*	.257
PDAreaNeutral	.094	.027
PDAreaCrimeNeutral	.373**	.352*
PDAreaCashExam	.684**	.268
PDLevelNeutral	-.202	-.195
PDLevelCrimeNeutral	.352*	.331*
PDLevelCashExam	.649**	.144
ItemBlinkRateNeutral	-.035	-.012
ItemBlinkRateCrimeNeutral	-.086	.147
ItemBlinkRateCashExam	.094	-.067
NextItemBlinkRateNeutral	-.112	.148
NextItemBlinkRateCrimeNeutral	.005	-.322*
NextItemBlinkRateCashExam	-.021	-.245
DifferencebetweenblinksNeutral	-.075	.142
DifferencebetweenblinksCrimeNeutral	.066	-.298*
DifferencebetweenblinksCashExam	-.073	-.122

*$p < .05$, **$p < .01$. *Note.* RT = response time per character, PropWrong = proportion wrong, NFix = number of fixations per character, FirstPass = time spend reading per character, SecondPass = second time reading per character, Reread = time spent rereading per character, PDAmplitude = pupil diameter peak amplitude, PDArea = pupil diameter area under the curve, ItemBlinkRate= number of blinks per second on each item type, NextItemBlinkRate = number of blinks per second on the item following each item type, Differencebetweenblinks = NextItemBlinkRate – ItemBlinkRate, Neutral = response for neutral items, CrimeNeutral = difference between crime-related and neutral items, and CashExam = difference between cash and exam items.

were not predicted nor had they been observed in any prior experiments.

The seven variables were submitted to a stepwise discriminant analysis. Results indicated that FirstPassNeutral, PDAreaCrimeNeutral, PDAreaCashExam, and PDLevelCashExam best predicted guilt for direct items and FirstPassNeutral, PDAreaCrimeNeutral, and PDAreaCashExam best predicted guilt for indirect items. Coefficients for variables in each discriminant function were statistically significant, $p < .05$. The standardized canonical discriminant function coefficients and the functions at group centroids are presented in Tables 6 and 7, respectively. Classification results and jackknifed classification results are presented in Table 8. Jackknifed classification results were obtained with the leave-one-out method; that is, each case was classified using discriminant coefficients for the predictor variables that were based on all cases except the one that was classified. Classification results for the logistic regression using the same variables as were included in the discriminant functions are presented in Table 9.

Table 6
Standardized Canonical Discriminant Function Coefficients

Relevant issue	Variable	Function
Direct	FirstPassNeutral	.441
	PDAreaCrimeNeutral	.298
	PDAreaCashExam	.703
	PDLevelCashExam	.426
Indirect	FirstPassNeutral	-.451
	PDAreaCrimeNeutral	.806
	PDAreaCashExam	.559

Table 7
Functions at Group Centroids

Relevant issue		Function
Direct	Innocent	-1.264
	Guilty	1.264
Indirect	Innocent	-.487
	Guilty	.487

Table 8
Frequencies (and Percentages) of Cases Correctly Classified With the Discriminant Function

		Actual Group Membership	Predicted Group Membership		Total Correct
Original			Innocent	Guilty	
	Direct	Innocent	23 (95.8)	1 (4.2)	
		Guilty	4 (16.7)	20 (83.3)	
		Total			89.6%
	Indirect	Innocent	19 (79.2)	5 (20.8)	
		Guilty	10 (41.7)	14 (58.3)	
		Total			68.8%
Jackknifed					
	Direct	Innocent	20 (83.3)	4 (16.7)	
		Guilty	4 (16.7)	20 (83.3)	
		Total			83.3%
	Indirect	Innocent	16 (66.7)	8 (33.3)	
		Guilty	10 (41.7)	14 (58.3)	
		Total			62.5%

Table 9

Frequencies (and Percentages) of Cases Correctly Classified With the Logistic Regression

		Actual Group Membership	Predicted Group Membership		Total Correct
Original			Innocent	Guilty	
	Direct	Innocent	21 (87.5)	3 (12.5)	
		Guilty	4 (16.7)	20 (83.3)	
		Total			85.4%
	Indirect	Innocent	19 (79.2)	5 (20.8)	
		Guilty	9 (37.5)	15 (62.5)	
		Total			70.8%

Behavioral Activation and Inhibition

Analyses were conducted to determine if BIS/BAS measures explained variance in the diagnostic value of discriminant scores within guilty and innocent groups. The scoring for the BAS/BIS scale is included in Appendix A. Since different discriminant functions were created for subjects in the direct and indirect conditions, correlational analyses were conducted separately for the two conditions.

