*Annals of Biomedical Engineering*, Vol. 34, No. 7, July 2006 (© *2006*) pp. 1182–1189
DOI: 10.1007/s10439-006-9143-3

# Facial Skin Surface Temperature Changes During a "Concealed Information" Test

DEAN A. POLLINA,[1] ANDREW B. DOLLINS,[1] STUART M. SENTER,[1] TROY E. BROWN,[1] IOANNIS PAVLIDIS,[2] JAMES A. LEVINE,[3] and ANDREW H. RYAN[1]

[1]Department of Defense Polygraph Institute, 7540 Pickens Ave., Fort Jackson, SC 29207; [2]Department of Computer Science, University of Houston, 501 Philip G. Hoffman Hall, Houston, TX 77204; and [3]Mayo Clinic, Endocrine Research Unit, Department of Medicine, Rochester, MN 55905

*(Received 14 March 2003; accepted 23 May 2006; published online: 20 June 2006)*

**Abstract—**When individuals who commit a crime are questioned, they often show involuntary physiological responses to remembered details of that crime. This phenomenon is the basis for the concealed information test, in which rarely occurring crime-related details are embedded in a series of more frequently occurring crime-irrelevant items while respiratory, cardiovascular, and electrodermal responses are recorded. Two experiments were completed to investigate the feasibility of using facial skin surface temperature (SST) measures recorded using high definition thermographic images as the physiological measure during a concealed information test. Participants were randomly assigned to nondeceptive or deceptive groups. Deceptive participants completed a mock-crime paradigm. A focal plane array thermal imaging radiometer was used to monitor SST while crime-relevant and crime-irrelevant items were verbally presented to each participant. During both experiments, there were significant facial SST differences between deceptive and nondeceptive participants early in the analysis interval. In the second experiment, hemifacial (i.e., "half-face" divided along the longitudinal axis) effects were combined with the bilateral responses to correctly classify 91.7% of participants. These results suggest that thermal image analysis can be effective in discriminating deceptive and nondeceptive individuals during a concealed information test.

**Keywords—**Imaging/Infrared thermography, Behavior/Physiologic behavior, Polygraph, Face temperature.

## INTRODUCTION

The use of polygraph instruments in criminal investigations has a long history, and several types of tests using the polygraph have been developed. In one such test, called the "concealed information test" (CIT), criminal suspects are presented with crime-relevant items and with items unrelated to the crime while measures of cardiovascular, respiratory, or electrodermal activity are recorded.[19,20] In a properly designed CIT, the crime-relevant items occur randomly, less frequently than the unrelated items, and are similar in all respects to the crime-irrelevant items except that they are related in some way to the crime under investigation.[8,10,17,18,22] Additionally, it is important that the crime-relevant details used in the CIT be known only by the perpetrator.[11] Typically, the CIT items are created by the examiner after an investigation of the crime scene. Crime-relevant items are selected from a category that represents a salient feature of the crime itself (e.g., the color of the getaway car), and then the crime-irrelevant items are created by using similar items from that category (e.g., other colors). If a suspect's physiological responses are consistently larger for the crime-relevant items, it is assumed that the suspect has recognized the crime-related information. The mechanisms underlying the differential responsiveness most likely involve the "orienting response" first reported by Pavlov, and detailed in Sokolov's orienting response model.[23,24] In Sokolov's conceptualization, the brain develops a "neuronal model" (or memory) of the stimulus environment, which must be updated when a stimulus of some significance is perceived. Autonomic nervous system (ANS) correlates of neuronal model update include respiratory, cardiovascular, and electrodermal (e.g., sweat gland activity) changes.

The CIT is extremely effective when traditional polygraph measures are used. However, the devices used to record these measures have changed very little in more than 20 years,[16] and most often include metal electrodes attached to the fingers, pneumatic tubes surrounding the thoracic and abdominal areas, and a pneumatic blood pressure cuff attached to the upper arm overlying the brachial artery. These sensors require time to attach and the blood pressure cuff may be uncomfortable to the examinee when inflated for more than approximately five minutes.[26] Additionally, ANS measures of the orienting response rely on cognitive phenomena such as memory updating rather than emotional responses to the test questions.[14,15] Many believe that increases in polygraph accuracy might be possible if the specific emotion experienced in response to particular

Address correspondence to Dean A. Pollina, Research Division, Department of Defense Polygraph Institute, 7540 Pickens Ave., Fort Jackson, SC 29207. Electronic mail: pollinad@jackson-dpi.army.mil

questions could be determined.[2] This would be useful not only in the CIT format, but in other polygraph test formats as well. Research has documented a link between behavior and the expression of specific emotions.[3–7] The cited studies typically involve detailed measures of facial muscles as specific emotions are invoked.

