REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE:	3. REPORT TYPE AND DATES COVERED Final Report 1-Apr-2003 - 31-Mar-2006

4. TITLE AND SUBTITLE
A New Approach to Detecting Deception Using Learning Theory

5. FUNDING NUMBERS
DAAD190310042

6. AUTHORS
John T. Cacioppo

8. PERFORMING ORGANIZATION REPORT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Chicago
University Research Administration
970 East 58th Street
Chicago, IL 60637 -

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
45232-LS-DRP.1

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
The abstract is below since many authors do not follow the 200 word limit

14. SUBJECT TERMS
DECEPTION, LEARNING THEORY

15. NUMBER OF PAGES
Unknown due to possible attachments

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev .2-89)
Prescribed by ANSI Std.
239-18 298-102

Report Title

A New Approach to Detecting Deception Using Learning Theory: End of Project Report

ABSTRACT

The scientific literature on the detection of deception indicates that the use various physiological signals and testing approaches such as the guilty knowledge, or control question tests, yield results better than chance though lacking in sensitivity, specificity, and resistance to countermeasures (Committee to Review the Scientific Evidence on the Polygraph, 2003, "The polygraph and lie detection." Washington, DC: National Academy Press). Recent approaches that use brain imaging and other new technologies still rely on the emergence of a "natural lie response" that is presumed intrinsic to all people. While some people do intrinsically emit anxiety during deception, data do not support the ubiquitous nature of such a response.

While serving on the National Academy of Sciences Committee to review the scientific evidence for the validity of the polygraph, we developed an alternative analytic approach to the detection of deception. The approach differs from previous approaches in two fundamental ways. First, we proposed to use Pavlovian conditioning techniques to instill a unique but innocuous physiological response (e.g., a micro-eye blink) when they are exposed to an untrue statement. Second, we proposed to develop a sensitive and specific digital signal processing algorithm for each person individually based on the pattern (e.g., timing, frequency components, symmetry across the right and left ocular regions) of responses that best discriminated that individual's perception of a true (e.g., I kick a ball with my leg") versus untrue (e.g., "I kick a ball with my arm")statement. If no such response template is found, evidence is secured that one cannot test for deception. If signal detection analysis suggests a response template is apparent, this template is used to evaluate whether subsequent test items (e.g., "I was born in June") are true or untrue. (Test items are personally relevant questions for which we have ground truth.) NOTE A DETAILED WORD DOCUMENT HAS BEEN ATTACHED TO THIS REPORT

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Number of Papers published in peer-reviewed journals: 0.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Papers presented at meetings, but not published in conference proceedings (N/A for none)

Number of Papers not Published: 0.00

(d) Manuscripts

Number of Manuscripts: 0.00

Number of Inventions:

Graduate Students

NAME	PERCENT_SUPPORTED
Derek Tucker	No
FTE Equivalent:	
Total Number:	**1**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED	
FTE Equivalent:		
Total Number:		

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	
John T. Cacioppo	0.10	Yes
FTE Equivalent:	**0.10**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME
Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED	
George Monteleone	1.00	No
John Stockton Irick	1.00	No
FTE Equivalent:	**2.00**	
Total Number:	**2**	

Sub Contractors (DD882)

Inventions (DD882)

End of Project Report

Recent instances of international espionage and terrorism have renewed scientific interest in the physiological detection of deception (PDD). Among the approaches to PDD that have been proposed are physiological measures of lying, physiological correlates of lying (e.g., arousal, guilt, fear), and physiological indices of memory (Ben-Shakar & Furedy, 1990; Lykken, 1998)). These approaches employ a variety of measures including cardiovascular, respiratory, thermography, voice stress, eye movements, event-related brain potential, and functional magnetic resonance imaging (Iacono, 2000). In classic and contemporary approaches to PDD, however, the investigator relies on naturally occurring changes in physiology to mark the occurrence of a lie. Physiological measurements of the autonomic nervous system (e.g. traditional polygraphy), for instance, assume that a guilty but not an innocent individual will exhibit a larger increase in autonomic activity to a relevant question (e.g., a question about a specific crime) than to a control question (e.g., a question about a misdeed that almost everyone has performed) because only the guilty individual should be more apprehensive about denying guilt or knowledge in response to the relevant than control question. Approaches using event related brain potentials (Rosenfeld et al., 1991; Farwell & Donchin, 1991; Allen & Iacono, 1997) depend upon the expression of an intrinsic brain responses associated with the recognition of stimuli linked to an aspect of the event under investigation (e.g. crime scene or victim information), and fMRI approaches to PDD seek to identify a set of brain regions active during lying (Spence et al., 2000).

These diverse approaches to PDD share a common problem: large individual differences exist in the presence, profile, and magnitude of naturally occurring physiological (including brain) responses, and in the psychological responses to relevant and control stimuli and questions (Iacono, 2000). Consequently, the recorded physiological responses can occur for reasons other than lying. Activity in the anterior cingulate gyrus found in fMRI studies of lie detection (Spence et al., 2000), for instance, can also occur in many circumstances including when individuals are conflicted (Milham et al., 2001) or dysphoric (Gehring & Willoughby, 2002). Electrodermal and cardiorespiratory measures, voice stress, and facial temperature may be sensitive to potentially irrelevant factors such as evaluation apprehension, task demands, time

pressure, anxiety, and a myriad of other conditions. Despite reassurances by an examiner, some individuals may be more anxious about and physiological reactive to questions about their behaviour related to a coveted position or situation than about generic misdeeds. The implication is that even though a false statement may contribute more to the physiological response to relevant than control question or stimulus, the detection of a larger physiological response to the relevant question does not logically mean that the cause of the larger physiological response was a lie or false statement (Cacioppo & Tassinary, 1990). For this reason, the common practice of validating a PDD approach by demonstrating larger physiological responses in guilty than innocent individuals or responses is inadequate. Also inadequate is the approach of reporting only correct detections or the percentage of correct categorizations in a PDD study since such approaches may also lead to excessive false positives.

To be specific, we take as given that:

$$\text{Lying} = f(\Phi) \quad (1),$$

where Φ represents physiological (e.g., brain) activity.

Traditional experimental approaches and statistical tests focused on:

$$(\Phi/\text{lying}) \qquad (2)$$

whereas the goal of the physiological detection of deception is the:

$$(\text{lying}/\Phi) \qquad (3)$$

It is simple to show that:

$$(\text{lying}/\Phi) = (\Phi/\text{lying}) \qquad (4)$$

only when dealing with one:one relationships. Among the implications is that as baserate declines, the likelihood of a false detection increases, *ceteris paribus*. This is because:

$$P(\text{lying}/\Phi) = P(\text{lying},\Phi) / \{P(\text{lying},\Phi) + P(\text{not-lying},\Phi)\} \quad (5)$$

$$\text{or } P(\text{lying}/\Phi) = P(\text{lying},\Psi) / P(\Phi) \quad (6)$$

Whereas t-tests, analyses of variance, and multivariate discriminant analyses speak to (2), signal detection theory provides a formal means of examining (3). Moreover, the deployment of countermeasures and interactions between the examiner and examinee during the test generally (e.g., evaluation anxiety, social intimidation), and especially the use of PDD for both detecting deception and interrogation, can alter P(not-lying,Φ) in ways that are not known. Therefore, the criterion for success should not

simply be a statistically significant difference in physiological response between the expression of lies and the expression of truths, but:

1. A **unique** physiological response (e.g., frequency, amplitude, waveform, or response syndrome) associated with deception {i.e., P(lying,Φ) = 1 & P(not-lying,Φ) = 0}

2. A sufficiently **large** physiological response that it is detectable following individual items/questions, or a procedure that allows signal:noise enhancement (e.g., ensemble averaging, deconvolution) to measure P(Φ/lying)

3. A physiological response that is not subject to voluntary motor or mental control – that is, it is **insensitive to countermeasures** {i.e., P(not-lying,Φ) = 0}

4. Either a physiological response that is **invariant** across individuals or a procedure for identifying/developing a large and unique involuntary physiological response for each examinee {i.e., an invariant response for a given individual)

5. **Standardized** examination for the exclusive purpose of detecting deception that minimizes extraneous influences of the examiner (computer-human interface)

6. A **quantitative** evaluation of the quality of signal detection for each examinee (signal detection theory), including the ability to specify a test as inconclusive for known reasons (e.g., examinee fails to correctly perform the designated task, examinee invokes somatic countermeasures, invariant response not identifiable for that examinee)

7. A procedure that can be tested and implemented in **field settings** (e.g., war games)

Additional desiderata might include the following:

1. Known **neurogenic** control of the response (e.g., baroreceptors, reflexive eyeblink, neurobiological circuit underlying lying)

2. Phasic response can be sculpted to have more unique **temporal response curve**

3. **Bidirectional conditioning** is possible

4. An array of physiological measures/parameters with which to develop an **idiographic** discriminant function using an adaptive decision algorithm

We have been pursuing research on the physiological detection of deception that would fit these criteria. We were unsuccessful, though not uninformative, in this effort. Specifically, we investigated four different response systems: peripheral vasomotor activity, baroreceptor activity, startle eyeblink, and hemodynamic responses of the brain using functional magnetic resonance imaging (fMRI). We summarize briefly our investigations in each of these domains.

