Eye Gaze Behavior as a Guilty Knowledge Test: Initial Exploration for Use in Automated, Kiosk-based Screening

Douglas C. Derrick
The University of Arizona
dderrick@cmi.arizona.edu

Kevin Moffitt
The University of Arizona
kmoffitt@cmi.arizona.edu

Dr. Jay F. Nunamaker, Jr.
The University of Arizona
jnunamaker@cmi.arizona.edu

Abstract

This paper presents findings from an eye tracker experiment based on the guilty knowledge test (GKT) paradigm. We conducted a two-group, posttest-only, randomized initial experiment in which the guilty group constructed an improvised explosive device (IED) and then was screened by an embodied conversational agent (ECA). Participants in the control condition went straight to screening. During the screening interview, participants were shown three images. The second image was a modified picture of an IED similar to the one constructed by the guilty participants. Eye gaze behavior was dramatically different between the groups. A classification algorithm correctly classified 100% of the participants as either part of the guilty or control group.

1. Introduction

The purpose of this study was to evaluate the potential of using eye gaze behavior to discover people with guilty knowledge. Recent research has indicated that people involuntarily give away indications of their familiarity with items through psychophysiological cues that can be detected by a variety of electronic sensors [1]. In this paper, we discuss eye-tracking technology as a means to determine if an individual has guilty knowledge of a prohibited improvised explosive device (IED).

Specifically, we evaluated eye-tracking technology to determine its ability to detect guilt and possible threats within an automated screening paradigm. Moreover, we are investigating the viability of this tool for environments with high-volume pedestrian traffic. Automating a portion of the screening process in high traffic areas could aid human agents by taking over low-skilled repetitive tasks and tasks a human is not capable of performing such as eye tracking. The effects would be increased efficiency and added depth to the screening task. The eye tracker may be able to support law enforcement and security screeners by providing behavioral feedback from interviewees. In this study, the eye tracker was embedded in a kiosk platform that interviewed people in a rapid assessment scenario and made initial assessments of a person's credibility, honesty, and intent. Avatar-like screeners that are controlled by intelligent agent software conduct the screening interviews. Our architecture is related to well-known intelligent agent architectures, with some key distinctions [2][3]. Like most intelligent agent systems, the paradigm for embodied-avatar interactions with humans involves an agent that perceives its environment through sensors, influences its environment via effectors, and has discrete goals. The purpose of this experiment was to evaluate the eye tracker as a potential sensor for use in automated, kiosk-based screening.

The remainder of the paper will proceed as follows: (1) we discuss using the eye tracker to conduct a Guilty Knowledge Test, (2) we explain the eye tracking technology, (3) we discuss the experimental methodology and (4), we review the results and limitations of the experiment.

2. Guilty Knowledge Test

Recent research supports the claim that eye gaze behavior can be used for memory assessment. For example, participants' eye movements when viewing familiar images differ from their ocular responses to unfamiliar images. These differences include less frequent fixations in fewer regions and more randomness in eye movement patterns. Althoff and Cohen [5] found differences in people's responses to familiar versus novel faces. Because of these findings, in this study we use eye tracker technology to conduct a version of the Guilty Knowledge Test (GKT) [6].

The GKT was developed to investigate situations in which only a guilty person would possess a particular piece of knowledge. A GKT is usually performed by asking multiple-choice questions with one correct answer and several plausible answers, also known as foils. For example in a GKT, a murder suspect might be asked to read out loud the names of six possible murder weapons. An increased physiological response in the examinee when he or she reads the name of the actual murder weapon should indicate guilty knowledge. A person who is innocent should exhibit a consistent physiological response to each possible weapon since to the innocent person the real weapon is just as plausible as any of the others in the list. Studies have shown that the GKT can accurately identify those who possess incriminating information while protecting the innocent [4]. However, in its current form, a GKT is ill suited to a rapid screening environment because (1) it requires a large number of questions, and (2) it must be administered by a trained professional.

In this study, we modeled an eye tracker test after the GKT for use in rapid screening. The goal of this experiment was to learn if eye behavior could be used as an indicator of guilty knowledge. In order to build a GTK eye behavior test, we created a scenario in which some participants constructed an IED and then viewed an altered image of the explosive device. An innocent participant (who did not construct the IED) should find the altered image unremarkable and that should be reflected in their gaze patterns. On the other hand, guilty participants (who did construct the IED) should be aware that something had been modified and therefore have their gaze drawn to the altered area of the image.