Within-group correlations are presented in Table 10. There was only one significant correlation and it was in the predicted direction; guilty subjects with large positive discriminant scores appeared more deceptive on the ODT and those discriminant scores were positively correlated with BIS scores.

Table 10
Correlations Between Discriminant Scores and BIS/BAS Scale Scores for Innocent and Guilty Subjects who Received Direct or Indirect Statements on the ODT

		BIS	Drive	Fun	Reward	Total BAS
Direct	Innocent	-.214	.047	-.093	-.070	-.058
	Guilty	.447*	-.046	-.135	.318	.073
Indirect	Innocent	.191	.044	-.124	.142	.027
	Guilty	.280	-.070	-.222	.161	-.027

*$p < .05$, **$p < .01$.

CHAPTER IV

DISCUSSION

The present study evaluated the effects of guilt, relevant issue, statement type, and questionnaire on ocular-motor and behavioral measures. Results for direct items with no questionnaire generally replicated Webb's dissertation.

There were no significant ANOVAs for the reading measures, but PD showed several significant effects. Overall classification rates exceeded 89% for the direct items and 68% for the indirect items. Direct item accuracy was similar to the accuracy rates reported by Webb (2008) and the indirect item accuracy rates resembled those obtained by Hacker et al. (2010), who correctly classified only 60% of the guilty and 78% of the innocent subjects. The discriminant functions for direct and indirect statements included both reading measures and changes in pupil size. The classification rates based on the functions in the present study and in Webb's dissertation suggest that a combination of PD and reading measures can be used to make accurate diagnoses of truth and deception.

The present study found that the accuracy on an ODT that asks directly if the person committed illicit acts differs from the accuracy of an ODT that indirectly asks if the person provided false information on a pre-ODT questionnaire. There was 89% overall accuracy for direct items versus 68% accuracy for indirect items. Differences between the point-biserial correlations for direct and indirect groups were consistent with

the differences between the groups in accuracy rates. There were more significant point-biserial correlations for the direct items (11 of 36) than the indirect items (4 of 36). In addition, the differences between cash and exam items were more diagnostic for subjects interrogated about the crime than for subjects interrogated about the questionnaire. The ODT uses the Relevant Comparison Test, and that test is based on the idea that the difference between crime-related items should be more diagnostic than the difference between crime-related and neutral items. Therefore, the results obtained with direct items were not only stronger than those obtained with indirect items but also more consistent with the rationale that underlies the ODT.

Although the accuracy rates for direct and indirect groups differed, only 1 of 10 statistical tests of the Guilt X Statement Type X Relevant Issue was significant. Since the point-biserial correlations revealed a general pattern of greater diagnostic validity for direct items as compared to indirect items, there may not have been enough power to reliably detect a three-way interaction.

PDAreaCashExam and PDLevelCashExam for direct items had validity coefficients that exceeded .64. The observed differences between groups in pupil size are consistent with the idea that deception requires more cognitive effort than does truthfulness. Apparently, the additional investment of cognitive resources was beneficial to guilty subjects, because their error rates were lower than those of innocent subjects. This finding differs from Webb (2008) who found that guilty subjects made significantly *more* mistakes than did innocent subjects. Since the effects on pupil size in the present study were somewhat greater than those in Webb's dissertation, subsequent research

should explore the possibility that evidence of effortful information processing may be associated with fewer mistakes on the ODT.

The present study also asked if the administration of a pre-ODT questionnaire about illicit activities covered by the ODT affected the accuracy of the subsequent ODT. Guilty subjects who complete a pre-ODT questionnaire could become habituated or sensitized to questions about the crimes and less affected by those questions on the subsequent ODT. The results from the present study suggest that a pre-ODT questionnaire does not significantly affect the ODT. There remains a possibility that a questionnaire composed of more than two items could affect subjects' performance on the subsequent ODT, but the pre-ODT questionnaire in the present study had no discernible effect on the ocular-motor measures.

