

Using Thermal Imaging to Detect Deception Following a Staged Theft

Cassie L. Harper^a, Madison R. Walcott^a, and Frank M. LoSchiavo^a

We conducted an experiment to test whether thermal imaging could detect deception. Half of the participants ($n = 42$) were randomly assigned to be “thieves” by stealing money from an unattended briefcase. The other half ($n = 41$) were randomly assigned to be “innocent suspects.” Immediately after the simulated theft, we interrogated all participants, although we instructed them to deny any involvement in the theft. During the interrogation, we measured each participant’s facial temperature using a consumer-grade, infrared thermometer. We hypothesized that lying would produce physiological reactions resulting in elevated skin temperatures. Although nearly all group means showed that thieves had higher skin temperatures than innocent suspects, most of the differences were not statistically significant. However, we found that thieves had statistically significantly higher skin temperatures when asked about the actual amount of money that was stolen. Thus, we found that thermal imaging was able to detect that thieves had concealed information regarding the mock crime. Thieves also reported statistically significantly more anxiety from the beginning to the end of the interrogation.

Keywords: thermal imaging, deception, mock crime, simulated theft

Because people are often motivated to lie (Ekman, 2001), the ability to detect deception is critical for effective social relations. For example, good parenting requires parents to know if their kids are lying about drug abuse. Otherwise, it would be difficult for them to determine if an intervention is necessary. Like parents, agents of the criminal justice system are particularly eager to distinguish between liars and truth-tellers, because without such knowledge, it would be difficult to determine who should be, and who should not be, charged with crimes against society.

Unfortunately, researchers have been unable to discover a unique lie response (Lykken, 1998), so instead of measuring deception directly, researchers and practitioners alike have focused on alternative measures that correlate with deception. For example, polygraphers focus on physiological measures such as heart rate, blood pressure, respiration, and skin conductivity, because each measure is at least somewhat related to normal bodily reactions following a lie (National Research Council, 2003).

Polygraph tests are used extensively in criminal justice environments, but several shortcomings render the general method problematic. For example, the polygraph requires an extensive set-up procedure, bulky equipment, intrusive electrodes, gangly wires, and an experienced examiner to administer the test and subjectively interpret the results. Furthermore, the polygraph can artifactually raise anxiety levels of innocent suspects, making results difficult to interpret, and guilty subjects can game the system using mental and physical countermeasures (Honts, Raskin, & Kircher, 1994). Thus, particularly in an era where terrorism is on the rise, there is a compelling need for additional research into new methods that utilize remote, rapid screening processes that require neither skilled staff nor direct physical contact with potential suspects (Warmelink, Vrij, Mann, Leal, Forrester, & Fisher, 2011). Because countermeasures can influence test results, new methods for detecting deception should focus on involuntary physiological responses that are nearly impossible to control.

One relatively new and potentially promising method uses thermal imaging to assess deception by measuring facial skin temperature. In this context, thermography assumes that liars will show instantaneous warming around the periorbital region of the face (Pavlidis, Eberhardt, Levine, 2002) and forehead

(Zhu, Tsiamyrtzis, & Pavlidis, 2007), due in part to a fight/flight response facilitated by the sympathetic nervous system. Because thermography can measure skin temperature without direct contact, it has several strategic advantages over techniques involving traditional polygraph procedures (Moliné et al., 2017; Park, Suk, Hwang, & Lee, 2013). Furthermore, countermeasures based on consciously controlling one’s skin temperature would likely be ineffective.

Research testing thermography following mock crimes staged at the US Department of Defense Polygraph Institute (Pavlidis et al., 2002) and at a university campus (Park et al., 2013) have yielded relatively positive results, on par with more traditional polygraph methods. However, when tested in a field study at a busy airport, thermal imaging proved less accurate than trained interviewers using a well-established interview protocol, although thermal imaging did reveal differences in skin temperature patterns between liars and truth-tellers (Warmelink et al., 2011). Thus, additional research is needed to determine the practicality of this relatively new method for detecting deception.