Vasomotor Activity

Our initial effort was to demonstrate an alternative approach to PDD that instils a physiological response specific to information known to be false to the subject. Our approach uses Pavlovian conditioning to create patterns of autonomic responses that would never occur naturally and pairing these unique physiological responses to true and false statements in a series of conditioning trials, where a trial is defined as the presentation of a conditioned stimulus (true/false statement) followed by an unconditioned stimulus. In the vasomotor studies, we induced vasomotor changes by heating the right and cooling the left index finger upon the presentation of false statements and reversing contingency during the presentation of true statements. Following a series of conditioning trials, test (unreinforced) trials are interspersed among conditioning trials to allow an assessment of the veracity of the test statements while minimizing extinction. Participants exposed to these conditioning procedures were hypothesized to exhibit the same vascular changes on test trials where true or false statements are presented but no heating or cooling is introduced.

Twelve male college students served as subjects. Following obtaining informed consent and the completion of several questionnaires, two conditioning interfaces were attached to the left and right index fingers of the subjects using hook and loop fasteners. These conditioning interfaces consisted of photoplethysmographs integrated into thermoelectric cooling devices (see Figure 1). These devices and their computer interface allowed control of the heating and cooling of the fingers during the presentation of text information on a computer screen.

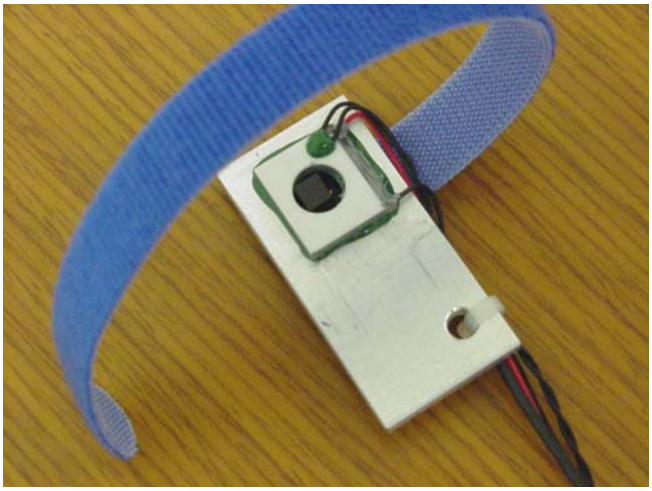


Figure 1. Photograph of the conditioning device. The photoplethysmograph is in the centre of the circular opening of the thermoelectric cooler. Both are attached to aluminium backing that acts as a heat sink. Attachment to the subject is made with a hook and loop fastener.

The conditioning procedures were presented in two phases. The first phase required that subjects view a series of statements on the computer screen. The statements were presented in two parts such as "I don't like being" followed by the word "honored" and "I support" followed by the word "terrorism." Four seconds after the word completing the sentence appeared, subjects were instructed to say if the completion was true or false. During the presentation of the sentence completion but prior to the verbal response (conditioned stimuli - CS), temperature stimulation (unconditioned stimuli - UCS) was applied on 80% of the trials. Thus, the perceived veracity of the word completion was conditioned to the temperature change, not the verbal response. In the second phase, the ratio of temperature-reinforced trials was reduced to 50%. Additionally, true and false sentence completions were added that had not been previously introduced to assess the effect of the conditioning trials on novel stimuli. Forty trials were presented in each phase. Vasomotor responses were continuously recorded from the two plethysmographs (unconditioned and conditioned responses - UR and CR) and the record was marked by the computer administering the textual stimuli to indicate the presentation of a completion and whether it was true or false.

The signals from the plethysmographs were normalized, filtered, and the responses from the left and right fingers subtracted. Such processing reduced the oscillatory activity due to cardiac output and potentiated the differences between the two fingers providing for the clear comparison of vasodilatation/constriction associated with false sentence completions versus the vasoconstriction/dilation associated with true sentence completions. Plethysmograph responses of the 12 subjects during the training trials where temperature inductions were associated with true and false statements (CS /

UCS pairings) can be seen in the left panel of Figure 2. The right panel contains the average response of the twelve subjects to previously unconditioned true and false sentence completions when no temperature changes were induced. As can be seen in this panel, responses are similar, though diminished, to the training trials. Statistical comparison of these waves indicated that they differed significantly from each other (p=0.01). These findings demonstrate the feasibility of conditioning unique physiological responses to true or false statements presented to the subject and establish the possibility of using such procedures to detect an individual's beliefs about the veracity of such statements.

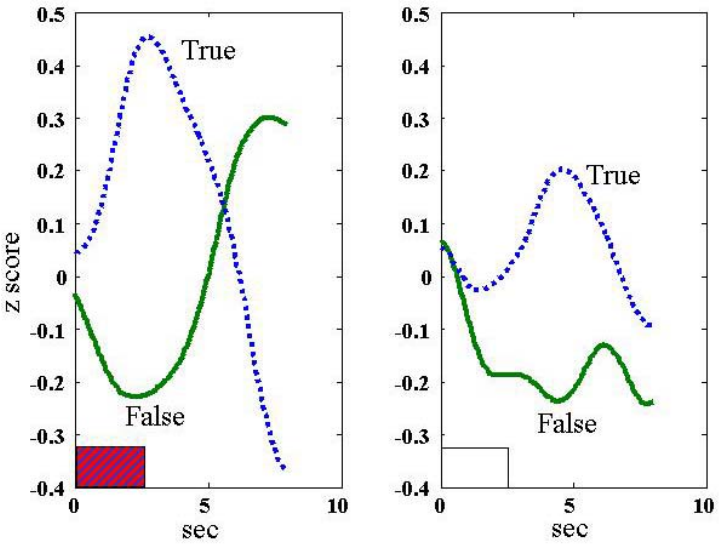


Figure 2. Averaged vascular responses from all subjects during conditioning (left panel) and testing (right panel). Each line represents the difference between the heated and cooled fingers. Timing of heating and cooling during the conditioning trials is shown by the block in the lower left. During testing, no heating or cooling took place and subject saw novel stimuli that were either true or false.

As noted above, however, the physiological differentiation of true and false statements does not necessarily imply that these physiological responses will provide a sensitive and specific marker of the veracity of statements.[10] For instance, although the conditioning procedures described here produced responses that differentiated true and false statements, they are not necessarily diagnostic. It is possible to be 100% accurate in finding falsehoods in this study by setting selective criteria for this outcome. Doing so however, leads to nearly 100% false positives as well. This problem is well known is sensory psychology and has led to the development of signal detection theory. This theory provides for a common language and formulae for describing the relationship between categorical decisions about distributions of data. For data such as these, a statistic comparing the correct detection of false statements to the false positives at various decision criteria (d'[11]) is most appropriate. A nearly perfect

testing instrument will yield a *d'* of approximately 4.0 and an instrument working at a random level will yield a *d'* of 0.

Techniques to reduce individual trial variability, such as aggregating similar individual trials, have been used in event related potential and functional magnetic resonance imaging studies to improve signal quality. Aggregation of trials in the same condition for each subject in the present study improved classification, as well, leading to the correct detection of 10 of 12 subjects for false statements and 9 or 12 for true statements. In terms of signal detection theory, the false positive rate was 25% (*d'*=1.64). It may be surprising that an effect can emerge at a probability of p=0.01 yet lead to modest predictive power. This is due to the fact that individual trials in the distribution which are proportionally more difficult to classify when variability is high.