The present study also tested if guilty subjects with relatively high BIS scores were more likely to fail the ODT than guilty subjects with relatively low BIS scores. As predicted, scores on the BIS scale were positively correlated with deceptiveness as measured by discriminant scores in the guilty group that received direct statements. People with high BIS sensitivity were expected to experience high levels of anxiety in the presence of cues of impending punishment. Theoretically, guilty subjects with high BIS scores should be more adversely affected by the commission of the mock crime and subsequent interrogation about the crime than guilty subjects who are less sensitive to cues for nonreward or punishment. Gray's (1987) theory posits that the output of the BIS is behavioral inhibition along with increased arousal and heighted attention. As a result, when guilty subjects with high BIS scores respond deceptively to test questions, they should be concerned they will get caught and not earn the bonus, which increases anxiety

and sensitivity to nonreward. Additional support for this hypothesis is the psychological set theory, which holds that when a person being examined fears punishment or anticipates serious consequences if they fail to deceive, then the fear or anticipation produces a measurable physiological reaction if the person answers deceptively (Barland, 1981). Therefore, the amplified anxiety of someone with a greater fear of consequences could be the cause of the greater physiological reaction we see with the ODT. The correlation between BIS scores and discriminant scores for guilty subjects who answered indirect items (r = .28) was not significant, but it was in the same direction as the correlation obtained for guilty subjects who answered direct items (r = .45). The failure to obtain a significant correlation for the indirect group may be related to the finding that the discriminant scores were less diagnostic of deception in that group.

A practical implication for this finding is that the BIS could be given to potential ODT examinees. A person with a low BIS score who passes the ODT might produce a false negative outcome because guilty people with low BIS scores appear less deceptive on the ODT. Conversely, a person with a high BIS score who passes the ODT is more likely to be truthful, since a guilty person with a high BIS score would be expected to fail the ODT. The potential usefulness of the BIS scale for field applications of the ODT requires additional study.

The last question we set out to answer is if innocent subjects with relatively high BAS scores are more likely to pass the ODT than innocent subjects with low BAS scores. I did not find that BAS scores were related to indications of truthfulness. A possible reason for not finding significant results is that Drive and Fun are 4-item scales and Reward Responsiveness is a 5-item scale. The number of items on the BAS scales may

be insufficient to achieve adequate reliability. However, correlations between discriminant scores and the sum of all BAS items were not significant either.

Beyond the research questions, an additional motivation for this study was to determine why the results of Hacker et al. (2010) study differed from those of Webb (2008). The present study administered a pre-ODT questionnaire to mimic the pre-ODT simulated application for a scholarship in the Hacker et al. study. Since there were no significant effects of the pre-ODT questionnaire on any of the ocular-motor measures, the results of the present study suggest that the administration of the pretest questionnaire in the Hacker et al. study was not responsible for their low accuracy rates.

The results of the present study suggest that the accuracy rates in the Hacker et al. (2010) were low because they used indirect items. Why would the use of indirect items result in less diagnostic ocular-motor measures? Differences in the semantic complexity of items on the two forms of the test might count for the loss of diagnostic validity (Appendix A). The relevant issue for a direct statement referred to the commission of a particular crime (an action). The relevant issue for an indirect statement referred to falsifying information on a questionnaire (one action) concerning the crime in question (another action). To answer an indirect statement correctly, the subject had to retain information concerning their possible involvement in the crime and how they responded on the questionnaire. Guilty subjects had the added burden of distinguishing between items answered truthfully and items answered deceptively. If there was a ceiling effect for guilty subjects, the additional burden of item complexity might raise the load on innocent subjects and reduce the difference between guilty and innocent subjects. Item difficulty was identified as a factor that influenced the diagnostic validity of reading

measures in Webb's (2008) dissertation, and it provides a plausible explanation for the results obtained by Hacker et al.

