A new approach to solving crimes

Do the eyes of a suspect reveal if they are guilty of a crime?

Bachelor Thesis

Clara Büttner

s1836706

c.buttner@student.utwente.nl

Behavioural, Management and Social Sciences

Department of Psychology of Conflict, Risk and Safety

1st supervisor: Anouk den Besten

2nd supervisor: Dr Mariëlle Stel

- 26.06.19 -

UNIVERSITY OF TWENTE.

Abstract

The aim of this study was to find out if eye-tracking might be a suitable alternative for polygraphs and if it can support finding a difference between guilty and innocent regarding the detection of deception. Therefore, a simulated interrogation in a lab setting was carried out. This study consisted of three conditions that were compared, i.e. a guilty condition, an innocent and an eyewitness condition. After participants read a scenario of a mock crime, the police interrogation took place in which participants needed to look at pictures of crime-relevant and crime-irrelevant objects as part of a Guilty Knowledge Test. During this task, their eye-fixation was measured. Overall, the outcomes show that eye fixation duration did differ between the innocent and the guilty condition, nevertheless, no significant difference was found for the witness condition. Results also support that participants in the guilty condition perceived the most stress during the interrogation. The findings indicate that eye-tracking might be a suitable measurement to differentiate between someone who is guilty and someone who is innocent and does not possess guilty knowledge, however, innocent people that do possess guilty knowledge are left out in this process. Hence, future research should take on here and see whether there might be better ways to integrate innocent people with guilty knowledge.

Keywords: deception, eye-tracking, eye fixation, Guilty Knowledge Test, stress

Table of Contents

Do the eyes of a suspect reveal if they are guilty of a crime?

Crime is something that exists throughout the history of our population. Even though it seems to have always been a problem, we still do not have accurately working solutions for solving crimes. Police officers daily face the challenge of determining if a suspect is guilty or not, especially since most methods to do so are not objective enough as shown in a study by Hartwig, Granhag, Strömwall and Vrij (2004). They compared two methods of judging if someone is guilty. In the first group, police officers reached a decision based on a live interrogation and in the other group, they judged the person based on a video interrogation. However, neither one of these groups detected deception better than chance. Still, research on the distinction between guilty and innocent goes back a long way.

As early as 1911, it was thought one could distinguish criminals from non-criminals. Lombroso (1911) stated in his theory of the criminal man that all criminals would look alike. Moreover, he even linked specific features, like a sloping forehead or thick lips to specific crimes, in this case, to murder and sex offending.

A similar theory is the constitutional one by Sheldon (1942). He concluded that the body type of offenders differed from that of non-offenders. However, Reed (1973) claimed that the study of Sheldon (1942) may have suffered from the 'halo' effect since body shape and temperament were evaluated by the same person. The 'halo' effect describes the phenomenon that an earlier judgement of a person also influences later judgement in a different area (Nisbett & Wilson, 1977). This again supports that finding an objective method of judgement is difficult. Hence, a new, more reliable and less subjective method needed to be found and one rather concentrated on detecting deception within police interrogations.

Lykken (1959) was the first to find a connection between the possession of guilty knowledge and physical responses to the accompanying emotions such as, *inter alia,* fear, nervousness, insecurity or guilt. From then onwards, people began to direct their research on the physiological responses of these emotions like increased heart rate and higher skin conductance. Due to that, one could say that lie detectors do not actually detect lies but rather the physiological responses to the feelings that are believed to accompany lying. Polygraphs, for instance, measure skin conductance and heart rate, which are believed to be indicators for the possession of guilty knowledge if they are increased (Lykken, 1959).

However, there are serious flaws in these designs. For instance, if one is simply nervous because of the attitude of the interviewer or because one fears to be wrongfully accused of something, this can also evoke increased physiological responses, which can then wrongfully be interpreted as lying (Barnes, 2008). Therefore, nowadays, polygraphs are not accepted as sufficient evidence anymore. In fact, they are believed to have a 50% rate of false positives (someone wrongfully detecting as a liar) (Horvath, 1977).

Apart from the polygraph, there are also other, newly evolving measures of deception. Working with eye-tracking technology is one of them. Several studies tested eye-tracking successfully (Russo & Leclerc, 1994; Russo & Rosen, 1975). However, this research was mostly done in the fields of marketing and consume. Little research is done on this topic in connection to social psychology and deception. Due to the promising results in other studies like the one by Russo and Leclerc (1994), it would be interesting to see if eye-tracking could maybe also pay a contribution to the work with guilty knowledge and may even provide a more objective method than working with the other physiological measures mentioned before. Therefore, this study aims to find out if eye-tracking can be the new objective and accurately working measurement in the detection of deception that has been searched for. Furthermore, it would be interesting to see how eye fixation behaviour differs between innocent and guilty people in interrogation situations. But in particular, how can witnesses be differentiated from the guilty and innocent group by their eye-fixation?