One technology that shows promise in overcoming some of the limitations of traditional polygraph measures is thermography. Thermography is a technique used for measuring heat (infrared) emission from the human body.[12,13] Using infrared (IR) radiometry, heat measurements from large areas of the body surface can be made without skin contact. Skin surface temperature (SST) is affected by changes in underlying muscle activity and microcirculation,[1,9] suggesting that it might bridge the gaps among behavioral studies of facial expression, emotion, and the ANS measures traditionally used to score polygraph tests. In the studies described below, facial SST was recorded from deceptive and nondeceptive individuals while a CIT was administered. It has been shown that the periorbital region around the eye is associated with increased SST during arousal and it might be associated with specific emotions.[21] Therefore, data analysis focused on this region. The hypothesis that skin temperature in the eye region could be used to discriminate deceptive and nondeceptive examinees in a manner similar to traditional polygraph measures was tested.

## EXPERIMENT 1

### *Methods*

*Participants*. Thirty participants (6 Female) between the ages of 19 and 28 (Mean = 21.2) were recruited from a sample of U.S. Army basic trainees stationed at Fort Jackson, South Carolina and assigned to duty at the Department of Defense Polygraph Institute (DoDPI). The percentage of female and male participants was based on the population of basic trainees at Fort Jackson selected by military personnel for assignment at the Polygraph Institute. All participants were given the option of participating in this research study, watching television, or reading in the DoDPI library for the day. Informed consent was obtained and documented for all participants. All participants were in good health by self-report, and none were taking medications except for analgesics (e.g., ibuprofen) for minor injuries sustained during basic training. Five participants were dropped from the study due to: incriminating statements made to the polygraph examiner ($n = 1$), sleeping during the polygraph examination ($n = 2$), or failure or unwillingness to commit a mock crime ($n = 2$). Data from ten other participants could not be analyzed because the storage media had become corrupted. This resulted in the inclusion of 15 participants (5 deceptive, 10 nondeceptive; age mean = 21.1; 3 Female) in the final data analyses.

*Apparatus*. A Raytheon model $256 \times 256$ focal plane array (FPA) thermal imaging radiometer (Sensitivity $< 0.10°C$) was used to monitor SST. The radiometer 12-bit digital output was connected to a high-speed digital video processing board supplied with software designed specifically for thermal imaging installed in a Pentium III 466 Mhz computer. The mock crime was committed in a room that included a plastic mannequin, purse, screwdriver, and chair. Physiological data were collected in a darkened, temperature controlled room (Range: 20–22°C).

### *General Procedure*

All interested participants were instructed to read a brief description of this research project and sign an informed consent form. Each participant then answered a series of biographical and medical questions to ensure that they were in good health and not taking medication that could interfere with the examination results. After all forms were completed, the investigator explained how the polygraph examination would be conducted. Each participant was randomly assigned to either the deceptive or nondeceptive group. Participants in the nondeceptive group were told that they would be taking a polygraph test as part of a research study and questioned about the "murder" of a woman that took place at the DoDPI earlier that day. Since they did not commit this crime, they were instructed to answer all questions truthfully during the polygraph examination. Participants assigned to the deceptive group were told that they would be involved in a pretend crime, and would lie about this during the polygraph examination in an attempt to appear innocent. Participants then either waited quietly to be brought to the polygraph examination room (nondeceptive group) or committed a pretend crime (deceptive group).

### *Procedure for Deceptive Group*

Prior to each participant's arrival at the DoDPI, a mock crime room was prepared. In the room, a plastic dummy was seated in a chair. A purse containing $20 U.S. was placed next to the dummy, and a screwdriver was placed on a table next to the purse. Participants in the deceptive group were instructed by the investigator to enter the mock-crime room without being seen, stab the dummy with the screwdriver, and steal the $20 from the purse. After committing the mock crime, each participant was asked details about the crime. Questions included, "Were you seen by anyone? Did you remember to steal the $20? What happened to the woman in the room?" Participants who failed to stab the dummy and steal the $20 were excluded from this study.

### *Data Collection Procedures*

At a prearranged time, each participant was met by a U.S. Government certified polygraph examiner, who was blind to the participant's group membership. Polygraph sensors were attached to the participant, and each participant was administered a Zone Comparison (ZCT) polygraph test. The results of the ZCT will be reported elsewhere. Next, the CIT was administered. Six questions were presented during a single series, and each series was repeated three times during the polygraph examination. Each question series consisted of one "relevant" item (the murder weapon) and five irrelevant items (weapons that were not used to commit the crime). For each participant, the relevant item ("Screwdriver") was the fifth item of the first series, the second item in the second series, and the sixth item in the third series. Each item was presented verbally by the polygraph examiner approximately 25 s after the onset of the previous question. Irrelevant items used in the CIT were: pencil, bayonet, scissors, letter opener, and bowie knife. All examinees were instructed to repeat the item spoken by the polygraph examiner immediately after hearing each item.