H1: Those familiar with the IED will spend more time looking at the altered region of the image than those unfamiliar with the IED.

3. Eye Tracking with Near Infrared

Our test design is based on orienting theory and predicts that measurable physiology accompanies and orienting reflex to familiar stimulus [9]. The orienting reflex was originally studied by Pavlov during his classical conditioning experiments [7]. This reflex orients attention to novel and familiar stimuli and is considered adaptive to the environment [8]. In order to capture the eye behavior responses, we used the EyeTech TM3 (see figure 1) mounted directly below a 19” computer monitor.

Figure 1 – EyeTech TM3

The TM3 has two infrared light sources and an integrated infrared camera. It connects via USB to a Windows computer and captures the eye gaze location (x, y coordinates) at each instance at a rate of approximately 33-34 frames per second.

4. Bomb Making Experiment

3.1 Sample

Volunteers were solicited at a large southwestern university undergraduate MIS classes. Before the test was conducted, four student volunteers validated the instructions and processes. After we were confident in the processes, we recruited 12 participants for this initial study.

3.2 Procedures

The experimental data were collected using a straightforward protocol that required some participants to assemble a realistic, but not operational, improvised explosive device (IED). The experimental design is a two-treatment, between-group design. The first treatment was the control group and participants in this group were in the non-bomb-making condition and therefore completely unfamiliar with the IED. In the second group, participants became familiar with a simulated bomb and bomb making materials and actually assembled the device. The treatments were randomly assigned and the experiment participants in the bomb-making condition received the bomb-making materials and a typed page with the assembled bomb image at the top (shown in figure 2). The instructions were as follows:

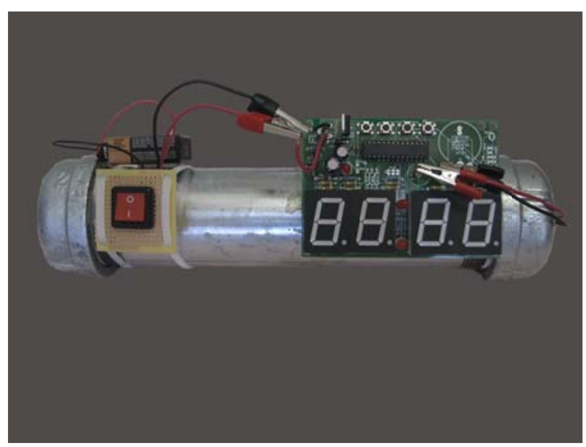

Figure 2 – Completed IED

“*You will construct the IED pictured above with the materials provided to you. Follow the steps below in exact order to replicate the IED shown above.*
Materials list:

- *Pipe*
- *Timer*
- *Battery*
- *Switch*
- *Zip ties*

1. *Orient the switch so the "1" is on the bottom. Firmly attach the switch to the left hand side of the pipe with two zip ties. Make sure the back of the switch is pressed against the white piece of Velcro already attached to the pipe. Make sure you don't break the switch module by tightening the zip ties too tight.*

2. *Attach the 9V battery to the 9V battery connector coming from the switch module.*

3. *Attach the Velcro on the 9V battery to the pipe above the switch module as shown in the picture. Make sure the connections on the battery are facing to the left (outward).*

4. *Orient the timer so the 4 digital numbers are at the bottom. Attach the timer to the pipe by placing the white Velcro on the back of the timer onto the black Velcro on the metal pipe. Position the right edge of the timer flush with the inside edge of the pipe cap.*

5. *Clip the red alligator clip coming out of the end of the pipe to the red wire coming out of the timer as shown above. Do the same with the black clip.*

6. *Clip the red alligator clip coming from the switch to the red wire coming from the left side of the timer as shown in the picture above. Do the same with the black alligator clip.*

7.

Participants took approximately 5-7 minutes to assemble the device. After the "bomb" was completed, the participants went to an automated screening station. Those in the control group did not construct an IED and went directly to the screening station to begin the automated interview. At the station, experiment personnel used a brief calibration program to calibrate the eye-tracking device to the participants' eyes.