Lying on a questionnaire may have been less arousing than the commission of a realistic mock theft. A subject who lied on the questionnaire wrote "No" to one question on a form. The deceptive subject had just planned and executed a theft from a temporarily unoccupied office of a secretary during normal working hours. Lying on the questionnaire may have been a mere afterthought, since the guilty subject may have been focused on denying culpability about the crime, not their response on the questionnaire. Writing "No" on the questionnaire was only the last of several illicit behaviors, and it may have been the least emotionally arousing of those behaviors because it posed less risk of discovery. Guilty subjects who were asked about their answers on the questionnaire may have been relieved that they were not asked if they had committed the crime.

There may be greater social stigma associated with lying about committing a theft than lying on a questionnaire. Five subjects withdrew from the study upon learning they had to steal $20 from a secretary's wallet, and six subjects chose not to steal the money but showed up for the ODT anyway. No one refused to lie on a questionnaire. The conditioned emotional response theory of deception detection would predict a stronger effect of deception on ocular-motor responses to statements about the crime than statements about the pre-ODT questionnaire (Davis, 1961). In addition, Levine, Shaw, and Shulman (2010) found that direct interrogative questioning of a potential liar is associated with detection accuracy rates substantially higher than is typical of the

literature due to 'leakage,' where lying on direct questions produces guilt and anxiety, which then reveals inadvertent cues that signal deceit.

In the present study, two factors were confounded - the conceptual proximity of the relevant issue to the crimes under investigation and emotional arousal. Arousal is important in regulating consciousness, attention, and information processing and it is crucial for motivating certain behaviors. Due to their fear of being caught while wearing an eye tracker, liars could have been more highly aroused when answering relevant statements, such as "I did not take the $20," than when answering statements about a questionnaire. Perhaps a future study could manipulate the arousal levels and the directness of the relevant issue independently to determine their individual and joint effects on ocular-motor measures. Further research will be needed to determine why indirect questioning is not as diagnostic of deception as direct questioning.

Ideally, high levels of accuracy would have been achieved with indirect statements about subjects' responses on a pre-ODT questionnaire. Given that high levels of accuracy were not achieved with indirect questions, the present findings do not support the development of a standardized ODT that would test if the subject falsified answers on a pre-ODT questionnaire. On the other hand, if the reason(s) the indirect interrogation failed can be established in future research, the possibility remains that a generic ODT could be developed that would prove effective for many different applications.

Limitations

The present study was a laboratory experiment. The ODT may be more or less effective in field situations where subjects may be more highly motivated to pass the test, but high levels of experimental control are often difficult to achieve.

Another limitation was that the sample consisted mostly of single Caucasian college students. This sample was representative of the University of Utah population, but generalizations to the general population may be limited. If the ODT is to be used for security screening, it is important to ensure the results generalize to the populations of interest. The mock crime procedures in the present study were designed to maximize differences between truthful and deceptive subjects on ocular-motor measures. The guilty subjects committed an emotionally engaging and realistic mock crime, and then they denied their involvement on a deception test that took place immediately after commission of the crime. These procedures have been found to produce physiological reactions in polygraph examinations that are indistinguishable in most respects to those obtained from suspects in actual criminal investigations (Kircher et al., 1994). Whether or not these procedures produce ODT outcomes that are representative of those obtained in the field is unknown.

Because the ODT is administered by a computer, a number of examinees could be tested simultaneously by a single proctor. In that scenario, the subject would work alone at a workstation until they complete the test. In the present study, the experimenter sat in the same small room with the subject while they completed the ODT. Whether the presence of the experimenter in the room contributes to evaluation apprehension and whether that affects the ODT also is unknown.

Implications and Future Directions

Results from the present study and Cook et al. (2012) suggest that a combination of behavioral and ocular-motor measures can be used to detect deception. These results were found in a mock-crime study similar to a forensic situation, but they also have potential for use in a security screening situation. In a security screening situation, subjects are asked questions about several issues, and they may or may not be deceptive about one or more issues on the test. Future work should test if there are advantages or disadvantages to adding issues to the test.