Recordings of facial temperature values were started at the onset of each of the relevant and irrelevant items using a 30 Hz sampling rate for five seconds (150 image frames) and a $256 \times 256$ FPA. Thermal image data collection was started with the press of a computer key after a prearranged signal (finger tap) from the polygraph examiner prior to reading each item in the sequence. The resulting ($256 \times 256 \times 150$) array of temperature values collected during the presentation of each item was converted to an ASCII text file and stored on a CD-R disk for off-line data analysis. Due to storage limitations, only data from the first series of crime-relevant and irrelevant items were collected.

*Data Reduction*. Given the theoretical framework worked out by Pavlidis *et al.*[21] increases in blood flow around the eyes should cause relatively 'hot' blood from nearer to the core of the body to flow toward the periphery, resulting in greater temperature changes in this region. To measure these changes, the thermal image frame containing the onset of the examinee's verbal response was determined by visual inspection. Next, the maximum and minimum (between-frames) temperature for each specific pixel in images recorded during each of three (0.33 s) time intervals preceding and three (0.33 s) time intervals following the examinee's verbal response was determined, using peak-hold and valley-hold algorithms. This resulted in six "maximum temperature" and six "minimum temperature" image frames for each participant. The average pixel intensity of two bilaterally symmetric regions of interest (ROIs) were then calculated for each peak-hold and valley-hold frame ($10 \times 10$ pixels each; located directly under the left and right lower eyelids) using a commercially available image analysis software package (National Instruments Labview IMAQ Version 6.0, Austin, TX). The length and width of the ROIs were adjusted slightly ($\pm 0.02\%$ of image area) for two participants according to the size of their faces within the images.

Separate $2 \times (2 \times 6 \times 2)$ mixed factor "maximum amplitude" and "minimum amplitude" ANOVAs were calculated using the mean ROI temperature in the left and right hemiface (i.e., "half-face" divided along the longitudinal axis) of the peak-hold (maximum amplitude) and valley-hold (minimum amplitude) images as the dependent variables. A third ANOVA was calculated on the difference between the mean ROI temperature of each maximum temperature and minimum temperature image (temperature range) in the left and right hemiface. Independent variables consisted of group (deceptive/nondeceptive), hemiface, time interval, and question type (crime-relevant/crime-irrelevant).

### *Results*

Figure 1 (left) shows a topographic map of the facial SST distribution of a single deceptive examinee at the onset of his verbal response. The region used in data analysis is shown as two rectangular regions; one just below each eye. Temperature gradients in the regions around the eyes and forehead can also be seen (Fig. 1, right). The ANOVA conducted on maximum amplitudes showed a significant four-way interaction $F(5, 45) = 3.27$, $p < .05$, $\eta^2 = .26$. This interaction was further examined by calculating two $2 \times (6 \times 2)$ factorial ANOVAs including only the right or the left hemiface in each analysis. The Group $\times$ (Interval $\times$ Question Type) interaction was statistically significant $F(5, 45) = 3.85$, $p < .05$, $\eta^2 = .30$ in the analysis conducted on right hemiface data, but there were no significant results in the analysis conducted on left hemiface data. The ANOVAs conducted on minimum and range of SST amplitudes did not indicate any significant main effects or interactions.

The Group $\times$ Hemiface $\times$ Time Interval $\times$ Question Type interaction for maximum response amplitudes is illustrated in Fig. 2. For nondeceptive participants, the between-interval pattern of responses appears similar in both the left and right hemiface. However, for deceptive participants, the response pattern appears more variable, particularly the right hemiface data. During the earliest analysis interval, beginning one second prior to response onset, the differences in maximum amplitude following irrelevant and relevant items appear to be larger in the positive direction in the deceptive group, particularly in the right hemiface. During the last analysis interval, beginning 0.66 s after response onset, this difference appears to be larger in the negative direction in the deceptive group; again the effect was largely in the right hemiface. However, further analyses of simple effects were not calculated because results with this small sample could be misleading.

### *Discussion*

The results of Experiment 1 suggest that thermal image analysis may be used to effectively discriminate deceptive from nondeceptive individuals during a CIT. Specifically, deceptive participants' maximum amplitude crime relevant/crime irrelevant question response differences were larger than those of nondeceptive participants early in the