One possibility is that, just as in event related potential research, averaging over repeated presentations of the same stimulus item (e.g., sentence completion) can improve signal detection. Subsequent data collection revealed this to yield only minimal improvements. We found a major limitation to using peripheral vasomotor activity is that peripheral vasomotor activity is under only limited central neurogenic control. We were not able to find a conditioning procedure that permitted effective classical conditioning of the vasomotor response in most of the subjects who were tested, and even when conditioning was achieved the signal discrimination continued to be no better than extant procedures.

Baroreceptor Response

We next considered classically conditioning the baroreceptor reflex because it is an autonomic response that is under tight central neurogenic control. The stimulation of the baroreceptors requires applying positive and negative pressure to both sides of the neck over the carotid sinus (see Figure 3).

Fabricated Paired Neck Chamber And Pressure/Vacuum Delivery Device

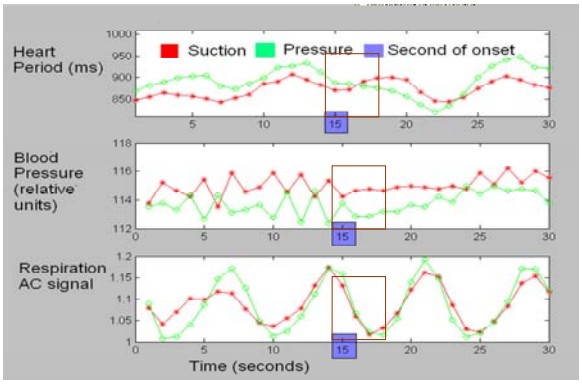
Figure 3. Equipment for baroreceptor conditioning.

Results indicated the expected cardiovascular

unconditioned responses to the application of the unconditioned stimuli (sunction, pressure; see Figure 4). There were two major limitations that we encountered early. First, there was a small risk that atherosclerotic build-up in the carotid sinus could be dislodged by the application of the unconditioned stimulus, placing the subject at risk for a stroke.

Second, the conditioning procedure was compromised by the fact that the application of positive pressure to the neck created the sensation of being strangled (see Figure 5). Piloting also suggested that the strength of the unconditioned stimuli would need to be intense to have reliably measurable effects on conditioned cardiovascular responses. After consultation with our sponsors about these limitations, the decision was made to not pursue the conditioning of baroreceptor responses but instead to focus on the startle blink, for which there is an experimental literature in the field of human classical conditioning.

Response After 0.6 s Stimulation.
One subject 9 trials.



Startle Eyeblink (Tucker, 2005)

Nineteen right-handed male participants, ages 18-25 (mean age 20.83 years), in good physical health and fluent in English were recruited from the University of Chicago (Tucker, 2005). Participants' task followed written informed consent, and their entire time in the lab was approximately 2 hours. Demographic information including recent alcohol, nicotine, herbal and prescription medication, history of illness and injury, and history of familial disorders was taken. Participants were compensated at the rate of five dollars per half hour for their participation in this study.

The UCS was a 5 psi air-puff that lasted 75 ms. and was delivered near the lateral corner of their left eye. The participants were fitted with safety goggles with a hole drilled through the protective glass through which tubing was attached. There were various levels of holes available to properly place the air puff for differently sized faces (see Figure 6).

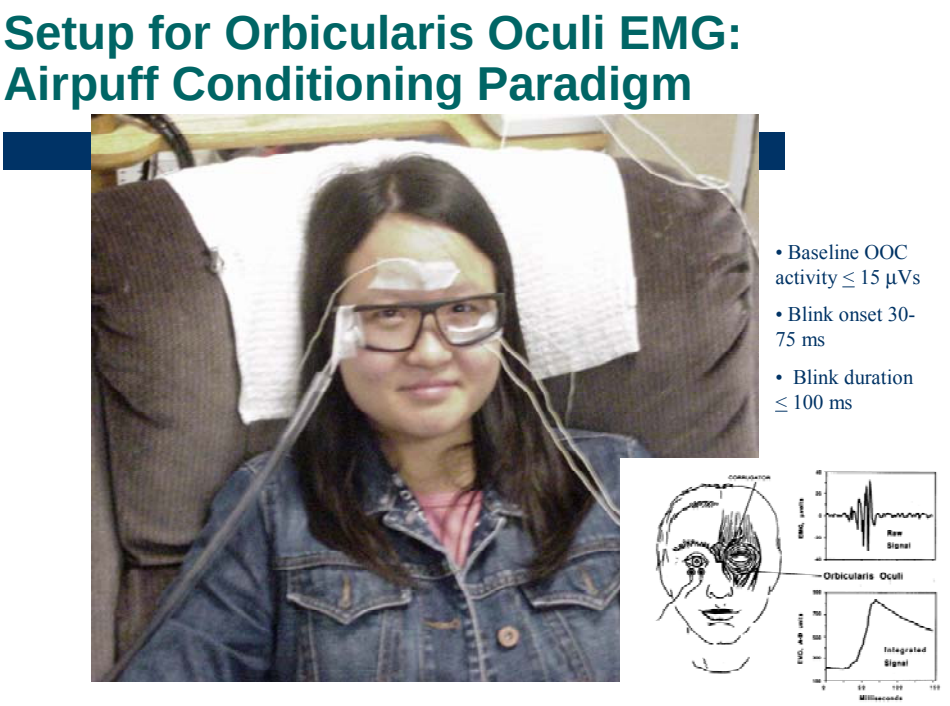


Figure 6.

The Conditioned Stimuli (CS) were the veracities of a statement which followed the final word that completed an obviously true or false statement. Each statement was delivered through headphones as a digitally recorded sound file. A fully representative sample of statements was presented to each participant prior to the procedure to verify agreement on the assumptions of being true or false. Each statement was presented in a two part, stem-and-completion format. The stem, (e.g.: "When heated melts") has no truth value on its own, it depends on the completion for its meaning. Two completions were designed for each stem, one true and one false (e.g.: "When heated melts," "ice," and, "wood"). Analogously, each completion had a paired stem such that each completion could either be true or false (e.g.: "When heated *burns*," "ice," and, "wood"). This way, the veracity of the statement could not be determined by either the stem or completion alone, and required semantic understanding by the participants. Moreover, the UCS was associated with neither the stem, nor the completion, but the abstraction (a false rather than true statement: the "AB+, CD+, AD-, CD-," design is that of biconditional discrimination (Lober & Lachnit, 2002). To further underscore this contingency, a screen appeared after the completion showing "TRUE" and "FALSE" in a green and red box, respectively, with the side of appearance (right or left) for each box varying randomly. The left most box corresponded to the "F" key, and the rightmost box corresponded to the "J" key. Participants were to respond by pressing "F" or "J" corresponding to the left or right appearance of the correct response, "True" or "False"

which varied randomly. The response screen randomization restrictions protected against more than 3 consecutive "F" or "J" responses, thus avoiding a possible confound by preserving the requirement of attention to predict accuracy.

Experimental control, visual, and audio presentation were performed by a custom program developed in the E-Prime environment from Psychology Software Tools, Inc., on a PC running the Windows 98SE Operating System. Eye movement and blink signal were measured using EMG over the right (contralateral to the air-puff) orbicularis occuli (OOC), and VEOG on the left (ipsilateral to the air-puff) eye. The tubing in the safety goggles fed back to a custom built computer controlled system to calibrate and time the air flow. Output of the EMG, VEOG, and air puff mechanism was recorded on a second Windows 98SE PC through the Acknowledge© program, and their veracity judgment was stored as a text file through E-Prime.