In the current study, this will be assessed with the help of a Guilty Knowledge Test (GKT). The GKT is a tool for the detection of guilty knowledge. It aims at provoking increased physiological responses to some of its items, which then indicate familiarity to these items or a connection in another way (Ben-Shakhar & Elaad, 2002). This will be discussed in more detail in the next section.

Theoretical Framework

The Guilty Knowledge Test

To evoke physiological responses that can be measured during, for instance, an interrogation, different techniques are in use. One of them is the Control Question Test (CQT). During the CQT, a person is asked questions that are directly related to the crime (e.g. “Did you

rob the bank?"), emotionally provocative control questions (e.g. "Have you ever hurt someone?"), and neutral questions such as "Is today Tuesday?". A larger physiological response to the crime-related question compared to the control and neutral one indicates that the subject is lying. Hence, a larger physiological response to the control or neutral question indicates the innocence of the subject (Meijer, Smulders, Johnston, & Merckelbach, 2007; Verschuere, Crombez, & Koster, 2004). However, the CQT was often part of the critique of lie detection. This mainly stems from the fact that it is not standardised and has no real theoretical framework and, thus, produces a high percentage of false positives.

Most studies nowadays, due to that, rather use the aforementioned Guilty Knowledge Test (GKT), which is also referred to as Concealed Information Test (CIT), to detect deception. Instead of using different kinds of questions, it only uses questions related to the incident under investigation. Moreover, it gives the person multiple, normally four, answer options of which one is correct. An example of a GKT question would be: What weapon was used during the robbery? a) knife b) handgun c) acid d) knuckle duster. These answers are then expected to evoke physiological responses in the interviewee. For innocent people, all of these answers must seem equally possible, however, for a guilty person or a person that possesses guilty knowledge, one answer is expected to be more physically arousing (Ben-Shakhar & Elaad, 2002).

Nevertheless, what is important here in connection with, for instance, eyewitnesses, is that if the test only measures the recognition of its items, it would be likely that people that are not guilty but do possess guilty knowledge, may be classified as guilty based on the test results. Bradley and Warfield (1984) conducted a study where they compared the results of a GKT (in combination with a polygraph) with several conditions. They had one condition where the participants were guilty of a mock crime, three conditions whose participants possessed guilty knowledge but were innocent (eye-witness; read about the crime; performed tasks with behaviours similar to the ones of the guilty condition) and one uninformed innocent condition that acted as the control group. After the participants read their scenarios, they were informed that they are suspects of that crime and a police interrogation was following. All conditions were instructed to deny any responsibility and that they will be rewarded with $5 if they pass the lie detector test. Bradley and Warfield (1984) found that the groups that were informed about the crime and, thus, possessed guilty knowledge, were not detected as guilty based on the test results. Moreover, the innocent condition, with no information about the crime scored even lower than the innocent informed condition. This indicates

that only guilty people are detected by the test. Regarding this, it would be interesting to see if similar results can be obtained with eye-tracking as the main measurement. Therefore, in the current study, also an innocent informed group will be included, namely, eyewitnesses.

Eyewitnesses

Eyewitnesses are of most importance in cases under investigation, however, they may not give up their information voluntarily. It may be that someone close to them committed the crime and, thus, they want to protect the suspect. On the other hand, it could also be that if they tell the truth they might lose something, or they were threatened to lie. In this case, it would be interesting to research how one can detect that they are lying since they might not have the same physiological responses, for instance in a polygraph, as people that are guilty themselves. Therefore, it will be interesting to see if also their eye behaviour might differ from people that are guilty or innocent without the possession of guilty knowledge.

Eye-tracking in connection with deception

As already mentioned, there are only a few studies that made the link between eye tracking and deception. One of these studies was conducted by Millen, Hope, Hillstrom, and Vrij (2017). They tested eye-movements regarding the recognition of faces with three different groups of faces; familiar faces, unfamiliar faces and control faces (e.g. celebrities). The outcomes of their study showed that when people saw familiar faces, they produced significantly fewer eye fixations compared to the unfamiliar or neutral stimuli. They claimed that this was due to the memory load connected to these faces. Despite this, the study also had some limitations. For instance, they could not be sure if the control faces were really working as such since they could not determine with certainty that every participant really knew the face of the celebrity.

Another study which used eye tracking in combination with deception was the one by Derrick, Moffitt and Nunamaker (2011), but instead of eye fixation, they measured eye gaze. They had two conditions, the first one assembled a fictional bomb based on instructions and the second one was the control group that was unfamiliar with this explosive device. All participants then saw pictures of two bomb-unrelated objects and a picture of the bomb that the first group assembled earlier. However, in this picture, a part of the bomb was missing to make this stimulus novel for the experimental condition. They found that people from the guilty condition gazed much more on

the area that was modified than the condition that was unfamiliar with the device. Even though they produced promising results, one has to take the small sample size of 12 participants into account. In this study, it would have also been interesting to consider eyewitnesses as a third condition and see how they might have differed from the other two groups.