Results from the present study suggest that it is easier to detect deception when the relevant issue directly addresses the behavior of interest than if the relevant issue indirectly addresses whether the subject was truthful or deceptive on a questionnaire. Although the indirect statements yielded less accurate classifications than the direct statements, discrimination between truthful and deceptive who received indirect statements was significantly greater than chance accuracy. In a field setting, the accuracy of the indirect ODT might improve and be adequate for a screening environment. Based on the present results, the highest accuracy is likely to be achieved with direct questions about specific illicit behaviors rather than indirect statements about responses on a pre-test questionnaire. More research is needed to determine why indirect questioning does not produce results that are as diagnostic as direct questioning.

Summary

Several behavioral and ocular-motor measures were diagnostic of deception, and a weighted combination of four variables for direct items correctly classified 95% of the innocent and 83% guilty. Three variables for indirect items correctly classified 79% of innocent and 58% of the guilty. Whereas administration of a pre-ODT questionnaire did not significantly affect the results obtained with the ODT, the relevant issues covered by the ODT affected the diagnostic validity of ocular-motor measures. Results from the present study suggest that the best ODT is one that directly addressed the relevant issues that motivate the use of the test.

APPENDIX A

TRUE/FALSE ITEMS

Direct Items

Cash Items

True
I was uninvolved in the theft of the $20.
The wallet in the office was untouched by me.
I did not take anything from the wallet in the office.
The theft involving the purse was not my fault.
The article from the purse was not stolen by me.
The secretary's property was not stolen by me.
I am innocent of taking the secretary's property from the office.
I had nothing to do with the theft of the $20.

False
I am guilty of taking something from the secretary's office.
The reason the $20 is gone is because I took it.
The item from the purse is hidden on my person.
I removed something from the purse in the office.
I am not innocent of stealing the $20.
The wallet was not recently tampered with by anyone but me.
The secretary's property was stolen by me.
I know what happened to the item missing from the wallet.

Exam Items

True
I took nothing from the professor's office.
The disk was untouched by me.
The loss of the professor's information is not my fault.
The information from the computer is not in my possession.
I did not take the exam from the office.
The information was not copied to a disk by me.
I did not take anything from the computer in the office.
I am not guilty of taking the exam from the office.

False
I made a copy of the professor's information.
The disk in my possession contains the copied information.
The loss of the exam is no one's fault but mine.
The missing exam is in my possession.
I am not innocent of copying the information to a disk.
I took the information from the computer in the office.
The professor's information is missing because of me.
I copied the information from the computer.

Indirect Items

Cash Items

True
I answered truthfully that I was uninvolved in the theft of the $20.
My response truthfully indicated that the wallet in the office was untouched by me.
I truthfully responded that I did not take anything from the wallet in the office.
My response that the theft involving the purse was not my fault is truthful.
I correctly reported that the article from the purse was not stolen by me.
My answer correctly suggested that the secretary's property was not stolen by me.
I correctly reported that I am innocent of taking the secretary's property from the office.
My response that I had nothing to do with the theft of the $20 is accurate.

False
I admitted that I am guilty of taking something from the secretary's office.
My answer correctly suggested that the $20 is gone because I took it.
I admitted that the item from the purse is hidden on my person.
My answer that I removed something from the purse in the office is valid.
I answered truthfully that I am not innocent of stealing the $20.
My response that the wallet was not recently tampered with by anyone but me was false.
I truthfully indicated that the secretary's property was stolen by me.
My answer correctly indicated that I knew what happened to the item missing from the wallet.

Exam Items

True
I correctly reported that I took nothing from the professor's office.
My answer accurately indicated that the disk was untouched by me.
I accurately claimed that the loss of the professor's information is not my fault.
My response correctly indicated that the information from the computer is not in my possession.
I correctly indicated that I did not take the exam from the office.
My answer that the information was copied to a disk by me is false.
I truthfully responded that I did not take anything from the computer in the office.
My answer that I am not guilty of taking the exam from the office is correct.

False
I lied that I did not make a copy of the professor's information.
My answer correctly indicated that the disk in my possession contains the copied information.
I admitted that the loss of the exam is no one's fault but mine.
The response that the missing exam is in my possession is true.
I truthfully reported that I am not innocent of copying the information to a disk.
My response that I took the information from the computer in the office was correct.