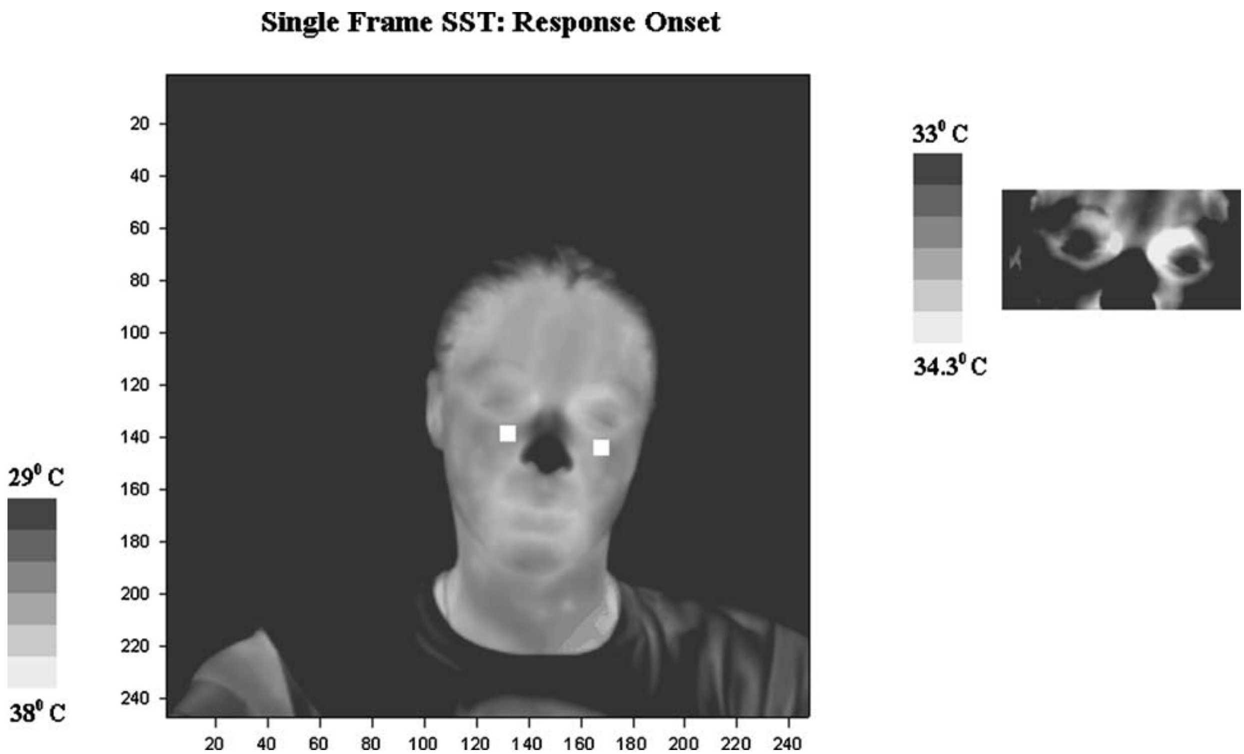


**FIGURE 1. (Left) Facial SST distribution of a single deceptive examinee at the onset of his verbal response. The regions used in data analysis are shown as two rectangular ROIs just below the left and right eyes. (Right) Facial SST distribution in the region around the eyes at response onset.**

response interval, but more so in the right hemiface. However, during the latter part of the response interval, deceptive participants' maximum amplitude SST response differences were larger than those of nondeceptive participants in the opposite direction. Again, this effect appeared to be more pronounced in the right hemiface.

The finding that the deceptive and nondeceptive group SST differences depend on the recording time interval and the type of stimulus presented suggests that these responses are related to stimulus information processing and to preparation for a response. Human vision is highly developed resulting in the ability to visually locate threats in the environment. From an evolutionary perspective, effective self-defense must be extremely rapid. The cost of failing to detect a threat could result in death. It is therefore not surprising that temperature changes in the area around the eyes, which most likely reflect increased blood flow, should occur extremely rapidly.[21] These blood flow increases are required to meet the metabolic requirements of the visual system in the moments after the detection of a threatening stimulus.

The results of Experiment 1 suggest that thermal image analysis may be useful in discriminating between deceptive and nondeceptive participants during a CIT. There were, however, sufficient deficiencies in Experiment 1 (e.g., small number of participants, no quantitative system of head-movement tracking) that examination of additional data is warranted before drawing any firm conclusions.

## EXPERIMENT 2

The participant treatment procedures of Experiment 1 were replicated in Experiment 2. Data were recorded with a different thermal camera. A more sophisticated method of tracking head movements in real time was used during data reduction. In Experiment 2, statistical analyses again focused on a two dimensional array of temperatures recorded from the periorbital regions under participants' left and right eyes. This allowed us to continue to explore the possibility that these areas within the periorbital region might be useful as thermal signatures in the detection of deception, and to reinvestigate the possibility that hemifacial differences might have contributed to the significant results seen in Experiment 1. Because the results from Experiment 1 do not make it clear how thermal measurements can be used in the individual case (in a manner similar to traditional polygraph), binary logistic regression analysis was used in Experiment 2 to classify individual study participants and determine the practical value of these thermal measurements for use in the psychophysiological detection of deception.