Furthermore, the study by Peth, Kim and Gamer (2013) did not only use the CIT in combination with eye fixation but also measured the number of eye blinks during the test. The study had two conditions, one that was guilty of a mock crime and one that was innocent. They conducted two CIT's, one directly after the mock crime was committed and one delayed one. The results showed longer fixation duration in the crime-relevant items than in the neutral ones and a reduction in eye blinks after the crime-relevant stimulus was presented for the guilty condition. Thus, they were able to differentiate between the guilty and the innocent participants in both CIT's.

Schwedes and Wentura (2016) found similar results. They instructed participants to steal specific objects in a virtual mock crime. Afterwards, participants had a learning phase in which they interacted with these stolen objects, and also with some control objects that were labelled by the experimenters as "gift" objects. This phase was followed by a simulated police interrogation in which they were instructed to identify the objects they received as gifts but hide their knowledge about the stolen objects. They found that the stolen items were fixated significantly longer than the gift items. This supports the findings of the aforementioned studies (Millen et al., 2017; Peth et al., 2013) and suggests that fixation duration is a suitable measurement to differentiate between guilty and innocent. Therefore, this will also be a main component of the current study.

Eye-fixation

Eye fixation, as the word implies, means that one is fixating a special point for longer. The eye movements in between these fixations are called saccades and are measured as well (Van der Lans, Wedel, & Pieters, 2011). This can especially be helpful in terms of attention. If there is stronger eye fixation at one point, this indicates that this point may be of special interest for the subject and that it caught the subject's attention (Spain & Perona, 2008). Humans are, in fact, able to control their attention but sometimes non-intentional eye movements are made that can reveal something about the viewer's state of mind or in this case, about the possession of guilty knowledge (Duchowski, 2007).

In the study by Millen and colleagues (2017), which was mentioned earlier, also eye fixation was used as the main measurement. They stated that "there are fewer fixations during processing of familiar than during processing of unfamiliar faces." (Millen et al., 2017, p. 932). This is further supported by the study conducted by Peth and colleagues (2013). They also found that the guilty condition had fewer eye-fixations in the areas that were related to the mock crime than in the innocent group. However, these were significantly longer. Therefore, in the current study, it is tested whether eye fixation duration is longer in areas/pictures that are crime related/familiar to the subject than in other areas.

The current study

The current study makes use of the GKT and eye-tracking technology. There will be three conditions: the guilty, the innocent and an eye-witness group. Guilty knowledge will be manipulated by using a scenario and a video fitting to a mock crime. The guilty group will see a gun in the video which has a blue dot on its grip, the witness group sees the gun without the blue dot and the innocent group does not see the gun at all. Afterwards, the groups are interrogated by showing them pictures of different guns, of which one was the gun they had seen in the video. In regard to the guilty condition, there is a slight change since the picture of the gun during the interrogation is missing the blue dot. During this part, eye-fixation will be measured with a wearable eye-tracker.

Based on the aforementioned studies and their outcomes, it is expected that the guilty condition has a longer fixation duration in the area of interest one (AOI_1), in this case, the area at the weapon where the blue dot should have been. Moreover, since the witness group does not really see the gun in all its details, it is expected that they do recognize the gun and, hence, stare longer at the picture of the gun used during the incident (AOI_2) than participants to whom the gun is not familiar (innocent condition). However, they do not specifically fixate AOI_1. For the innocent group, it is expected that their eye fixation does not significantly differ regarding the pictures.

Hypotheses:

1. Eye fixation duration is longer in the AOI_1 for the guilty condition in comparison to the witness and innocent group

2. Eye fixations on the AOI_2 are more spread in the innocent condition compared to the witness and guilty condition

Emotion: stress

Research suggests that the perceived level of stress has an influence on the detection of deception. Especially situations like interrogations, detentions or testing (e.g. polygraph tests but also exams *et cetera*) can evoke high levels of stress in the person of attention (Tyson, 1983). There is evidence that stress also has biological effects on the body as changes in cortisol levels, skin conductance or blood pressure (Groer et al., 2010). These would, in turn, have an influence on a polygraph test where skin conductance is measured. One could not differentiate if the increased skin conductance stems from the lying process itself or from their perceived stress (Barnes, 2008). Moreover, according to Deffenbacher, Bornstein, Penrod, and McGorty (2004), the fidelity of eye-witness testimonies is negatively influenced by the stress perceived by them. Thus, witnesses may make flaws in identifying the suspect or describing the details of a crime. Due to that, in the current study, perceived stress will be measured in a post-experiment questionnaire to determine if it might have had an effect on the eye-movements.

Furthermore, to look at guilt may also be interesting, however, MacLaren (2001) claimed that even though it makes the participant more prone to the detection of guilty knowledge if he/she has committed a mock crime, the feeling of guilt is not necessary for the detection. Therefore, in the current study, guilt will not be taken into account and it will be solely concentrated on the perceived stress during the interrogation.

In the current study, it is expected that the guilty group does experience more stress than the other two groups during the interrogation. Additionally, it is expected that participants in the innocent group will experience the least amount of stress. This leads to the third hypothesis.