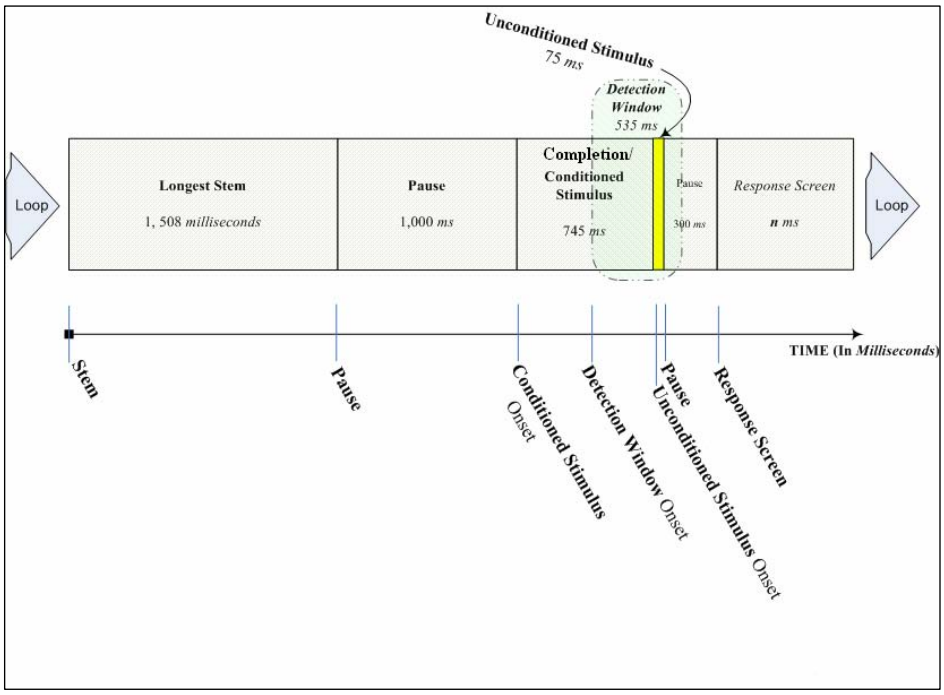
We incorporated a bi-conditional discrimination eye-blink conditioning paradigm

Phase	**Number of Trials**	**Design of Stems**	**Design of Trials**
Adaptation	4 true statements	4 stems from 1-12	All CS-
Training	120 (60 true stimuli, 60 false stimuli)	12 stems (1-12)	90% False reinforced
Assessment	60 (30 true stimuli, 30 false stimuli)	6 stems (3, 4, 7, 8, 9 & 10)	50% False reinforced
Test	120 (60 true stimuli, 60 false stimuli)	12 stems (old stems 2-11; critical stems 13 & 14)	50% False reinforced

using the abstraction of statement veracity, or whether it is "true" or "false," as the differentiating variable, into a four phase procedure: Adaptation, Training, Assessment, and Test. Each phase consisted of a specified number of trial stimuli involving a stem + completion pair whose combination would either be of true or false veracity. The conditioned stimulus was the veracity of the statement, recognized subsequent to the completion's presentation. The unconditioned stimulus (UCS: air-puff) was delivered

745 ms. following the onset of the completion, which never overlapped or preceded the completion's articulation (see Figure 6).

A "Response Screen" appeared 300ms following the termination of the time slot for the air-puff, indicating which button to press to indicate "true" or "false". The participant's response to the statement was included as a potential method to assess participant attention throughout the experiment, but the response itself is theoretically unimportant for the conditioning procedure. The "Response Screen" stayed on until the participant responded. An inter-trial interval of eight to fourteen seconds with a mean of 12 seconds, was randomized throughout the experiment, all but the last second of which



was filled with music. True stimuli that were temporally correlated over the course of the experiment with false stimuli that received an air-puff were termed "CS-*pUCS*" for analysis purposes. False stimuli that were reinforced were identified as, "CS+UCS" False stimuli that were not reinforced were identified as "CS+" while true stimuli temporally correlated with them were identified as "CS-."

During the Adaptation phase, participants were presented with four true statements to assess baseline eye-blink responses. During the Training phase, 120 stimuli were presented, 60 true and 60 false. According to Levond and Steinmetz (2002), humans usually take between 25 and 50 trials in order to learn the association between a CS and US. To be conservative, given the Trace and semantic nature of the procedure, we assumed it would take 60 trials to create a reliable response. Therefore, since each

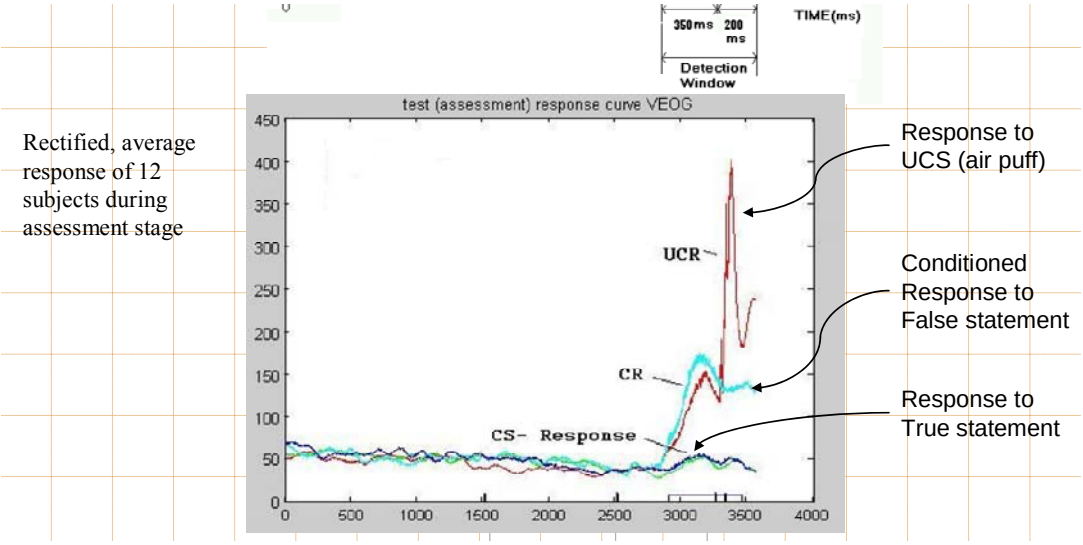
CS+ trial was paired with a CS- trial, 120 trials were continuously presented in the Training Phase. Of the false stimuli, 90% (54) were reinforced with an air-puff. During the Assessment phase, 60 stimuli were presented, 30 true and 30 false. Of the false stimuli, 50% (15) were reinforced. The Testing phase was designed to appear as two Assessment phases, 120 stimuli, 60 true, 60 false, 30 false stimuli were reinforced. The additional element in the Test phase was that critical items that had never before been seen during the procedure, and were never reinforced, were embedded, thus allowing an analysis of the detection of the veracity of the statement. The "true" critical stimuli were the stem + completions: "Used for hugging – arm," and "Used for kicking – foot." The "false" critical stimuli were "Used for hugging – foot," and "Used for kicking – arm". The stem, "Used for hugging," is referred to in the analysis as stem "M," or, "13," while, "Used for kicking," is referred to as stem "N," or, "14." The Dependent measures were the presence of a blink and its' magnitude, as detected by vertical electro-oculogram (VEOG) on the eye that was being presented with the UCS, and electromyograph (EMG) on the eye contralateral to the air-puff.

Participants were greeted upon arrival at the lab by the experimenter who brought them into the room where preparation for EMG and VEOG recording was administered. After informed consent was obtained and demographics completed, participants were read a description of their task while the electrodes were being placed. The electrodes were affixed as described in Tassinary & Cacioppo (2000). The impedance between the electrodes was verified as less than 5 kiloohms. Next, participants were seated in an overstuffed chair in the testing chamber, reclined to 45°, fitted with the air-hose safety goggles, head phones, and given a cup of water. They were also given a remote keyboard on which they were to make the judgments to the statements. The experimenter then summarized the experimental procedure, reminding participants that air puffs would only be received if the completion following the stem lead to a false statement. That is, the conditioning contingency was made explicit to participants.

The experimenter was located in a separate control room monitoring the participant during program execution. Next, an automated experimental control presentation program was started that was used to verify a good signal from the EMG and VEOG, and asses normal blinks from the participant. After verification of setup, an

adaptation program using 4 true stem-completion pairs was used to accustom the participant to the environment, and to make sure they had full comprehension of the task. Any questions or concerns by the participant were addressed, and the Training – Assessment – Test sequence was initiated. The participants were allowed a break between each phase, and also between 60 trial segments of the Test phase.

Illustrative results from the study are depicted in Figure 7.



During the Training phase, there was a main effect of the air-puff on VEOG response ($F[1,11] = 252.17$ $p<0.001$) which remained through the Assessment phase ($F[1,11] = 484.80$ $p<0.001$) and persisted through the Test phase ($F[1,11] = 1{,}001.19$ $p<0.001$) (Tucker, 2005). These results indicate the effectiveness of the air-puff in eliciting a differential blink response from no air-puff.