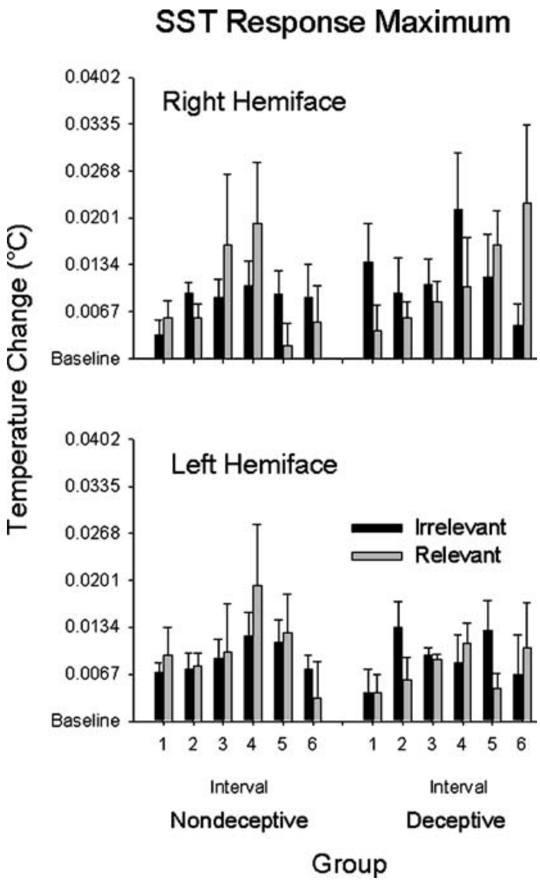


**FIGURE 2. Mean (+S.E.M.) Group × Hemiface × Time Interval × Question Type SST responses. Six time intervals are shown. Three (0.33 s) time intervals preceded and three (0.33 s) time intervals followed the examinee's verbal response (Response Baseline).**

## Methods

*Participants*. Thirty-nine participants (21 Female) between the ages of 19 and 35 (Mean = 22.3) were recruited from a sample of U.S. Army basic trainees stationed at Fort Jackson, South Carolina and assigned to duty at the DoDPI. None of them had participated in Experiment 1. All participants were given the option of participating in this research study, watching television, or reading in the DoDPI library for the day. Informed consent was obtained and documented for all participants. All participants were in good health by self-report, and none were taking medications except for analgesics ($n = 1$) and nasal decongestants ($n = 1$). One participant was dropped from the study due to incriminating statements made to the polygraph examiner. Data from twelve other participants were discarded due to excessive head movements ($n = 10$) or to the inability to measure head movement reliably ($n = 2$). Data from 2 additional participants were dropped from the statistical analysis because they were outliers (greater than 2 s.d. above their group's mean). This resulted in the inclusion of 24 participants (12 deceptive, 12 nondeceptive; age mean = 22.8; 14 Female) in the final data analyses.

*Apparatus*. A FLIR Systems (model SC1000) 256 × 256 focal plane array thermal imaging radiometer (Sensitivity = 0.10°C) placed directly in front of each participant at a distance of 1.5 m and focused directly at the face was used to monitor SST. The radiometer 8-bit digital output was connected to a high-speed digital video processing board (National Instruments Model NI-PXI 1411) supplied with software designed specifically for image processing (National Instruments IMAQ) installed in a Pentium 1.8 GHz computer.

*General Procedure and Procedure for Deceptive Group*. Participant instructions and procedures were the same as in Experiment 1. In this Experiment, each mock-crime was recorded to verify that it was successfully committed. Once again, Participants in the deceptive group who failed to stab the dummy and steal the $20 were excluded from the study.

*Data Collection Procedures*. At a prearranged time, each participant was met by a U.S. Government trained polygraph examiner, who was blind to the participant's group membership. Polygraph sensors were attached to the participant. An additional sensor that recorded near-infrared responses from the area under the forehead was also attached at this time using a neoprene strap wrapped around the forehead and tied at the back of the head (pilot data were collected using this sensor, but the results are not reported here). Next, each participant was administered a ZCT polygraph test. The results of the ZCT will be reported elsewhere. Next, the CIT was administered. Once again, six questions were presented during a single series. However, in this experiment each series was repeated four times during the polygraph examination. As in Experiment 1, each question series consisted of one "relevant" item (the murder weapon) and five irrelevant items (weapons that were not used to commit the crime). For each participant, the relevant item ("Screwdriver") was the fourth item of the first series, the third item in the second series, the second item in the third series, and the fifth item in the fourth series. Each item was presented to participants via computer using auditory stimuli (.wav files) previously recorded by one of the experimenters (T.B.). The interval between the presentations of each item was held constant at 15 s. Irrelevant items used in the CIT were the same as those used in Experiment 1: pencil, bayonet, scissors, letter opener, and bowie knife. All examinees were instructed to repeat the item spoken by the polygraph examiner immediately after hearing each item.