Hypothesis:

3. Participants of the guilty and eye-witness condition score higher on perceived stress than the innocent condition. This effect is expected to be stronger for the guilty condition than for the witness condition

Lastly, the 'interviewer effect' is something that should be avoided when conducting experiments, including interrogations, wherefore, in the current study, this possibility will be controlled for in the form of questions about the experimenter's attitude and behaviour towards the participants.

Methods

Participants

In this study, 38 participants were included. Most of them were students from the University of Twente (50% male, 50% female). Their age ranged from 20 to 55 (m=23.66., sd=5.92) and they were of different nationalities (73.7% German, 7.9% Dutch, 2.6% Spanish, 2.6% Hungarian, 2.6% Syrian, 2.6% Cypriot, 2.6% Ecuadorian, 2.6% Ukrainian and 2.6% Greek). Participants were collected by asking people in the library and by using the network SONA, where participants received 0.5 SONA credit points for their participation. Moreover, inclusion criteria were normal or corrected-to-normal vision and sufficient understanding of the English language. The data of one participant from the guilty condition needed to be excluded due to the lack of eye-tracking data/gaze points. Therefore, a total of 37 participants was used for the data analysis of the eye-tracking results. For the questionnaire, the original data set of 38 people was used.

Design

This study used a randomised, between-subject design with three conditions: guilty (n=12/13), innocent (n=12) and eyewitnesses (n=13). The guilty and witness conditions were manipulated in guilty knowledge by a mock crime scenario, whereas the innocent condition did not possess guilty knowledge. Participants in all conditions read a scenario and watched a first-person-perspective (FPP) video. The dependent variables were the total eye fixation duration at the first area of interest (AOI_1), the second area of interest (AOI_2) (see Figure 1) and the perceived stress during the interrogation. Because this study controlled for the interviewer effect, warmth, coldness and competence of the interviewer were taken into account as explorative dependent variables and were measured in a post-experiment questionnaire.

Figure 1: The picture shows the two Areas of Interest, AOI_1 being marked in blue and AOI_2 in red.

Materials

Materials used in this study were the wearable Tobii pro 2 eye-tracker (see Figure 2), a total of three scenarios, three FPP videos and three instructions for the interrogation, different ones for each condition (see Appendix B). Moreover, four pictures of different guns were used that were shown to the participants during the GKT. After the experiment, a questionnaire was used to assess the demographics of the participants, the perceived stress and the interviewer effect during the interrogation. For the questionnaire, the online survey tool 'Qualtrics' was used. For the data analysis, the programmes Tobii Pro Lab, Excel and IBM SPSS Statistics 25 (IBM Corp. Somers, NY) were utilised.

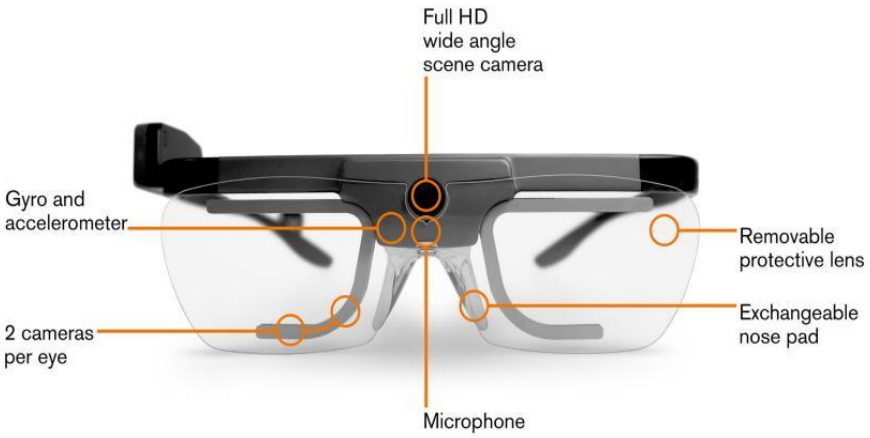

Figure 2. Wearable Tobii eye-tracker. If someone has poor eyesight, the lenses can be exchanged through corrective lenses. Source: https://www.tobiipro.com/de/produkte/tobii-pro-glasses-2/

Procedure

After the study was ethically approved by the Ethics Committee of the Faculty of Behavioural Sciences at the University of Twente (UT), equipment and a room in the library of the UT were reserved. Each session took about 10-15 minutes.

The experiment consisted of three parts in total. The first part was the manipulation of guilty knowledge, thus, reading the scenarios and watching the video of the mock crime. The second part was the GKT, i.e. the interrogation, and the third part was the questionnaire. This experiment was conducted in cooperation with another researcher.