The automated agent communicated the following messages:

1. Please state your full name.
2. Are you a citizen of the United States?
3. Where are you travelling from?
4. What was your business there?
5. Do you have anything to declare?
6. Please carefully examine the following images. < The following images (figures 3, 4, and 5) were then displayed for 12 seconds each. >

Figure 3 – First Image Shown to Participants

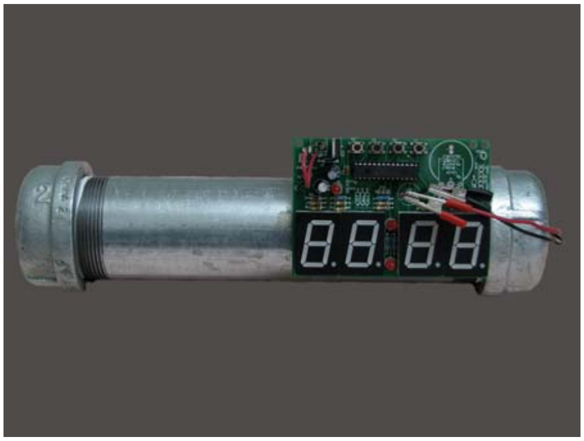

Figure 4 – Second Image Shown to Participants

Figure 5 – Third Image Shown to Participants

7. Have you ever seen a device similar to this image? < Figure 4 is repeated >
8. Please see the officer at the next available station. Thank you for your cooperation.

The participants were then debriefed, dismissed and the bomb was disassembled. The first and third images were used as basic foils and to allow for task acclimation. The key image of interest is figure 4. Please note the differences in figures 2 and 4. Figure 2 shows the IED that the participants assembled. Figure 4 is the same device, but the button, the battery, and the connecting leads have been removed to make this image novel to participants who assembled the device.

4. Analysis

As mentioned earlier, the images were shown for 12 seconds each and the participants' eyes were sampled at a rate of every ~30ms (33-34 samples per second). For each sample, we captured the (x, y) gaze location on the screen (in pixels) for both eyes and this four-tuple is denoted as s_n below. We captured the total number of samples for the participant denoted by set *P*. Based on the image, screen size and resolution, we determined that the region of interest was any pixel on the x-axis less than 650 (the region where the switch was located). For every sample (s_n), we calculated the average x coordinate using the gaze position of each eye (x_{an}) and then determined if it was in the region of interest denoted by the set *I*.

$s_n = (x_{n\text{-left-eye}}, x_{n\text{-right-eye}}, y_{n\text{-left-eye}}, y_{n\text{-right-eye}})$
$P = \{s_1, s_2, \dots, s_n\}$
$x_{an} = (x_{n\text{-left-eye}} + x_{n\text{-right-eye}}) / 2$
$s_n \in I \textit{ iff } x_{an} < 650$

We then calculated the percentage of the samples in region of interest using the formula below:

$$\frac{|I|}{|P|}$$

5. Results

Of the 12 participants in the experiment, only 11 were usable. The eye tracker was unable to detect and track the 12th participant's eyes. During the post interview, the participant said that she suffered from lazy eye and this may explain why the technology failed. Using the metric created above, we created a box plot that compares the two groups (see figure 7).

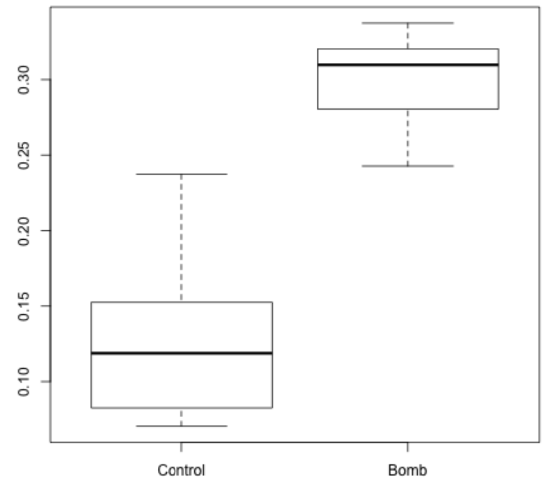

Figure 7 – Box Plot of Results

According to the plot, there is a marked difference in eye gaze behavior between the two groups. Analysis shows that *Mcontrol* = 13.00%, *SDcontrol* = 6.06%, while *Mbomb* = 29.81%, and *SDbomb* = 3.72%. A Welch two-sample T-test shows that the hypothesis was strongly supported T = -6.64, df = 8.403, $p < .001$. The participants in the bomb-making condition gazed much more at the altered portion of the image. Based, on inspection of the data, we created a classification cut line to discriminate between the two groups at 24%. Those who had more than 24% of their samples in the region of interest were classified as in the bomb group, while those with 24% and less were classified in the control group. Using this line, the eye behavior based GTK yielded 100% accuracy in discriminating between those that had built the bomb and those that had not.