Using a paired samples two tailed t-test for each phase, the %CR's (both raw and corrected for Assessment and Test) were compared between the EMG and VEOG measures of eye-blink responding. Three comparisons were significantly different at the 0.05 level, and one approached significance. Each significant difference was found in a CS+ condition for critical trials with the VEOG score reporting a greater %CR than EMG. The critical pool raw CS+ VEOG measurement (M=74.17% SE=3.98%) was greater, $t(11) = 2.374$ ($p=0.037$) than the same EMG measurement (M=65.58% SE=4.82%). The raw critical question "N" CS+ VEOG measurement (M=78.33% SE=6.26%) was greater, $t(11)=2.219$ ($p=0.048$), than the same EMG measurement (M=62.92% SE=7.32%). The critical pool corrected CS+ VEOG measurements'

(M=46.58% SE=7.61%) difference from the EMG measurement (M=31.42% SE=7.48%) approached significance t(11)=2.031, (p=0.067). The corrected critical question "N" CS+ VEOG measurement (M=58.17% SE=28.33%) was greater, t(11)=2.353 (p=0.038) than the corresponding EMG measurement (M=28.33% SE=9.75%).

Each of the aforementioned significant differences in the %CR between VEOG and EMG measures of the identical trials are driven by the difference between VEOG and EMG in their responses to the statement, "Used for kicking: Arm," or the CS+ for critical question, "N." That is, the critical pool difference noted above was only present as a consequence of the difference present in false (CS+) "Used for kicking" instantiation, and not its' true counterpart ("Used for kicking: foot"; t[11]=0.057 p=0.956), or either the true or false instantiation of its' companion stimulus ("Used for hugging," "arm/foot" t[11]=0.958 p=0.359 and t[11]=-.178 p=0.862 respectively). Since the VEOG was greater in each of these differences, this suggests that the neuro-cognitive processes underlying the realization that "Used for kicking: Arm" is a false statement requires more executive processes than the other critical statements which did not show a measure difference.

During the Training phase, false statements showed significantly greater, t(11) =5.191 ($p<0.001$) %CR's (M=74.83% SE=6.69%) than true statements (M=26.42% SE=6.644%). The raw Assessment items also showed this difference, t(11)=10.225 ($p<0.001$), with false statements showing greater %CRs (M=67.50% SE=4.86%) than true statements (M=17.25 SE=2.71%). A similar pattern was demonstrated for the corrected Assessment items, with the %CRs to false statements (M=66.50% SE=3.79%) being greater, t(11)=8.858 ($p<0.001$), than true items (M=24.75% SE=3.51%) . These results indicate successful differential conditioning to statement veracity.

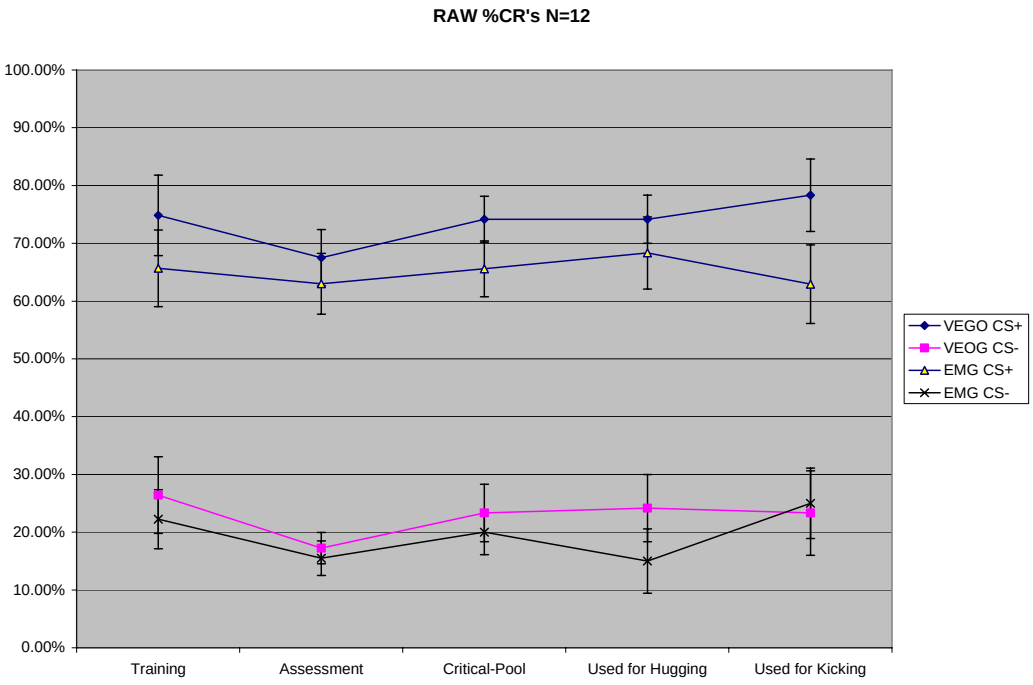


Figure 7: Raw %CR's Across Experiment including all participants

The novel false items that were presented as critical statements in the raw Test phase elicited greater VEOG, t(11)=7.249 ($p<0.001$), %CR's (M=74.17% SE=3.98%) than the novel true items that were presented as critical items (M=23.33% SE=4.98%). This was also true for the EMG, t(11)=5.902 ($p<0.001$), with false items showing an average of 65.58% CR's (SE=4.83%), and true items showing an average of 20.00% (SE=3.98%).

Using the corrected %CR's as a measure of the differential responses to the critical statements in the test phase, the same result pattern was demonstrated. False statements with the corrected VEOG measure generated a %CR (M=46.58& SE=7.61%) greater (t[11]=4.577 p=0.001) than true statements (M=12.17% SE=3.93%). The same was true for the corrected EMG in that the %CR to false statements (M=31.42% SE=7.48%) was greater (t[11]=3.095 p=0.01) than the %CR to true statements (M=9.5% SE=2.67%). These results indicate the successful generalization of the differential CR elicitation to novel presentations of statement veracity.

For the critical item, "Used for Hugging," the raw VEOG %CR for the false (CS+) completion "Foot" (M=74.17% SE=4.17%) was greater than (t[11]=6.770 $p<0.001$) the corresponding %CR for the true (CS-) completion "Arm" (M=24.17 SE=5.83) with the EMG result being virtually identical. For the same corrected VEOG statement, the CS+ %CR (M=36.08% SE=11.45%) was somewhat greater than (t[11]=6.770 p=0.051) the CS- %CR (M=10.83% SE=5.14%). The corrected EMG

measure for the same statement showed that the %CR to the CS+ (M=37.42% SE=10.29%) was significantly greater (t[11]=3.850 p=0.003) than the %CR to the CS- (M=5.00% SE=2.61%).

The corrected EMG %CR's were significant at p<0.005, while the VEOG %CR's merely approached significance at p=0.051, even though there was no significant difference between the two measures (see results section 2) in their reports for the %CR to CS+ (t[11]=-0.178 p=0.862) or to the CS- (t[11]=0.958 p=0.359). Possible insight toward the etiology of this measure difference in differential responding is the finding that the correlation between the VEOG and EMG measures for the corrected responses to the CS+ statement "Used for Hugging: Foot" were significantly correlated (r=0.768 p=0.004), while the VEOG and EMG measures for the corrected responses to the CS- statement "Used for Hugging: Arm" were not correlated (r=-0.141 p=0.662). Even when removing the participants that did not show differential conditioning (see section 7), the disparity in correlation of VEOG and EMG for the CS+ and CS- remained (r=0.785 p=0.007; r=-0.014 p=0.969 respectively). The relevant correlations for the raw scores showed a similar pattern, with the raw CS+ VEOG and EMG %CR correlation achieving significance weaker than the corrected score (r=0.635 p=0.027) and the CS- correlation remained non-significant (r=0.384 p=0.217). These results suggest that the statement "Used for hugging: Foot" is neuro-cognitively tightly bound to being a false statement, and the CR was likely of the C-type. Further evidence supporting this claim is that the corrected EMG differential conditioning (p=0.003) was more effective than the corrected VEOG differential conditioning (p=0.051).

For the critical item, "Used for Kicking," the raw VEOG %CR for the false (CS+) completion "Arm" (M=78.33% SE=6.26%) was greater than (t[11]=4.750 p=0.001) the corresponding %CR for the true (CS-) completion "Foot" (M=23.33% SE=7.31%). The same was true for the raw EMG measure of this statement: the CS+ %CR (M=62.92% SE=6.81%) was greater than (t[11]=3.796 p=0.003) the %CR to the CS- (M=25.00% SE=6.09%). The same was true for the corrected VEOG measure: the %CR to CS+ (M=58.17% SE=9.81%) was greater than (t[11]=3.894 p=0.003) the %CR to CS- (M=15.42% SE=5.49%). However, for the corrected EMG responses for this statement,

the %CR to CS+ (M=28.33% SE=9.75%) failed to obtain a significantly greater response (t[11]=1.267 p=0.231) than the %CR to CS- (M=15.00 SE=4.48%).