First of all, the participants received information about the procedure of this study. It was explained to them that they needed to watch a small PowerPoint presentation with a scenario and a video. At the end of the presentation, they would receive instructions on how they needed to behave during the rest of the experiment. After participants signed the informed consent to agree to participation (see Appendix A), they were randomly assigned to one of the three conditions (guilty, witness, innocent). Following, the eye tracker was placed and calibrated in order to test if it was able to track the participant's eyes. They were instructed to leave on the eye-tracker until the researcher told them otherwise. Moreover, their eyes were not actually tracked during the first part

of the experiment, but it should just lead the participants to believe that they are already observed in their eye behaviour. Participants in each condition then watched the corresponding scenario and first-person-perspective (FPP) video (see Appendix B), which was timed so participants could solely concentrate on its content and the researchers were able to leave the room during this process.

Participants in the guilty condition received a story where they heard that they were responsible for a shooting in the Ravelijn building at the UT. They were told that they are now on their way to the Cubicus building, where they want to hide the gun. After that, a short video was shown to them, in which they are walking through a hallway and hiding the gun in a room. This gun had a blue dot on it and the camera focussed multiple times on the gun itself during the video.

Participants of the witness condition read a story in which they are heading to the Cubicus building in order to attend a tutorial. On their way, someone told them that a shooting has taken place in the Ravelijn. In the next hallway, they see someone carrying a gun, whom they are going to follow. In their video, they observed someone hiding a gun, but they could not see the blue dot or the gun in detail in the video. After this, they ran away, unnoticed by the offender.

For the third group, the participants of the innocent condition, the scenario was that they walk by a new study area in the Cubicus and decide to take a look inside. They watched a video where they walk through the same hallways as the other conditions, open the door to the study room (in which the gun was hidden in the two other videos) and take a look around. However, participants in this condition were not exposed to any crime related material in the video and, thus, did not possess guilty knowledge.

After the videos, participants in all three conditions received written instructions on how to behave in the following part of the experiment. These stated that when they were leaving the Cubicus, the police were waiting outside and want to interrogate everyone in the building since they received information that the suspected shooter headed there after the shooting. Moreover, they are all suspects and are requested to deny all responsibility for the shooting and claim that they know nothing about it. In the end, participants were asked to tell the researcher that they had finished reading the scenario and watching the video. The researcher was waiting outside with the co-researcher, who was then introduced as a police officer and asked the participant to take part in an interrogation. Sometimes the roles of researcher and interrogator were switched during the study based on the personal relationship with the participants (the person less familiar to the participant then took on the role of the interrogator).

The interrogation started with some basic questions, based on a study by den Besten (2017), that were used to make the situation seem more realistic (see Appendix C). Moreover, these questions were used with every participant and thus, the interrogation was standardised. After that, it was explained to the participants that they will now see four pictures on the computer screen for exactly ten seconds [These 10 seconds were based on the study by Derrick et al. (2011), which was mentioned earlier. Anyhow, the 12 seconds used by Derrick et al. (2011) were minimised by two seconds based on a pilot test with the researchers]. Participants were instructed to hold their head steady and look at the picture for the full time. When the screen would turn black, the interrogator will ask them a question on which they can answer with a yes or no ("Have you recently seen this object?").

Figure 3. The four guns presented in the GKT.

Figure 4. The weapon as the witness condition saw it (on top) and as the guilty condition saw it (bottom).

Following, their eyes were calibrated again, and the eye tracker recording was activated. Four pictures of different guns were shown to the participants (see Figure 3). Option d) was the weapon used in the incident and was presented as the third picture during the GKT. However, here the blue dot, that was originally on the gun in the FPP video of the guilty condition, was missing (see Figure 4). After the GKT, participants were told to take off the eye-tracker and were asked to fill in a questionnaire. They were now instructed to answer these questions honestly, regardless of the instructions they had received during the presentation. The

questionnaire measured their perceived level of stress during the interrogation and the interviewer's attitude.

To finish the experiment, a final debriefing was conducted to leave the participants with full information about the experiment. Within this debriefing, final questions by the participants were answered and it was directly stated that the incident used in the scenarios and videos was not real. Moreover, they were told in which condition they were, what was actually measured and what the hypotheses of this experiment were. Participants then had the opportunity to choose if they want to withdrawal from the study based on the newly received information. If that would have been the case, the participant's data would have been excluded from the sample.

Questionnaire

The dependent variables, apart from the eye fixations, are measured in a questionnaire. At first, the demographics were measured, namely, gender, age, nationality, education and occupation. After this, perceived stress was measured. For measuring this construct, seven items were used. The items were adapted from other studies (Verschuere, Meijer, & De Clercq, 2011; Sauerland et al., 2016; Smeets et al., 2011). Example items would be "When I think back to the interrogation, I feel stressed." or "During the experiment, I felt confident with my ability to handle my task successfully." (see Appendix D). These statements were then rated on a 5-point Likert scale, 1 indicating strongly disagree and 5 indicating strongly agree.