*Data Reduction*. Facial skin surface temperature values were recorded continuously throughout each stimulus series using a 30 Hz sampling rate, interpolated to 15 Hz by averaging across each two successive frames. Each of the averaged frames was then converted to 240 × 240 pixel JPEG format in real-time and stored on computer disk for off-line analysis. The JPEG images were saved as 8-bit gray scale; thus pixel intensity was linearly related to the temperature recorded by the thermal camera. The onset of each stimulus was accompanied by the opening of a shutter that exposed a heat-emitting diode. The diode heat was recorded by the thermal camera and was visible in the corresponding JPEG file as a high-intensity spot in the upper-left corner of the image. The high-intensity spot was used to synchronize the thermal data with the onset of each stimulus.

For each participant, the average pixel intensity of two bilaterally symmetric regions of interest ($10 \times 10$ pixels each; located directly under the left and right lower eyelids) was calculated for each frame using a procedure similar to Experiment 1. A computer program developed for this study using commercially available software (National Instruments Labview Version 6.0, Austin, TX) was used to extract these measures from 70 sequential images (i.e., 4.66 s) beginning at the onset of each crime-relevant and crime-irrelevant item. Due to the findings of Experiment 1, which showed greater discrimination occurring rapidly after stimulus onset, SST data collected more than 4.66 s after the onset of each item were not analyzed in this study. This data analysis strategy was similar to that used in Experiment 1. Head movement was also tracked continuously using a commercially available pattern-matching algorithm (National Instruments IMAQ), and the left and right regions of interest were adjusted for each image according to the recorded movements of the head.

The resulting array of temperature values from the left and right regions of interest for each of the four repetitions of crime-relevant and crime-irrelevant items was categorized by repetition number and stimulus type. Each SST waveform was then baseline corrected such that the temperature value corresponding to stimulus onset was subtracted from each of the time points on the waveform. Responses occurring within the analysis windows that contained vertical or horizontal head movements greater than 1/24 of the radiometer lens field of view (10 pixels in the JPEG images) were deleted prior to statistical analysis.

For each participant, mean responses at each of the 70 time points were obtained by collapsing across the four question series and the five crime-irrelevant items. For each of 14 successive (0.33 s) time intervals along each crime-relevant and crime-irrelevant waveform derived from the left and right ROIs, minimum and maximum SSTs were recorded. SST range was also calculated as the difference, in Celsius degrees, between the minimum and maximum SSTs in the interval. Difference scores were then obtained for each participant, measure, and interval by subtracting each maximum, minimum, and range of the "crime-relevant" measure from the corresponding "crime-irrelevant" measure. This resulted in six measures for each participant at each 0.33 s interval: left hemiface maximum amplitude (LH Max), minimum amplitude (LH Min), and range (LH Range), and right hemiface maximum amplitude (RH Max), minimum amplitude (RH Min), and range (RH Range). SST

**TABLE 1. Experiment 2: classification table: binary logistic regression.**

| Observed | Predicted | | Percentage correct |
|---|---|---|---|
| | Nondeceptive | Deceptive | |
| Nondeceptive | 11 | 1 | 91.7 |
| Deceptive | 1 | 11 | 91.7 |

*Note*. Classifications were made using a case inclusion cutoff probability of .5, which forced a classification of each case. All six subtraction measures (LH Max, LH Min, LH Range, RH Max, RH Min, and RH Range) were entered on step 1 (simultaneous analysis). Of these, only LH Max (*Wald* = 2.42), LH Min (*Wald* = 3.30), RH Max (*Wald* = 4.38) and RH Min (*Wald* = 1.76) were included in the regression equation. In this analysis, only the coefficients for RH Max appeared to be significantly different from 0 as a stand-alone measure, using a significance level of .05.

responses less than 29°C or greater than 38°C were considered outliers and were rejected prior to any statistical analysis.

A logistic regression procedure was used to determine the extent to which the periorbital skin surface temperature could be used to accurately classify each participant as either deceptive or nondeceptive. The analysis used Deception/Nondeception as a dichotomous dependent variable, and LH Max, LH Min, LH Range, RH Max, RH Min, and RH Range as each of six covariates. A separate logistic regression analysis was run at each successive 0.33 s interval from 0 to 4.66 s, and the criterion for statistical significance was set at $p < .01$.

## Results

*Binary Logistic Regression*. Classification accuracy was significantly above chance for deceptive and nondeceptive groups in the first 0.33 s interval (Table 1). Mean LH Max, RH Max, and RH Min values were larger in the deceptive group, and this pattern was probably responsible for the significant results seen. Within the first interval, the four SST maximum and minimum measures accounted for a significant proportion ($R^2 = .62$) of the variation in the regression model, $X^2$ (4, $N = 24$) = 23.0, $p < .001$. For a given participant, classifications were made using the logistic regression equation,[25]

$$\text{Prob(Deception)} = \frac{1}{1 + e^{-Z}}$$

where, Z = −7.512 + (139.595)(LH Min) + (56.589)(LH Max) + (−42.248)(RH Min) + (208.393)(RH Max).