The corrected EMG differential %CR's for the stem, "Used for kicking" did not achieve significance while the corresponding corrected VEOG differential CR's did, and that the opposite trend was shown for the corrected VEOG and EMG responses for the stem, "Used for hugging." Furthermore, while the correlation between the corrected EMG and VEOG %CR responses to the false (CS+) statement, "Used for hugging: foot" was significant (r=0.768 p=0.004, see table 4), the analogous correlation for the statement, "Used for kicking: arm," failed to achieve significance (r=0.160 p=0.619), and similar to the CS- counterpart for, "Used for hugging," the corrected EMG correlation with corrected VEOG %CR's did not show a significant correlation at the 0.05 level (r=-0.069 p=0.830). In contrast, the raw VEOG EMG correlation for the statement, "Used for kicking: foot" did show a significant correlation (r=0.646 p=0.023) while the raw VEOG EMG correlation for the statement "Used for kicking: arm" was less strong (r=0.437 p=0.155).

This evidence suggests that even though conditioning was successful, at the group level, the specific questions differed in their concordance with the differential conditioning seen at the nomethetic level. Data from this study were therefore sent to Scott Arouh for signal processing to determine whether an independent and disinterested investigator could identify reliable conditioned responses to the CS. A time series approach was used to identify conditioned responses. The detailed report is provided in the filename Eyeblink_Detection_Results_Addendum2. Briefly, nine out of twelve subjects showed reasonably good discriminant conditioning.

The largest limitation, however, is the voluntary control subjects have over their skeletomuscular system. For instance, the successful differential conditioning was characterized by a V-type response. This response suggests subjects were closing their eyes prior to the expected air puff to avoid the irritation to the eye. This suggests subjects could avoid closing their eye if they were instructed to deceive the experimenter. Subsequent studies confirmed this concern. Subjects could easily inhibit or mask the eyeblink when they sought to maintain secrecy about their lying despite the variations in conditioning and measurement that were applied.

Brain Response

We first sought to determine the neural correlates of eyeblink conditioning. Nine subjects underwent eyeblink conditioning and nine did not. All eighteen subjects then underwent an fMRI study in which they responded to true and false statements following the procedures outlined above except unconditioned stimuli were not used. No group differences in brain activity were found, which suggests little persistence or generalization in the conditioned eyeblink response.

Others have investigated the neurobiological substrates of lying using fMRI under the assumption that measures of the underlying neurobiology would overcome some of these limitations. It is naïve to think that the brain's response is not under voluntary control. Sensory cortices can be attuned to stimuli to which you might wish to attend, the voluntary control over motor responses are mediated through the control of inputs to the motor cortex, and much of mentation and emotion are classified as "controlled" processes because people can exert voluntary control over these operations. To the extent that lying, and/or the deployment of countermeasures, is under intentional control, we might expect subjects to be able to alter or mask many of the neural responses associated with lying. Moreover, careful analyses of whether fMRI can be used to classify truth and lies are needed. We began with the latter task.

Specifically, prior research has provided fMRI evidence for neural activation related to deception in VLPFC, DLPFC, MPFC, MSFG, and STS. In an illustrative study, Phan et al. (2005) reported these areas of activation in nomethetic analyses of 14 Ss who were given a modified version of the Guilty Knowledge Test. Using the same dataset, we approached the question of classifying individuals as guilty based on their neural responses related to deception. The classification algorithm was based on nomethetic maps made from a Lie-Truth response contrast based on 13 of the 14 Ss, which were then used as ROIs for predicting the Lie-Truth contrast of the remaining subject. This analysis was iterated 14 times, once for each subject. Functional ROIs were obtained at both the group and individual levels by applying individual voxel thresholds with a clustering criterion of five contiguous voxels.

As reported in Phan et al. (2005), fourteen healthy, right-handed volunteers (7 males and 7 females; mean age, 32 years; age range, 23–48 years) participated in the

fMRI study. All participants were recruited on a volunteer basis, without monetary or other compensation, and no reward was given for their task performance. All subjects were without a history of head injury, learning disability, or neurologic or psychiatric illness, as verified by a semi-structured clinical interview modified from the Structured Clinical Interview from the Diagnostic and Statistical Manual of the American Psychiatric Association, 4th Revision (DSM-IV) (16), and had normal or corrected-to-normal visual acuity.

The study design was adapted from the "high-motivation" GKT task using playing cards described by Langleben and colleagues (2002). At the start of the experiment, before scanning began, each subject received the task instructions and was shown the workstation that would be used to analyze the subject's fMRI data in real time, using the TurboFIRE software (Phan et al., 2004). Example scans of previous participants made during the task were displayed on the work-station screen, and subjects were informed that their brain activation would be monitored by the research team while they performed the task in the scanner. Although we used TurboFIRE to monitor brain activation in real-time, the number of trials conducted in this pilot study did not have adequate statistical power for formal data analyses. Subjects were given a response pad and told that their button-press responses would also be monitored while they performed the task in the scanner. In order to make the task simulate a "real-life" experience, each subject was given two playing cards—the 5 of Clubs (5♣) and the 2 of Hearts (2♥)—and was asked to briefly study these cards and then place them in the subject's pocket for the duration of the scan. Subjects were told that they would be asked to lie about possessing one card and to tell the truth about the other, indicating their responses by button-pressing (thumb = "No", index finger = "Yes"); this assignment was counterbalanced across subjects such that half were instructed to lie about the 5 of clubs and half were instructed to lie about the 2 of hearts. This 2-card design was implemented so that the subject, when asked about a card in the subject's possession, had to make a Yes/No decision, without any object-recognition or card-specific (ie, color or number) effect. While in the scanner, subjects were presented with playing cards as separate events within four different categories of cards/events which prompted four different responses: 5♣ (lie/truth), 2♥ (truth/lie), 10 of Spades (10♠; control), and random cards from the rest of the 49-card

deck (non-target responses). Screens with the lie, truth, and non-target cards were accompanied by the question, shown above each card, "Do you have this card?" while the screen for the control card carried the question, "Is this the 10 of spades?" The control and non-target cards were intended to promote alertness and attention to the task and to minimize repetition of the lie-truth cards, while the inclusion of the control card forced subjects to read the question posed above all cards rather than provide indiscriminate, automatic "No" responses. For example, if a subject was instructed to lie about the 5♣, then the correct responses for each card type would be as follows: 5♣ = No; 2♥ = Yes; and 10♠ = Yes. Cards other than the 5♣, 2♥, or 10♠ were to be given "No" responses.

On each imaging run (of 2 total runs), subjects saw randomized presentations of 38 separate trials of lie, truth, control, and non-target cards. Each card was presented for 8 seconds, followed by an 8-second interstimulus interval during which the reverse side of the card was shown. Stimuli were presented via MR-compatible LCD goggles (Resonance Technology Inc., Northridge, CA), and button-press responses were recorded using Presentation software (Neurobehavioral Systems, Inc., Albany, CA). It should be noted that in contrast to the task developed by Langleben and colleagues (2002), the subjects in our study had actual possession of the test cards, were told to lie about either the 5♣ or 2♥ and received no financial reward or punishment for their performance. They were told that a research investigator blinded to the assignment of truth/lie cards would monitor the accuracy of their button-press responses and their brain activity with real-time fMRI technology (TurboFIRE). In our attempt to simulate a polygraph-like environment, we told subjects that their performance and brain responses were being monitored closely during the course of the experiment.

The subjects were scanned with a 4-T MedSpec MRI scanner (Bruker, Ettlingen, Germany) on a Siemens Syngo platform (Siemens Medical Systems, Erlangen, Germany) with a standard RF coil. After a T1-weighted, high-resolution anatomical scan, fMRI data were acquired through single-shot multi-echo echoplanar imaging (EPI) with 7 evenly spaced TEs ranging from 11–78 ms (TR = 2000 ms; FOV = 192 mm; 32 x 32 matrix; 16 slices; 6-mm slice thickness; 0.6-mm slice gap; flip angle = 90°) (Posse et al., 1999). Slices were oriented axially or nearly axially along the AC-PC line at the level of the amygdala.