I correctly indicated that the professor's information is missing because of me.
My answer that I copied the information from the computer is accurate.

Neutral Items

True
I was born prior to the year 2000
The sky is blue on sunny days.
Cats and dogs are often kept as pets.
Dinosaurs used to roam the earth.
I am reading this on a day other than Sunday.
Polar bears do not roam freely in Mexico.
Global warming is a concern for many people.
Large SUVs often get lower gas mileage than newer compact cars.

False
I am reading this sentence on March 12, 2002.
San Francisco is in the state of Nevada
There are only 35 states in the United States.
Road construction is fast and convenient for motorists.
I have never listened to radio or watched TV.
Whales do not live in any of the world's oceans.
Trees are never harvested for lumber.
Morbid obesity is not a health concern in the United States.

APPENDIX B

PRE-ODT QUESTIONNAIRE

Exam Form

Your answer to this question is important.

1. Did you take the exam from the professor's office?

Cash Form

Your answer to this question is important.

1. Did you take $20 from the secretary's wallet?

APPENDIX C

BIS/BAS SCALES

BIS/BAS Scales

Each item of this questionnaire is a statement that a person may either agree with or disagree with. For each item, indicate how much you agree or disagree with what the item says. Please respond to all the items; do not leave any blank. Choose only one response to each statement. Please be as accurate and honest as you can be. Respond to each item as if it were the only item. That is, don't worry about being "consistent" in your responses. Choose from the following four response options:

1 = very true for me
2 = somewhat true for me
3 = somewhat false for me
4 = very false for me

1. A person's family is the most important thing in life.	1	2	3	4
2. Even if something bad is about to happen to me, I rarely experience fear or nervousness.	1	2	3	4
3. I go out of my way to get things I want.	1	2	3	4
4. When I'm doing well at something I love to keep at it.	1	2	3	4
5. I'm always willing to try something new if I think it will be fun.	1	2	3	4
6. How I dress is important to me.	1	2	3	4
7. When I get something I want, I feel excited and energized.	1	2	3	4
8. Criticism or scolding hurts me quite a bit.	1	2	3	4
9. When I want something I usually go all-out to get it.	1	2	3	4
10. I will often do things for no other reason than that they might be fun.	1	2	3	4
11. It is hard for me to find the time to do things such as get a haircut.	1	2	3	4
12. If I see a chance to get something I want I move on it right away.	1	2	3	4
13. I feel pretty worried or upset when I think or know somebody is angry at me.	1	2	3	4
14. When I see an opportunity for something I like I get excited right away.	1	2	3	4
15. I often act on the spur of the moment.	1	2	3	4
16. If I think something unpleasant is going to happen I usually get pretty "worked up."	1	2	3	4
17. I often wonder why people act the way they do.	1	2	3	4
18. When good things happen to me, it affects me strongly.	1	2	3	4
19. I feel worried when I think I have done poorly at something important.	1	2	3	4
20. I crave excitement and new sensations.	1	2	3	4
21. When I go after something I use a "no holds barred" approach.	1	2	3	4
22. I have very few fears compared to my friends.	1	2	3	4
23. It would excite me to win a contest.	1	2	3	4
24. I worry about making mistakes.	1	2	3	4

Scoring

Items other than 2 and 22 are reverse-scored.

BAS Drive: 3, 9, 12, 21
BAS Fun Seeking: 5, 10, 15, 20
BAS Reward Responsiveness: 4, 7, 14, 18, 23
BIS: 2, 8, 13, 16, 19, 22, 24

APPENDIX D

DEMOGRAPHIC QUESTIONNAIRE

Subject ID #_____

1. Age:_____

2. Sex: (circle one) Male Female

3. Marital status: (circle one)

Single
Married
Divorced
Widowed
Separated

4. Racial/Ethnic Origin: (circle one)

African American
Asian
South Pacific Islander
Latino/a
American Indian
Middle Eastern
Caucasian
Other (please explain):_____

5. What is your status? (circle one)

Student
Staff
Other

6. If you are a student, what is your college major? _____

7. If you are a student, what is your class standing? (circle one)

Freshman
Sophomore
Junior
Senior
Graduate

8. If you are a student, what is your enrollment status? (circle one)

Full-time
Part-time
Other (please explain):_____

10. If you are not a student, what is the highest level of school or degree you have completed? (circle one)

High school
Trade school
Associate's degree
Bachelor's degree
Master's degree
Professional degree
Doctorate degree