To determine if the questionnaire really just measured that one variable 'perceived stress', a factor analysis was conducted. The results showed that, actually, the questionnaire measured two factors. Two statements, that were formulated positively, showed one factor with an Eigenvalue of 1.66 (e.g. "During the interrogation, I felt I could cope with all the things going on") and the other five statements, which were formulated negatively (e.g. "When I think back to the interrogation, I feel stressed") built the second factor (Eigenvalue = 3.22). Moreover, the variable 'perceived stress' had a Cronbach's alpha of 0.77 and, thus, good reliability.

In the end, the interviewer effect was measured with items like "During the interview, I found the interviewer to be warm" or "During the interview, I found the interviewer to be intimidating" (see Appendix D). These items were rated on a scale from 0 to 100, with 0 indicating 'not at all' and 100 indicating 'strongly'. They were adapted from a study by Fiske, Xu, Cuddy and Glick (1999).

During a factor analysis, two factors were found with an Eigenvalue above 1 (5.50, 1.75). Following, three new variables were computed out of the different items that were measured. It was decided to use three variables since there was nearly a third factor found during the factor analysis (Eigenvalue of 0.995, if rounded it has an Eigenvalue of 1). Thus, warmth ("The interviewer was… warm, friendly, kind, likeable"), competence ("The interviewer was… confident, competent, intelligent") and coldness ("The interviewer was dominant, cold, intimidating") were the variables built from the items. The Cronbach's alpha of the interviewer effect, namely warmth (α=0.96), coldness (α=0.83) and competence (α=0.71) also showed excellent to adequate reliability.

Data analysis

Eye-tracking data. First, the recordings from the Tobii glasses were imported into the programme Tobii Pro Lab. There, also the picture of the weapon from the shooting was inserted and the two areas of interest were marked (see Figure 1). Following, the attention filter was used to map the eye fixations of the participants onto the picture with the AOI's. During this process, it became clear that in some of the cases there seemed to have been an error during the data collection. When one would have expected people to have looked at the interrogator, since that person was asking them a question, the gaze point actually appeared a few centimetres below the face of the interrogator. After careful consideration, it was decided that this must have been a calibration error and if left unfixed, this could have consequences for the results. Therefore, it was decided to map the fixations of the participants where this error derived manually. This manual mapping was conducted with data of 15 of the 37 participants.

Moreover, heatmaps for every condition have been created. The heatmaps show a visualization of which points on the picture were fixated longest by the participants (a red dot indicated long and multiple fixations and thus, high attention, whereas a green or yellow dot indicated less attention on that point) and, thus, may have caught their attention. After that, the fixation duration data were exported in an Excel file and then into an SPSS file to analyse the data together with the data from the questionnaire.

To analyse the fixation duration data of the GKT, first, a normality test is conducted, to see whether there is a normal distribution. This is done independently for both AOI's. The data of both AOI's was not normally distributed, hence, a Kruskal-Wallis test is conducted to determine if the

three conditions differed from each other in their eye-fixation in the AOI's. In the end, using a Bonferroni test, it is calculated between which groups a significant difference exists.

Questionnaire data. In terms of the questionnaire, for the analysis of the variable 'perceived stress', two items needed to be coded reverse ("During the interrogation, I felt I could cope with all the things going on" and "During the experiment, I felt confident with my ability to handle my instruction successfully"). These were the ones that were formulated more positively. Then, a one-way ANOVA will be executed for the variable 'perceived stress' and also for the three variables of the interviewer effect. This is to determine if there was a significant difference between the three conditions and, again, if that is the case, a Bonferroni test will be conducted to determine which groups differ from each other.

Results

Analysis of the GKT

First of all, the mean scores of the two variables AOI_1 (related to hypothesis 1) and AOI_2 (related to hypothesis 2) were calculated (see Table 1). Here, it can be seen that participants of the guilty condition fixated AOI_1 the longest, followed by the innocent condition and then the witnesses. Regarding AOI_2, the fixation duration was the longest for the innocent condition, followed by the guilty condition and then the witness condition.

Table 1

Means and Standard Deviations for the Variables 'duration_AOI_1' and 'duration_AOI_2' for the three conditions.

Eye-fixation duration in...	Condition	mean	sd
	Guilty	0.50	0.53
...AOI_1	Eyewitness	0.02	0.09
	Innocent	0.11	0.21
	Guilty	9.08	1.61
...AOI_2	Eyewitness	8.74	2.02
	Innocent	9.15	1.11

These differences, especially regarding AOI_1, are supported by the heatmaps. As can be seen in Figure 5, in the guilty condition, the gaze points hit the point at the gun, whereas in the other condition, this is not the case. Thus, the area is red in the guilty condition, which indicates multiple and longer fixations, whereas it is green/yellow in the other two conditions. However, a further analysis had to be conducted to see whether the differences that manifested themselves in the means and visualizations were also significant differences.