	Bomb	***Control***
Bomb	5	0
Control	0	6

Table 1 – Classification Results

We also explored using the control group as a way to perform classification. Using a cut line of 19.06% (the control group mean + one standard deviation) we achieved accuracy of 91% with only one misclassified participant (a false positive). Using a cut line of 25.12% (the control group mean + two standard deviations) we achieved accuracy of 91% with only one misclassified participant (a false negative).

5. Discussion and Future Directions

While preliminary, the results are very promising for a GKT test. We demonstrated with 100% accuracy the ability to discriminate between those who had constructed and were familiar with the IED from those who were unfamiliar with the device. To further validate our findings we will continue to recruit participants to increase the sample size. Because of the encouraging results of this study, in the future this technology will be field tested at U.S./Mexican border crossings with where high-traffic pedestrian screenings are the norm.

In addition, this test protocol will be applied to images of places and people's faces. For example, in a digitally altered photo of a crime scene, a criminal may focus on the area of the image where an object should have been. In a photograph of a face with a digitally removed a mole or scar, a person familiar with that face may focus on the altered area longer than a person who is unfamiliar with that face. These types of test could help agents and investigators determine what knowledge an individual has about a person, or place.

6. Limitations

Despite the small sample size, there are several limitations to this research. First, there should be an examination of the temporal aspect of the effect. This involves two main analyses:

- How long does a person need to examine the image in order to determine familiarity with the image?
- Does the guilty knowledge effect last over time?

In our experiment, the participants went to screening immediately after assembling the bomb. We need to understand if the effect lasts after an hour, a day, or a week. Second, in this experiment, we used a nearly exact image of the IED. The phenomenon needs to be examined with images that are close or related, but not identical. In this study, all of the participants were cooperative during the eye calibration process. This process needs to be automated before the system can be used in the field. We also need to consider methods for detecting and/or counteracting possible countermeasures for fooling the system. Because of the small sample size of this study, we were unable to determine any effects due to impaired vision, gender, age, or cultural effects. Some final concerns with this technology include the possibility that people may be less able to discriminate known versus novel stimuli when these objects originate from their non-native culture. Additionally, it is unknown how well matched test items must be to target items for the appropriate response to register.

7. Conclusion

It is believed that the results from this experiment will provide a method for border security agents to rapidly identify individuals who may have information relating to dangerous or illegal materials Its use in rapid screening would require much greater control over conditions for interviewee responding than was achieved here in the semi-structured phase of the interview.

In sum, eye tracking as a GTK is a promising technology with some limitations. It deserves further study and scientific examination in a rapid credibility assessment environment.

8. Acknowledgements

This research is supported by the US Department of Homeland Security through the National Center for Border Security and Immigration under grant number 2008-ST-061-BS0002. However, any opinions, findings, and conclusions or recommendations herein are those of the author and do not necessarily reflect views of the US Department of Homeland Security. The views, opinions, and/or findings in this report are those of the authors and should not be construed as an official U.S. Government position, policy, or decision.

9. References

[1] D.C. Derrick, A.C. Elkins, J.K. Burgoon, J.F.N. Jr, and D.D. Zeng, "Border Security Credibility Assessments via Heterogeneous Sensor Fusion," *IEEE Intelligent Systems*, vol. 25, 2010, pp. 41-49.

[2] M. Wooldridge, *Intelligent Agents*, 1999.

[3] M. Wooldridge and N.R. Jennings, "Intelligent agents: Theory and practice," *The Knowledge Engineering Review*, vol. 10, 1995, pp. 115-152.

[4] G. Ben-Shakhar and E. Elaad, "The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review.," *Journal of Applied Psychology*, vol. 88, Feb. 2003, pp. 131-151.

[5] R.R. Althoff and N.J. Cohen, "Eye-movement-based memory effect: A reprocessing effect in face perception.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 25, Jul. 1999,

pp. 997-1010.

[6] A.H. Ryan Jr, I. Pavlidis, J.W. Rohrbaugh, F. Marchak, and F.A. Kozel, “Credibility assessments: operational issues and technology impact for law enforcement applications,” E.M. Carapezza, Ed., Orlando, FL, USA: SPIE, 2003, pp. 168-182.

[7] I. Pavlov, *Conditioned Reflex*, Oxford, England: Clarendon Press, 1927.

[8] Y. Sokolov, *Perception and the Conditioned Reflex*, Oxford, England: Permagon Press, 1963.

[9] A. Vrij, *Detecting lies and deceit: Pitfalls and opportunities*, West Sussux, England: Wiley-Interscience, 2008.