Data sets from all 14 subjects met our criteria for high quality and scan stability with minimum motion correction (< 2 mm displacement in any one direction), and were subsequently included in fMRI analyses. Image processing and data analysis was done with the statistical parametric mapping software package SPM99 (Wellcome Department of Cognitive Neurology, London; www.fil.ion.ucl.ac.uk/spm). Standard pre-processing was applied, comprising slice-time correction, realignment, and spatial normalization to the Montreal Neurological Institute (MNI) high-resolution T1 template. Images were resampled into this space with 2-mm isotropic voxels, and were smoothed with a gaussian kernel of 6 mm full-width at half-maximum to minimize noise and residual differences in gyral anatomy, resulting in an effective spatial resolution of 12.8 × 14.4 × 14.9 mm. Each normalized image was bandpass-filtered (high-pass filter = 32 seconds) to remove low-frequency noise.

For the statistical parametric mapping (SPM) analysis, a general linear model was applied from which statistical inferences were based on the theory of random gaussian fields, and changes relative to the experimental conditions were modeled by convolution with the canonical hemodynamic response function (HRF) in order to approximate the activation patterns (Friston et al., 1995). Statistical parametric maps (SPMs) representing the association between the observed time series (eg, blood-oxygenation-level-dependent [BOLD] signal) and one or a linear combination of the regressors were generated for each subject. Within-subject contrasts were derived for brain activity related to the following comparisons: lie > truth, lie > control, truth > lie, and truth > control. These contrast images were then entered into a one-sample t-test across the 14 subjects in a second-level, random-effects analysis to allow for inferences applying to the general population (Holmes & Friston, 1998). This produced statistical parametric maps of the t statistic at each voxel, which were subsequently transformed to the Z distribution. From voxel-wise comparisons, activation foci were considered significant in regions in which we had an a priori hypothesis (ACC, MPFC, DLPFC, VLPFC), and whose activation surpassed a height threshold of $P < .001$ uncorrected ($t > 3.85$), with an extent of at least 5 contiguous voxels. These thresholds are commonly applied in the literature, and were intended to strike a balance between rates of type I and type II error. Reported activations outside these a priori regions had to exceed a threshold of $P < .05$, corrected for multiple

comparisons. Results revealed deceptive responses were specifically associated with activation of the VLPFC, DLPFC, DMPFC, and superior temporal sulcus (STS).

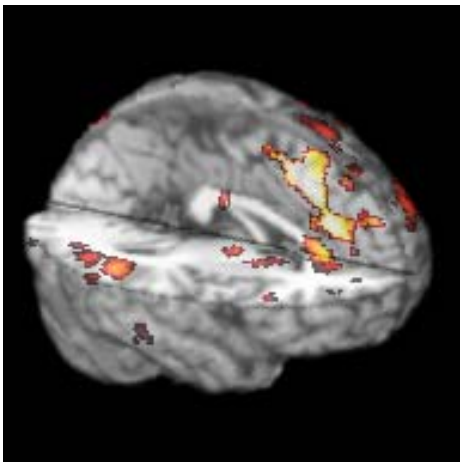
Data were re-analyzed to make individual-subject predictions from group response data. Preprocessed data from Phan et. Al. (2005) were converted from ANALYZE (spm99) to a format for use with the analysis software AFNI (Analysis of Functional Neuroimages). A canonical hemodynamic response function was convolved with the experimental conditions using the AFNI tool WAVER, and this model was regressed against the experimental data at each voxel to provide a within-subjects statistical maps of the responses for the lie > truth contrast as described above (Monteleone et al., 2006).

For the nomothetic assessments, 13 of the 14 subjects' contrast images were entered into the second stage of a random effects analysis as carried out in the original study (one-sample t-test, 2-tailed, $t=2.585$, $df=13$, $p<.01$). This process was performed for each of the fourteen subjects, resulting in 14 group response maps, which would be compared to the remaining individual response map. Each remaining individual subject map was submitted to the same threshold based on the coefficient of the least-squares estimate of the empirical data to the model ($p<.01$, $t=2.585$).

Significant regions were determined by applying an individual voxel probability threshold of $p<.01$ with a minimum cluster volume of 1072 microliters based on corner-to-corner connectivity in 3D space, which was equivalent to a connectivity radius of 3.46 mm. The cluster volume was chosen as the means to correct for multiple comparisons at a level of $alpha<.05$. Cluster volume threshold was determined with a Monte-Carlo simulation for which the input parameters modeled the analysis (voxel size 2x2x2mm, connectivity radius 3.46mm, FWHM gaussian smoothing at 6mm, individual voxel $p=.01$) executed within a mask of the entire brain (231766 voxels) for 1000 iterations using the AFNI program AlphaSim. The Monte Carlo simulation randomly generates "active" voxels within the mask according the probability and spatial parameters for the specified number of iterations, ultimately calculating the probability that a cluster of size X would occur by chance. The volume X is then used as a selection criterion on the experimental data to obtain activity clusters that meet the corrected Alpha level.

Masks were made from each resulting map of significant clusters. The individual mask was overlaid on the group predictor mask to identify points of coexisting significant activity in group and individual analyses.

Group responses were assessed in nine regions of interest: MPFC, DLPFC, VLPFC, ACC, Medial and Superior Frontal Cortex (Brodmann's areas 9 and 10), Temporal Gyrus, the Tempero-Parietal Junction of the superior temporal lobe, Cuneus/Precuneus, and sections of the anterior Basal Ganglia in the region including the caudate and putamen (see Figure 10). Thus, a reanalysis of the Phan et al. (2005) data using a conservative signal processing procedure replicated activation in the VLPFC, DLPFC, and DMPFC.



An analysis was applied using these 9 regions to determine, on a subject-by-subject basis, whether activation in each of these nine regions replicated the pattern of activation observed when data from the remaining 13 subjects were aggregated. For each region, a tally was kept of the number of false positive, false negatives, and hits across all 14 subjects. A hit was recorded at each cluster that showed overlap of significant group and individual responses indicating Lie > True. A false negative was recorded if no overlap was present, either due to absence of activation in the group or individual map. A false alarm was recorded if the group map predicted Lie > True and the within-subject response was significant for the opposite valence of True > Lie. A correct rejection was scored if the significant group prediction was True > Lie, and there was an overlapping individual response of the same valence.

Nine ROIs were significant (ps < .01). Individual ROI maps were overlaid on group ROIs to find points of regional overlap, indicating regions where significant activation co-existed in the group and individual analyses. Regions showing the best overlap between group and individual ROIs were MPFC, MSFG, DLPFC, and VLPFC. Classification results indicated that 57% of the Ss showed the predicted activation in at least 5 of the 9 ROIs, whereas 29% showed activation in 0 or 1 of these ROIs and would be considered false negatives. No false positives were observed, likely due to our use of

false discovery rate correction procedures during signal processing. Results were similar when classification was limited to the five most common ROIs, suggesting that individual classification of guilt or innocence using fMRI in the GKT may be subject to considerable error.

Resulting frequency distributions of successful classification were compared to chance using an analytical simulation of chance responses based on the observed data. Chance response frequencies of hits (H) and false alarms (FA) were modeled with the equation (H + FA)/2, based on the assumption that hits and false alarms would be equally distributed given random selection of the stimuli in the analysis. Simulated chance frequency distributions were compared to observed data, and of the 9 ROIs of interest, only MPFC and MSFG significantly differed from the simulated chance distribution (chi-square test, $X2=6.00$ and 6.67, respectively, $df=2$, $p<.05$).

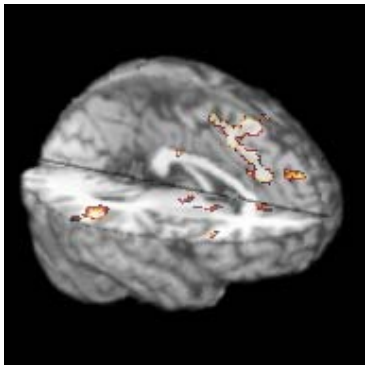
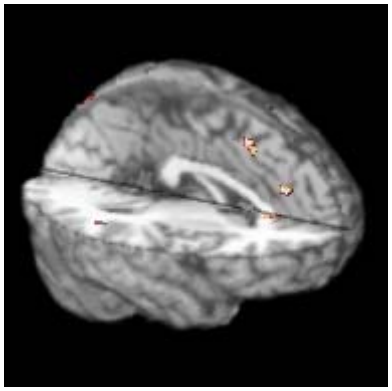
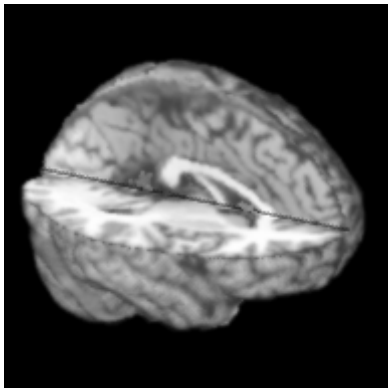


Figure 11 displays the best-case of classification of an individual subject. Depicted in Figure 11 is the overlap between the aggregate results obtained from aggregating the data from the other 13 subjects and the results obtained for the Lie > Truth contrast on this individual subject. As is apparent, all nine ROIs were observed in this individual subject.