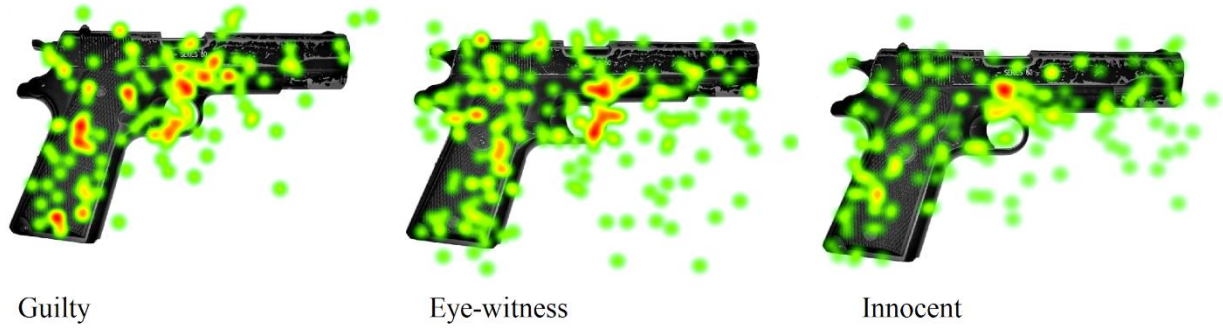

Figure 5. The three heatmaps of the three conditions.

To test whether **hypothesis one** *"Eye fixation duration is longer in the AOI_1 for the guilty condition in comparison to the witness and innocent group"* can be accepted, first a Shapiro-Wilk test was conducted. Results showed that the data for AOI_1 was not normally distributed ($p<.01$). Due to that, it was necessary to conduct a Kruskal-Wallis test, determining the differences between the three conditions. The results of the Kruskal-Wallis showed a significant difference between the three conditions in terms of AOI_1 [$X^2(2)=8.75$, $p=.01$] with a mean rank of 24.83 for the guilty condition, 14.65 for the witness condition and 17.88 for the innocent condition. To detect between which conditions the difference existed, a Bonferroni was conducted. The post-hoc test showed that there was a significant difference between the witness and the guilty condition ($p=.01$). However, there was no significant difference found between the guilty and innocent ($p=.15$) and the witness and innocent condition ($p=1$)[1]. Hence, the first hypothesis was not confirmed.

[1]When analysing the data with a one-way ANOVA and Bonferroni, which is normally only used for normally distributed data, another significant difference was found, specifically, between the guilty and innocent group ($p=.02$). The p-values for the other comparisons stay the same as with the Kruskal-Wallis and Bonferroni. This would have confirmed hypothesis one.

When testing the **second hypothesis** *"Eye fixations on the AOI_2 are more spread in the innocent condition compared to the witness and guilty condition"*, the results of the Shapiro-Wilk test implied that the eye-fixation for AOI_2 was also not normally distributed (p<.01). Based on that, a Kruskal-Wallis was executed to test for differences between the conditions. The results indicated that there is no significant difference between the groups [X^2(2)=.94, p=.95]. This contradicts the second hypothesis and therefore, it could not be confirmed.

In addition, the fixation count in AOI_2 has been taken into account.[2]

Analysis of the questionnaire

First of all, the means of the perceived stress regarding the three conditions were calculated (see Table 3). Here, it can be seen that the guilty condition scored highest in perceived stress (m=3.01) compared to the witness (m=2.79) and the innocent condition (m=2.20), which is in line with the third hypothesis.

Table 3

Means and Standard Deviations for the Variables 'stress' for the three conditions.

Variable	Condition	mean	sd
Stress	Guilty	3.01	0.43
	Eyewitness	2.79	0.78
	Innocent	2.20	0.77

However, to test whether these results also show a statistical significance and to see whether **hypothesis three** *"Participants of the guilty and eye-witness condition score higher on perceived stress than the innocent condition. This effect is expected to be stronger for the guilty condition than for the witness condition."* can be proven correct, a Shapiro-Wilk test for normality was performed. The variable was found to be non-normally distributed (p<.01), which is why then a Kruskal-Wallis needed to be conducted. The results of this test showed that there is a significant

[2] Fixation count for AOI_2: witness (m=17, sd=6.76), innocent (m=16.50, sd=6.56), guilty (m=12.42, sd=8.48)

difference between the three conditions [$X^2(2)=8.58$, $p=.01$] with a mean rank of 24.12 for the guilty condition, 21.92 for the eyewitnesses and 11.88 for the innocent condition. Using a pairwise comparison, the only significant difference found was between the guilty and the innocent condition ($p=.02$). No significant difference was found between the guilty and the witness group ($p=1$). The same applies to the comparison of the witness and innocent group ($p=.07$). Even though there was a significant difference found for the guilty condition, there was no evidence that the witness condition differed from the innocent condition. Thus, the third hypothesis was supported by the means but, still, could not be confirmed.