11. Is English your primary language? (circle one) Yes No

If you circled No, what is your primary language?_____

12. Do you wear any of the following for vision correction for reading? (circle one)

Glasses
Contacts
Neither

APPENDIX E

EFFECT SIZES FOR EACH DEPENDENT VARIABLE

Table 11
Effect Sizes for Response Time

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.070
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 12
Effect Sizes for Proportion Wrong

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.071
Questionnaire (direct only)	
Sex	.049
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	.098
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	.095
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 13
Effect Sizes for Number of Fixations

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.139
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 14
Effect Sizes for First Pass Duration

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.121
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	.083
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 15
Effect Sizes for Second Pass Duration

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.056
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 16
Effect Sizes for Reread Duration

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.114
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	.038
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 17
Effect Sizes for PD Peak Amplitude

Source	Effect Size
Guilt	
Relevant issue	.053
Statement type	.101
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 18
Effect Sizes for PD Area

Source	Effect Size
Guilt	
Relevant issue	.044
Statement type	.382
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	.171
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 19
Effect Sizes for PD Level

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.531
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	.155
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	.082
Relevant issue X Sex	.048
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	.042
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 20
Effect Sizes for PD

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.428
Questionnaire (direct only)	
Sex	
Time	.373
Guilt X Relevant issue	
Guilt X Statement type	.123
Guilt X Questionnaire	
Guilt X Sex	
Guilt X Time	
Relevant issue X Statement type	.037
Relevant issue X Sex	.070
Relevant issue X Time	.100
Statement type X Questionnaire	
Statement type X Sex	
Statement type X Time	
Questionnaire X Sex	
Questionnaire X Time	.071
Sex X Time	
Guilt X Relevant issue X Statement type	.037
Guilt X Relevant issue X Sex	
Guilt X Relevant issue X Time	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	.046
Guilt X Statement type X Time	.039
Guilt X Sex X Time	
Relevant issue X Statement type X Sex	
Relevant issue X Statement type X Time	
Relevant issue X Sex X Time	
Statement type X Questionnaire X Sex	
Statement type X Questionnaire X Time	
Statement type X Sex X Time	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Time	.029
Guilt X Statement type X Questionnaire X Time	

Table 20 Continued

Source	Effect Size
Guilt X Relevant issue X Sex X Time	
Guilt X Statement type X Sex X Time	
Guilt X Statement type X Questionnaire X Sex X Time	

Table 21
Effect Sizes for Blink Rate

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.053
Questionnaire (direct only)	
Sex	
Guilt X Relevant issue	
Guilt X Statement type	.036
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

Table 22
Effect Sizes for Next Item Blink Rate

Source	Effect Size
Guilt	
Relevant issue	
Statement type	.137
Questionnaire (direct only)	
Sex	.058
Guilt X Relevant issue	
Guilt X Statement type	
Guilt X Questionnaire	
Guilt X Sex	
Relevant issue X Statement type	.044
Relevant issue X Sex	
Statement type X Questionnaire	
Statement type X Sex	
Questionnaire X Sex	
Guilt X Relevant issue X Statement type	
Guilt X Relevant issue X Sex	
Guilt X Statement type X Questionnaire	
Guilt X Statement type X Sex	
Relevant issue X Statement type X Sex	
Statement type X Questionnaire X Sex	
Guilt X Relevant issue X Statement type X Sex	
Guilt X Statement type X Questionnaire X Sex	

REFERENCES

Barland, G.H. (1981). *A validation and reliability study of counter intelligence screening tests.* Unpublished manuscript. Security Support Battalion 902nd Military Intelligence Group. U.S. Army, Ft. Meade, Maryland.

Baumeister, R. F. (2002). Ego-depletion and self-control failure: An energy model of the self's executive function. *Self and Identity, 1,* 129-136.

Baumeister, R. F., Muraven, M., & Tice, D. M. (2000). Ego depletion: a resource model of volition, self-regulation, and controlled processing. *Social Cognition, 18,* 130-150.