To test how effective thermal imaging is at detecting deception, we staged a mock theft by randomly assigning participants to be “thieves” or “innocent suspects.” During a subsequent interrogation, we measured each participant’s facial temperature as they denied their involvement in the theft. We hypothesized that lying would produce physiological reactions resulting in elevated skin temperatures. Specifically, we predicted that thieves would show higher temperatures than innocent suspects when asked questions pertaining to the theft. We also predicted that elevated skin temperatures would reveal that thieves, but not innocent suspects, harbor concealed information about the simulated crime.

Experimental Procedures Participants

Participants were 101 Ohio University Zanesville students with ages ranging from 16 to 60 years old. After data collection began, our research team realized that eyeglasses adversely affected facial-temperature readings, so we discarded data from participants who wore eyeglasses during the interrogation, leaving us with 83 participants overall (58 female, 25 male). In exchange for their participation, students received extra-credit

a. Department of Psychology, Ohio University Zanesville, Zanesville, OH, 43701
Correspondence: loschiav@ohio.edu

points equivalent to approximately 1% of the total points possible for their class.

Thermal Imaging Equipment

We measured facial temperatures with a Fluke VT04 Visual Infrared Thermometer/Camera using the “far” setting, as recommended in the user’s manual because participants were sitting more than 9 inches away. The camera is capable of measuring temperatures ranging from 14°F to 482°F (-10°C to 250°C). We positioned the camera across a table, approximately 2.5 feet from where participants were seated, and we aimed the camera directly between each participant’s eyes, just above the bridge of their nose.

Baseline Phase

Upon arriving at our laboratory, participants provided informed consent (or submitted parental consent along with minor assent) and completed an extra-credit request form that they would submit later, as part of the experimental procedure. After participants rated their baseline anxiety on a 10-point scale (1 = no anxiety, 10 = extremely anxious), the experimenter asked a series of simple “yes” or “no” questions about the person’s life (e.g., Do you have any siblings?). Once participants responded, the experimenter waited 3 seconds and then recorded their skin temperature using the thermal device.

After the baseline temperature readings were completed, the experimenter randomly assigned participants to conditions by selecting slips of paper from a nearby container. Participants assigned to be thieves were instructed to walk to the main office to submit their extra-credit request forms. Before returning to the laboratory, thieves were instructed to walk past the faculty mailboxes and to steal all the money from an unattended briefcase left on a nearby bench. When at a safe distance, thieves were instructed to count the stolen money and hide it somewhere in their clothing. Thieves were told that they would be interrogated about the crime shortly after returning to the laboratory, but they were instructed to deny their involvement in, or any knowledge of, the theft.

Participants assigned to be innocent suspects were instructed to walk to the main office to submit their extra-credit request forms and then to return to the laboratory. Innocent suspects were informed that a theft occurred and that they would be interrogated about the crime shortly after returning, but because they were not involved in the theft, they were instructed to answer truthfully by denying their involvement in, or any knowledge of, the theft.

Stimulation Test

As soon as the participant returned to the laboratory, the experimenter called an interrogator via cell phone and informed them that the participant was ready for questioning. Shortly thereafter, the interrogator entered the laboratory, introduced themselves, and stated that they were going to ask some questions about a theft that occurred by the faculty mailboxes. The interrogator also stated that the participant was named as a potential suspect in the theft.

Then, ostensibly to ensure that the thermal device was adjusted properly, the interrogator performed a stimulation test by asking participants to select one card among three options, memorize the number printed on the card, and place it in their lap face down. Importantly, the interrogator then instructed participants to answer “No” to each of the next four questions,

regardless of which card they actually selected, starting with “Did you select card number 1?” and continuing by asking about each other possible card number. After each response, the experimenter recorded the participant’s facial temperature.

After the final temperature was recorded, the interrogator instructed the participant to return their selected card to the experimenter. The experimenter recorded the card number, examined the pattern of facial temperatures, and then nodded affirmatively to the interrogator, who then stated, “OK. The calibration process is complete. Now we know how your skin reacts when you’re telling the truth, and also when you’re lying.” The stimulation test is designed to “stimulate” anxiety in guilty suspects by leading them to believe, whether or not it is true, that the interrogator can determine when they are lying.