Most cases were as impressive as this best-case finding. In Figure 12, we depict the results for the median subject in terms of overlap. The overlap was limited to 3 regions of interest, resulting in poor classification of deceptive responding based on the fMRI results for this individual subject.



For completeness, Figure 13 illustrates the results for the worst-case subject. As is apparent in this figure, there was no overlap in the activation pattern found for the Lie vs. Truth contrast.

We next limited this analysis to the three regions that were reported by Phan et al. (2005) and replicated in our re-analyses of these data, namely, the MPFC, DLPFC, and VLPFC. Results revealed deceptive responses were associated with activation of the VLPFC, DLPFC, MPFC in 36% of the subjects, deceptive responses were associated with activation of two of the regions in 14% of the subjects, deceptive responses were associated with activation of one of the regions in another 21% of the subjects, and deceptive responses were associated with no differences in activation of these regions in 29% of the subjects – again suggesting a high rate of false negatives despite the plurality of the subjects showing the same pattern of activation as found in the nomethetic analysis and, therefore, permitting accurate classification of deceptive responding.

In sum, fMRI analyses permitted the differentiation of deceptive and truthful responding at the aggregate level, but individual differences in patterns of brain activation were observed despite similarities in behavior on the task. These results suggest that, while fMRI may permit investigation of the neural correlates of lying, it does not appear to provide invariant markers of lying that generalize across individuals. This might be expected given the functions associated, for instance, with the three ROIs that were most robust. The MPFC has been associated with mentalizing and theory of mind (e.g., Frith & Frith, 2003; Saxe, 2004), processes that are involved in but are not unique to intentional deceptive responding. The DLPFC has been associated with working memory (Blumenfeld & Ranganath, 2006), again a process that may be involved to a greater degree whey responding deceptively than truthfully, at least when the lie has not been extensively rehearsed prior to testing as in the current study. Finally, the VLPFC has been associated with response inhibition and interference monitoring and suppression (Blasi et al., 2006) and with the presence of a target regardless of context (Rahm et al., 2006), processes again that may be more likely when responding deceptively than truthfully but processes that are not unique to lying. The close matching of deceptive and truthful conditions in Phan et al. (2005) and our use of false discovery rate corrections may have contributed to absence of false alarms. However, the fact that these regions are associated with cognitive operations that may emerge during truthful responding in stressful interrogations suggests that concerns about false alarms cannot yet be laid to rest in fMRI studies.

Two other issues warrant commentary. The lies in Phan et al. (2005) were not extensively rehearsed. This feature of Phan et al. (2005) should increase the likelihood that subjects would show greater activation in these ROIs. Thus, greater attention needs to be given to different kinds of deceptive responding, such as spontaneous lies versus rehearsed lies. Second, subjects in the study were not implementing countermeasures to mask their deceptive responding. The present results suggest that effective cognitive countermeasures should be possible to develop. For instance, if unbeknownst to the examiner the subjects were to intently think about the mental state of the examiner and to concentrate on inhibiting competing thoughts and ideas when making truthful responses, the activation of the MPFC, DLPFC, and VLPFC should be boosted, thereby making it more difficult to detect differences in the deceptive and truthful conditions. Such a hypothesis requires testing, however.

The most important finding in this study, however, is that even under among the best of conditions the fMRI activation observed on an individual-by-individual basis yielded an unacceptable rate of false-negatives. Lowering the threshold for classifying a region as activated did not improve classification much, suggesting the false-negatives had more to do with differences in the information processing operations underlying deceptive and nondeceptive responding rather than in conservative decision rules for identifying activated areas per se.

References

Allen, J.J.B. & Iacono, W.G. A comparison of methods for the analysis of event-related potentials in deception detection. *Psychophysiology*. **34**:234-40, (1997).

Ben-Shakar, G. & Fruedy, J.J. *Theories and application in detection of deception*. Springer-Verlag: New York. (1990).

Blasi G. Goldberg TE. Weickert T. Das S. Kohn P. Zoltick B. Bertolino A. Callicott JH. Weinberger DR. Mattay VS. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *European Journal of Neuroscience. 23, 1658-1664.*

Blumenfeld RS. Ranganath C. (2006). Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. *Journal of Neuroscience. 26(3), 916-925.*

Cacioppo, J.T. & Tassinary, L.G. Infering psychological significance from physiological signals. *American Psychologist*, **45**:16-28, (1990).

Farwell, L.A., Donchin, E. The truth will out: interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*. **28**:531-47, (1991).

Friston KJ et al. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp*, 189-210.

Frith U. Frith CD. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences. 358(1431), 459-473.*

Gehring, W.J. & Willoughby, A.R. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*. **295**:2279-82, (2002).

Holmes AP, & Friston KJ (1998). Generalisability, randome effects and population inference. *Neuroimage, 7*, S754.

Iacono, W. The detection of deception. In Cacioppo, J.T., Tassinary, L.G., & Berntson, G.G. (Eds). *The handbook of psychophysiology, 2nd*. pp772-793. Cambridge University Press: New York. (2000).

Langleban, DD, Schroeder L, Maldijian, JA, et al. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *Neuroimage, 15*, 727-732.

Langleben DD, Loughead JW, Bliker WB, Ruparel K, Childress AR, Busch SI, Gur RC. (2005). Telling Truth From Lie in Individual Subjects With Fast Event-Related fMRI. Human Brain Mapping 26, 262-272.

Lavond, D.G., & Steinmetz, J.E. *Handbook of Classical Conditioning*. Klewer Academic Publishers NY. 2002.

Lober, K., Lachnit, H. *Configural Learning in Human Pavlovian Conditioning: Acquisition of a Biconditional Discrimination.* Biological Psychology 59 (2002) 163-168.

Lykken, D.T. *A tremor in the blood: Uses and abuses of the lie detector. 2nd* Plenum: New York. (1998).

Milham, M.P., Banich, M.T., Webb, A., Barad, V., Cohen, N.J., Wszalek, T., Kramer, A.F. The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Res Cogn Brain Res*.**12**:467-73, (2001).

Phan KL, Fitzgerald DA, Gao K, et al. (2004). Real-time fMRI of cortico-limbic brain activity during emotional processing. *Neuroreport, 15*, 527-532.

Phan KL, Magalhaes A, Ziemlewicz TJ, Fitzgerald DA, Green C, Smith W. (2005). Neural Correlates of Telling Lies: A Functional Magnetic Resonance Imaging Study at 4 Tesla. Acad Radiol, 12, 164-172.

Posse S et al. (1999). Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magn Reson Med, 42*, 87-97.

Rahm B. Opwis K. Kaller CP. Spreer J. Schwarzwald R. Seifritz E. Halsband U. Unterrainer JM. (2006). Tracking the subprocesses of decision-based action in the human frontal lobes. *Neuroimage. 30, 656-667.*

Rosenfeld, J.P., Angell, A., Johnson, M., Qian, J.H. An ERP-based, control-question lie detector analog: algorithms for discriminating effects within individuals' average waveforms.*Psychophysiology*. **28**:319-35, (1991).

Speck O, et al. (1998). Functional imaging by 10- and T2* parameter mapping using multi-image EPI. *Magn Reson Med, 40*, 243-248.

Spence, S.A., Farrow, T.F., Herford, A.E., Wilkinson, I.D., Zheng, Y., Woodruff, P,W. Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*. **12**:2849-53, (2000).

Swets, J.A. Measuring the accuracy of diagnostic systems. *Science*. **240**:1285-93, (1988).

Tucker, D. C. (2005). Applying classical conditioning toward the physiological detection of concealed information: Beyond native responses. Trial Research Project, University of Chicago, 8 June 2005.