*Group Analyses*. To facilitate a direct comparison with the results of Experiment 1, deceptive and nondeceptive group means were compared within the first 0.33 s interval. A separate (Deceptive/Nondeceptive Group × Left/Right Hemiface × Crime Relevant/Crime Irrelevant Question

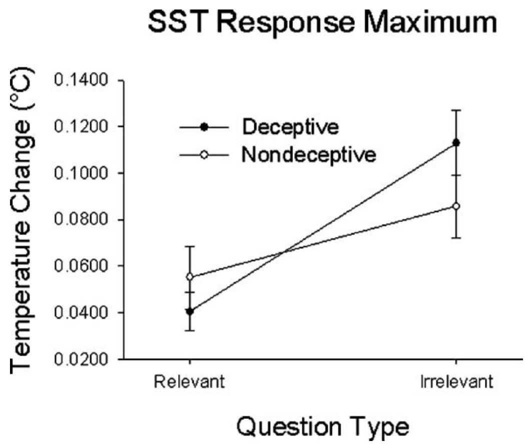


**FIGURE 3. Mean (+S.E.M.) Group × Question Type SST responses during the first 0.33 s time interval after stimulus onset. The two question types were crime-relevant and average of crime-irrelevant items.**

Type) analysis was calculated using maximum, minimum, and range scores. The results of the Maximum Amplitude ANOVA showed a significant main effect of Question Type, $F(1,23) = 58.21$, $p < .01$, $\eta^2 = .72$, and a significant Group × Question Type interaction $F(1,23) = 9.56$, $p < .01$, $\eta^2 = .29$. The significant interaction was due to greater SST differences between irrelevant items and the relevant item in the deceptive group, versus the nondeceptive group, during this time interval (Fig. 3). The results of the Minimum Amplitude, $F(1,23) = 25.20$, $p < .01$, $\eta^2 = .52$, and Range, $F(1,23) = 47.32$, $p < .01$, $\eta^2 = .68$, ANOVAs both revealed only a significant main effect of Question Type.

## GENERAL DISCUSSION

The results of our two experiments suggest that areas within the periorbital regions are useful as thermal signatures in the detection of deception, and that hemifacial differences may contribute to the effective classification of individual cases. The statistically significant group differences seen in Experiment 1 were replicated in Experiment 2, which also incorporated a more sophisticated method of tracking head movements and a greater number of question series repetitions. In both studies, the SST responses in discrete facial regions occurred very rapidly (within about 1 s) after stimulus presentation.

The finding that the right and left sides of the face can produce different SST responses is interesting. The results of the experiments reported here suggest that these SST changes are occurring rapidly, and in a region that is in close proximity to the anterior ethmoidal blood vessels. The ethmoidal artery is one branch of the ophthalmic artery, which supplies blood to most structures in the orbit as a branch of the internal carotid artery. The carotid artery also supplies structures in the brain. These hemifacial effects could therefore be explained as the differential activation of the right hemisphere as part of a right hemisphere-mediated emotional response,[27] or the result of (asymmetric) sympathetically-mediated neural activity whose function is to deliver a rapid supply of oxygenated blood to the eyes.[21] The hypothesis that these findings are related to autonomic activity in the peripheral nervous system is supported by the rapid time course and distribution of the SST responses.

Despite these promising results, this study has some limitations that should be addressed in future studies. The results of Experiment 1 show statistically significant differences between deceptive and nondeceptive groups, and the results of Experiment 2 extend these findings to the individual case. However, because of possible training effects, logistic regression without cross-validation should be viewed with caution until follow-up studies can be completed using similar tests and measures. The extent to which thermography will increase accuracy beyond that which is possible using traditional polygraph measures is not yet known. Future researchers will likely focus on the increases in sensitivity and specificity that result from combining thermography and traditional polygraph measures, relative to those using either approach alone. It is also likely that the dependent measures used in this study, minimum and maximum response amplitudes in the region immediately around the eyes, will be sub-optimal when used in isolation. Thermal imaging researchers using data transformations that more effectively isolate and discriminate the SST response could lead to further accuracy increases in the thermal detection of deception.

## ACKNOWLEDGMENTS

The authors would like to thank Kay Williams, Betty Rodriguez, and Rose Swinford of the DoDPI Research Staff for their assistance with data collection procedures. We would also like to thank Gordon Barland, Esther Harwell, Ron Kiefer, and Don Krapohl, all of whom administered the polygraph exams to our study participants. We are also grateful to Johnnie Rodgerson for his advice concerning instructions given to deceptive participants. This project was funded by the Department of Defense Polygraph Institute as project numbers DoDPI00-P-0011 and DoDPI02-P-0012. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of Defense or the U.S. Government

## REFERENCES

[1] Drummond, P., and J. Lance. Facial flushing and sweating mediated by the sympathetic nervous system. *Brain* 110:793–803, 1987.