Primary Interrogation

Following the stimulation test, the interrogator asked participants three theft-related questions: 1) “Do you know who stole the money from the briefcase by the faculty mailboxes?” 2) “Did you steal the money from the briefcase by the faculty mailboxes?” and 3) “Do you have the stolen money with you?” Participants were instructed to respond either “Yes” or “No,” and after each response, the experimenter recorded the participant’s facial temperature. We predicted that thieves would have higher temperatures than innocent suspects.

Embedded within the primary interrogation were two concealed information tests, which were designed to assess if participants had “guilty knowledge” of the theft. For example, only thieves would know that \$30 was stolen from a red wallet. The first concealed information test focused on the amount of money stolen by asking, “Do you know if the total amount of money stolen was \$10?” followed by separate questions asking about \$20, \$30, \$40, and \$50. The second concealed information test focused on the color of the wallet by asking, “Do you know if the wallet was black?” followed by questions asking if the wallet was brown, red, blue, and white. Participants were instructed to respond either “Yes” or “No,” and after each response, the experimenter recorded the participant’s facial temperature. Because only thieves knew the correct answers, we hypothesized that thieves would have higher temperatures than innocent suspects and that thieves would have the highest temperature when asked about the actual amount of money stolen (\$30) and when asked about the actual color of the wallet (red).

Following the concealed information tests, the interrogator asked each of the three theft-related questions once again. The interrogator also asked, “Have you told me the truth throughout this entire interrogation?” Participants then rated their post-interrogation anxiety on a 10-point scale. For each measure, we predicted that thieves would have higher values than innocent suspects.

Debriefing

Immediately following the interrogation, we informed all participants of the specific purpose of the study, and we thanked them for their participation. When time permitted, we also asked several additional questions for quality-control purposes, typically regarding the clarity of the instructions and regarding stress or guilt the students experienced while participating in the study. And finally, students interested in receiving a summary of the results were afforded an opportunity to submit their email addresses.

Results

Our primary goal of the study was to determine if thermal imaging could detect deception following a mock theft. Based on using random assignment, the experiment ended with 42 thieves and 41 innocent suspects. Because most of our analyses compared these two groups, we used independent-samples *t*-tests in most situations with a two-tailed alpha level set at 0.05.

Baseline Phase

The baseline phase occurred prior to condition assignment, so we expected to find no statistically significant differences between the participants who would later become thieves and participants who would later become innocent suspects. As predicted, there were no statistically significant differences between thieves ($M = 3.74$, $SD = 2.44$) and innocent suspects ($M = 4.24$, $SD = 2.36$) in terms of their baseline level of anxiety, $t(81) = 0.96$, $p = .341$. Furthermore, thieves ($M = 89.36$, $SD = 2.80$) and innocent suspects ($M = 89.20$, $SD = 2.12$) did not differ in terms of their baseline temperature readings, $t(81) = 0.29$, $p = .770$. Thus, random assignment resulted in relatively equal groups at the beginning of the study.

Stimulation Test

During the stimulation test, all participants lied about the card they selected. Although the stimulation test is a procedural hoax designed to increase anxiety in guilty suspects, we can still test if participants had the highest temperatures when asked about the card they selected. This proved not to be the case, as highest temperatures matched the selected cards just 24.1% of the time. Because participants were able to select among three available cards, the rate of agreement between the highest temperature and the card selected does not exceed levels expected by chance.

More importantly, we expected that thieves would have higher temperature readings than innocent suspects. Averaging across the stimulation test questions, thieves ($M = 88.10$, $SD = 2.64$) had higher temperatures than innocent suspects ($M = 87.99$, $SD = 1.79$), but the difference between groups was small and was not statistically significant, $t(81) = 0.21$, $p = .834$.

Primary Interrogation

The primary interrogation consisted of several theft-related questions. We predicted thieves would have higher temperature readings than innocent suspects for each question. When asked if they know who stole the money, thieves ($M = 88.49$, $SD = 2.51$) had higher temperatures than innocent suspects ($M = 88.26$, $SD = 1.70$), but the difference was not statistically significant, $t(81) = 0.47$, $p = .638$. Similarly, thieves ($M = 88.54$, $SD = 2.52$) had higher temperatures than innocent suspects ($M = 88.31$, $SD = 1.65$) when asked if they stole the money, but again the difference was not statistically significant, $t(71) = 0.49$, $p = .623$. The same pattern emerged when asked if they had the stolen money in their possession. Thieves ($M = 88.63$, $SD = 2.57$) had higher temperatures than innocent suspects ($M = 88.44$, $SD = 1.63$), but the difference was not statistically significant, $t(70) = 0.39$, $p = .699$.