Bradley, M. T., & Janisse, M. P. (1979). Pupil size and lie detection: The effect of certainty on deception. *Psychology: A Quarterly Journal of Human Behavior, 16,* 33-39.

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

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319-333.

Cook, A. E., Hacker, D. J., Webb, A., Osher, D., Kristjansson, S., Woltz, D. J., et al. (2012). Lyin' eyes: Ocular-motor measures of reading reveal deception. *Journal of Experimental Psychology Applied, 18*(3), 301-313.

Davis, R.C. (1961). Physiological responses as a means of evaluating information. In A. Biederman & H. Zimmer (Eds.), *Manipulation of human behavior* (pp. 142-168). New York: Wiley.

Dionisio, D. P., Granholm, E., Hillix, W. A., & Perrine, W. F. (2001). Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology, 38,* 205-211.

Ellson, D. G., Davis, R. C., Saltzman, I. J., & Burke, C. J. (1952). A report on research on detection of deception. Contract N6onr-18011 with Office of Naval Research. Bloomington: Department of Psychology, Indiana University.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30,* 725-745.

Ferreira, F., & Henderson, J. M. (1993). Reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology, 47,* 247-275.

Gray, J.A. (1987). *The psychology of fear and stress*. Great Britain: University of Cambridge. ILR School. (Posted by). (2009, January 30). *How do you know if people are lying on a Survey?* Retrieved March 23, 2013, from http://www.cornell.edu/video/surveys-how-to-know-when-people-are-lying.

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

Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research, 24,* 386-404.

Just, M.A., & Carpenter, P.A. (1993). The intensity of dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology, 47,* 310-339.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154,* 1583-1585.

Kircher, J. C. (1981, June). *Computerized chart evaluation in the detection of deception*. Master's thesis, University of Utah.

Kircher, J. C., Raskin, D. C., Honts, C. R., & Horowitz, S. W. (1994). Genereralizability of statistical classifiers for the detection of deception. *Psychophysiology, 31,* S73. (Abstract)

Krapohl, D. J. (2002). The polygraph in personnel screening. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 217-236). San Diego, CA: Academic Press.

Kuhlman, B. B., Webb, A. K., Patnaik, P., Cook, A. E., Woltz, D. J., Hacker, D. J., & Kircher, J. C. (2011, September). *Evoked Pupil Responses Habituate During an Ocular-motor Test for Deception.* Poster presented at the Society for Psychophysiological Research convention, Boston, MA.

Levine, T.R., Shaw, A., & Shulman, H.C. (2010). Increasing deception detection

accuracy with strategic questioning. *Human Communication Research, 36,* 216-231.

Loewenfeld, I. E. (1999). *The pupil: Anatomy, physiology, and clinical applications* (Vol. 1). Boston: Butterworth-Heinemann.

Lubow, R. E., & Fein, O. (1996). Pupillary size in response to a visual guilty knowledge test: New technique for the detection of deception. *Journal of Experimental Psychology: Applied, 2,* 164-177.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126,* 247-259.

Muraven, M., Shmueli, D., & Burkley, E. (2006). Conserving self-control strength. *Journal of Personality and Social Psychology, 91,* 524-537.

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29,* 894-906.

National Research Council. (2003). *The polygraph and lie detection.* Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10,* 241-255.

Ross, S.C., Millis, S.R., Bonebright, T.L., & Bailley, S.E. (2002). Confirmatory factor analysis of the behavioral inhibition and activation scales. *Personality and Individual Differences, 33,* 861-865.

Schwarz, N. & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication and questionnaire construction. *American Journal of Evaluation, 22,* 127-160.

Steller, M. (1989). Criteria-based statement analysis. Psychological methods in criminal investigation and evidence. In D.C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York, NY: Spring Publishing.

Stern, J.A., Skelly, J.J. (1984). The eye blink and workload considerations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 28,* 942-944.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, *10*, 141–142.

Webb, A. K, Honts, C. R., Kircher, J. C., Bernhardt, P.C., & Cook, A. E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. *Legal and Criminal Psychology, 14(2), 279-292.*