[2] Ekman, P. Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage. New York: W. W. Norton & Company, 1992.

[3]Ekman, P., W. Friesen, and S. Ancoli. Facial signs of emotional experience. *J. Pers. Soc. Psychol.* 39:1125–1134, 1980.
[4]Ekman, P., W. Friesen, and P. Ellsworth. Emotion in the Human Face. Elmsford NY: Pergamon Press, 1972.
[5]Ekman, P., J. Hager, and W. Friesen. The symmetry of emotional and deliberate facial actions. *Psychophysiology* 18:101–106, 1981.
[6]Ekman, P., R. Levenson, and W. Friesen. Autonomic nervous system activity distinguishes among emotions. *Science* 22:1208–1210, 1983.
[7]Ekman, P., and M. O'Sullivan. Who can catch a liar. *Am. Psychol.* 46:913–920, 1991.
[8]Elaad, E., and G. Ben-Shakhar. Effects of motivation and verbal response type on psychophysiological detection of information. *Psychophysiology* 26:442–451, 1989.
[9]Fox, R., R. Goldsmith, and D. Kidd. Cutaneous vasomotor control in the human head, neck, and upper chest. *J Physiol.* 161:298–312, 1962.
[10]Furedy, J. J., and G. Ben-Shakhar. The roles of deception, intention to deceive, and motivation to avoid detection in the psychophysiological detection of guilty knowledge. *Psychophysiology* 28:163–171, 1991.
[11]Furedy, J. J., C. Davis, and M. Gurevich. Differentiation of deception as a psychological process: A psychophysiological approach. *Psychophysiology* 25:683–688, 1988.
[12]Gorbach, A. M. Infrared imaging of brain function. In: Optical Imaging of Brain Function and Metabolism, edited by U. Dirnagl, A. Villringer, and K. M. Einhaupl. New York: Plenum Press, 1993, pp. 95–123.
[13]Grayson, J. Responses of the microcirculation to hot and cold environments. In: Thermoregulation: Physiology and Biochemistry, edited by W. C. Bowman, E. Schönbaum, and P. Lomax. New York: Pergamon Press, 1990, pp. 221–234.
[14]Horneman, C. J., and J. G. O'Gorman. Detectability in the card test as a function of the subject's verbal response. *Psychophysiology* 22:330–333, 1985.
[15]Iacono, W. G., G. Boisvenu, and J. A. Fleming. Effects of diazepam and methylphenidate on the electrodermal detection of guilty knowledge. *J. Appl. Psychol.* 69:289–299, 1984.
[16]Kleinmuntz, B., and J. J. Szucko. Lie detection in ancient and modern times. *Am. Psychol.* 39:766–776, 1984.
[17]Lieblich, I. Manipulation of contrast between differential GSR responses through the use of ordered tasks of information detection. *Psychophysiology* 6:70–77, 1969.
[18]Lieblich, I., S. Kugelmass, and G. Ben-Shakhar. Efficiency of GSR detection of information as a function of stimulus set size. *Psychophysiology* 6:601–608, 1970.
[19]Lykken, D. T. The GSR in the detection of guilt. *J. Appl. Psychol.* 43:385–388, 1959.
[20]Lykken, D. T. The validity of the guilty knowledge technique: The effects of faking. *J. Appl. Psychol.* 44:258–262, 1960.
[21]Pavlidis, I., N. L. Eberhardt, and J. Levine. Human behavior: Seeing through the face of deception. *Nature* 415:35, 2002.
[22]Seymour, T. L., M. Shafto, and A. L. Mosmann. Using response time measures to assess "Guilty Knowledge." *J. Appl. Psychol.* 85:30–37, 2000.
[23]Sokolov, E. Perception and the Conditioned Reflex. New York: Macmillan, 1963.
[24]Sokolov, E., and Cacioppo, J. "Orienting and defense reflexes: vector coding the cardiac response". In: Attention and Orienting: Sensory and Motivational Processes, edited by P. Lang, R. Simons, and M. Balaban. New Jersey: Lawrence Erlbaum Associates, 1997, pp 1–22.
[25]SPSS Inc. SPSS Regression Models 10.0. Chicago: SPSS Inc., 1999.
[26]Yankee, W. J. An investigation of sphygmomanometer discomfort thresholds in polygraph examinations. *Police* 9:12–18, 1965.
[27]Zajonc, R. B., Murphy, S. T., and M. Inglehart. Feeling and facial efference: Implications of the vascular theory of emotion. *Psychol. Rev.* 96:395–416, 1989.