Interestingly, the exact same pattern of results emerged when the interrogator asked each of the three theft-related questions once again, a bit later in the experimental procedure. All the group means were in the predicted direction (i.e., thieves had higher temperatures than innocent suspects), but no

statistically significant differences were found. Specifically, all *t*-values were less than 1.0 and all *p*-values were greater than 0.05.

Concealed Information Test

To assess if thermal imaging could uncover the fact that thieves were harboring guilty knowledge regarding the specific details of the theft, we looked to see if thieves had the highest temperatures when asked about the actual color of the wallet (red) and when asked about the actual amount of money stolen (\$30). Although thieves failed to register their highest facial temperature when asked about the actual color of the wallet, they did register their highest facial temperature when asked about the actual amount of money stolen.

Thus, among the thieves, we compared their temperature to the \$30 question ($M = 88.85$, $SD = 2.55$) with their average temperature for the other remaining options ($M = 88.73$, $SD = 2.52$) using a related-samples *t*-test and found statistically significant results, $t(41) = 2.15$, $p = .038$. Although the temperature difference is quite small, it coincides with a standardized effect size of $d = .33$, which according to Cohen (1988) falls somewhere between small and medium.

Although this within-subjects analysis was sensitive enough to detect subtle differences in facial temperatures of thieves, between-subjects analyses comparing the two groups directly continued to fall short. Similar to the pattern of results obtained during the primary interrogation, the group means for the concealed information test were in the predicted direction (i.e., thieves had higher temperatures than innocent suspects, suggesting thieves possessed guilty knowledge of the mock crime), but no statistically significant differences were found. Specifically, all *t*-values were less than 1.0 and all *p*-values were greater than 0.05.

Posttest Anxiety Levels

Earlier we reported no differences between thieves and innocent suspects in terms of their baseline anxiety levels. But after random assignment to conditions and the stress of the interrogation, we expected that thieves would report greater anxiety. Although thieves ($M = 4.74$, $SD = 2.28$) reported higher levels of anxiety than innocent suspects ($M = 3.98$, $SD = 2.12$), an independent-samples *t*-test failed to reach traditional levels of statistical significance, $t(81) = 1.58$, $p = .118$.

However, within-subjects analyses tell a more nuanced story. From pretest ($M = 4.24$, $SD = 2.36$) to posttest ($M = 3.98$, $SD = 2.12$), anxiety levels of innocent suspects did not change, $t(40) = 1.12$, $p = .269$, presumably because these participants harbored no guilty knowledge nor did they fear interrogation, because they did not commit a crime. Thieves, on the other hand, had valid reasons for being anxious, and as predicted, they did report more anxiety from pretest ($M = 3.74$, $SD = 2.44$) to posttest ($M = 4.74$, $SD = 2.28$), $t(41) = 3.28$, $p = .002$, suggesting that our manipulation was at least somewhat effective. This increase in anxiety coincides with a standardized effect size of $d = .51$, which according to Cohen (1988) is medium in its magnitude.

Discussion

Although some researchers have suggested that thermal imaging may be a promising new method for detecting deception (e.g., Rajoub & Zwiggelaar, 2014), our mixed results

Research Article (2019)

leave us somewhat skeptical. While all of our between-group comparisons showed group means in the predicted direction, not one comparison resulted in a statistically significant difference between the facial temperatures of thieves and innocent suspects. Based on these findings, it is difficult for us to see how thermal imaging can be useful in remote, real-time, real-world settings where measured differences between liars and truth-tellers must be large enough to generate confidence in one's decisions and clear enough to be actionable.

That said, our study had several limitations, and it is likely that those limitations contributed to our mixed results. For example, all of our data came from a convenience sample consisting of undergraduate students from a small, midwestern campus. A wider sampling of participants would likely represent the population more appropriately. Perhaps more importantly, our sample size left us with insufficient statistical power. Systematic differences between liars and truth-tellers have proven to be elusive, suggesting that whatever true effect sizes exist, they are likely to be small, and large samples may be necessary to measure them.

Although the procedure was carried out by novice researchers (i.e., students in a research methods course), one promise of thermal imaging is that it is user friendly with basic instruction, making it relatively easy for anyone to use—just point and shoot. However, not all thermal imaging devices are created equally. We selected a unit that is available commercially, and although it was not inexpensive (it retails for approximately \$500), we chose to test a device that is widely accessible beyond police departments and other large government agencies (e.g., FBI, CIA). For simplicity, we used the device to capture general surface temperatures only. It is possible that we would have been more successful if we used the device to capture video during the interrogation. It is also possible that we would have been more successful if participants were able to respond freely during the interrogation instead of being limited to simple “yes” or “no” responses. However, analyzing long video segments and complex verbal responses was beyond the scope of our course.

Although thieves reported statistically significantly more anxiety from the baseline phase to the posttest phase, and although the theft was about as realistic as it could have been on a college campus, the mock crime was ultimately an artificial event. It is possible that in real-world environments, thermal imaging might produce more valid results. Thus, future research should be conducted in more applied settings, such as during interrogations with actual criminal suspects. Criminal interrogations are video recorded on a regular basis. Thermal data could be recorded just as easily.

We also recommend that future research focus on concealed information tests, as that was one area of inquiry where we found statistically significant results. Having actually committed the simulated theft, thieves possessed guilty knowledge of the crime scene, and our thermal device was able to detect it. Concealed information tests make sense for a variety of reasons. First and foremost, they can be conducted on an individual suspect. In other words, it is not necessary to compare a suspect with anyone other than themselves. Thus, concealed information tests naturally lend themselves to real-world investigations, where a suspect is interviewed, and a determination of the suspect's truthfulness must be made.

Over recent years, several studies have investigated thermal imaging as a lie-detection tool, and just like our own

study, there is room for improvement in terms of devices tested and methods used. Until there is more positive data, we will remain skeptical, but optimistic about the potential promise of this relatively new technique. As researchers continue to search for a valid method of detecting deceit, it is important to remember that no matter what tool or technique is used, a unique lie response might not exist, so detecting deception will almost certainly remain elusive.

Acknowledgments

This research was conducted as part of a class project during the 2019 spring semester at Ohio University Zanesville. The authors would like to thank the other members of the research team, including Mackenzie Howe, Merry Murphy, Ryan Rider, and Kevin Winkle.

This manuscript has not been published elsewhere, nor is it currently under review elsewhere, nor does it pose any personal or professional conflict of interest.

References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (3rd ed.). New York: Norton.
- Honts, C. R., Raskin, D. C. & Kircher, J. C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, 79, 252–259. doi:10.1037/0021-9010.79.2.252
- Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector*. New York: Plenum.
- Moliné, A., Gálvez-García, G., Fernández-Gómez, J., De la Fuente, J., Iborra, O., Tornay, F., ... Gómez Milán, E. (2017). The Pinocchio effect and the Cold Stress Test: Lies and thermography. *Psychophysiology*, 54, 1621–1631. doi:10.1111/psyp.12956
- National Research Council. (2003). *The Polygraph and Lie Detection*. Washington, DC: The National Academic Press.
- Park, K. K., Suk, H. W., Hwang, H. & Lee, J-H. (2013). A functional analysis of deception detection of a mock crime using infrared thermal imaging and the Concealed Information Test. *Frontiers in Human Neuroscience*, 7(70), 1-17. doi:10.3389/fnhum.2013.00070
- Pavlidis, I., Eberhardt, N. L. & Levine, J. A. (2002). Seeing through the face of deception. *Nature*, 415, 35.
- Rajoub, B. A., & Zwiggelaar, R. (2014). Thermal facial analysis for deception detection. *IEEE Transactions on Information Forensics and Security*, 9, 1015-1023. doi:10.1109/TIFS.2014.2317309
- Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D. & Fisher, R. P. (2011). Thermal imaging as a lie detection tool at airports. *Law & Human Behavior*, 35, 40-48. doi: 10.1007/s10979-010-9251-3
- Zhu, Z., Tsiamyrtzis, P., & Pavlidis, I. (2007). Forehead thermal signature extraction in lie detection. *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 243-246. doi:10.1109/IEMBS.2007.435226