Regarding the interviewer effect, a Shapiro-Wilks was again used to test for normality (warmth $p=.24$, competence $p=.11$, coldness $p=.11$). All variables were found to be normally distributed and, thus, three separate ANOVAS were conducted. The results showed no significant differences between the conditions for warmth [$F(2, 35)=3.20$, $p=.053$] and coldness [$F(2, 35)=1.45$, $p=.25$]. However, in the variable competence, a significant difference has been found [$F(2, 35)=3.61$, $p=.04$] and, therefore, a Bonferroni was conducted as a post-hoc test. The perceived competence of the interviewer only differed significantly between the witness and the innocent condition ($p=.03$) but not between the witness and guilty condition ($p=.72$) or innocent and guilty condition ($p=.42$). Moreover, even though, the main effect in the ANOVA was not significant, by using a Bonferroni post-hoc test, another significant difference was found. This was also between the witness and innocent condition, but in the variable warmth ($p=.048$).

Discussion

Discussion of results

The aim of the current study was to see if eye-tracking can be used in the future as a more objective and reliable instrument to determine if a suspect is guilty or not. This was done by combining eye-tracking with a GKT. Moreover, this experiment concentrated on how eye fixation behaviour might differ between innocent and guilty people in interrogation situations and also how witnesses can be differentiated from them.

It was expected that the eye fixation duration is longer in the AOI_1 for the guilty condition in comparison to the witness and innocent group. Even though, when looking at the means and the heatmaps there was a difference visible, this difference was not big enough to be statistically significant, except between the guilty and innocent condition. When looking at the means, this

indicates that participants from the guilty condition fixated the blue dot much longer than people from the innocent condition did, similar to what was the case in the study of Derrick and others (2011). Moreover, this implies that participants recognised, consciously or unconsciously, that the blue dot was missing during the interrogation situation. The heatmaps also clearly support this.

Furthermore, as mentioned, a one-way ANOVA was conducted as well, next to the Kruskal-Wallis test. When also taking into account these results, which showed another difference between the guilty and the witness condition, this also supports what was originally expected: that people of the guilty condition fixate AOI_1 much longer than participants of the other conditions. Yet, an ANOVA is usually only used with normally distributed data, which, here, was not the case. Normally, it is said that constructing a one-way ANOVA with non-normally distributed data makes the test invalid. However, a study by Blanca, Alarcón, Arnau, Bono, and Bendayan (2017) manipulated data in terms of its homogeneity of variance and normality and then tested if the ANOVA would still be a valid option. They found that this is indeed the case since the Type I error rate (i.e. rejecting the Null Hypothesis even though it is true) was still acceptable. Thus, it was considered important to also take into account the results of the ANOVA, especially since they approve the expectations.

When comparing the results of the current study with that of other studies, it can be said that many similarities were found. For instance, in the study of Bradley and Warfield (1984), it was found that innocent people with guilty knowledge (in the current study referred to as witness condition) were not detected as guilty based on the results of the polygraph and GKT. When comparing this to the current results, it can be seen that the conditions scored in the same order as in Bradley and Warfield's (1984) study (i.e. the guilty group fixated the longest, followed by the witnesses and then the innocent). However, to see whether only guilty people were detected as guilty, it would have been necessary to determine cut-off scores for fixation duration. Apart from this, the results of the current study are also comparable to other results of eye-tracking studies (Millen et al., 2017; Peth et al., 2013; Schwedes & Wentura, 2016). Namely, that objects that were familiar to the person were fixated longer than items that were not familiar. However, to strengthen this argument, the results of the fixation duration in AOI_2 should also be considered.

Here, other than expected, there was no evidence of a difference between the three conditions. This also becomes apparent when looking at the heatmaps, where, in all three conditions, the fixation points were widely spread over the whole picture and the points were rather

green than yellow or red, which implies that nothing really caught the participants attention in the AOI_2. That the means were quite close to each other also supports this. Hence, this contradicts the argument that objects that are familiar to a person are fixated for longer than unfamiliar objects since, in that case, witnesses and guilty people should have differed from innocent people to whom the gun was unfamiliar. However, this could be explained by the fact that some participants stated they would know the gun, even though they were in the innocent condition. They stated they would know all of the four guns from different ego-shooter games. This may not be applicable to the specific gun but still, to the kind of guns and this may have had an effect on the eye fixation behaviour. Since they were, then, already familiar with the gun, this gives them the same premise as the witness group. They would be, still, innocent but familiar with the gun, which would lead to different expectations, namely, that they have longer fixation durations and less fixation count on the gun. This problem is comparable to the one in the study of Millen and colleagues (2017), where they could not be sure that participants would know the celebrities that were chosen as neutral stimuli.

Regarding the fixation count, it was stated earlier that there were fewer fixations but rather longer ones in exchange on familiar objects (Millen et al., 2017). When looking at the results of fixation count, it can be seen that the overall gun is fixated most often by the witness condition, shortly followed by participants of the innocent condition and then, with some distance follows the guilty condition. This supports the findings of Millen and colleagues (2017) and also Peth and others (2013), that guilty people might not avoid looking at these crime-related objects but rather feel no need to explore the picture since they already encountered and explored it before. Therefore, longer fixation duration and less fixation count might be an indicator for the possession of guilty knowledge.

Furthermore, another aim of the study was to see whether the perceived stress might be related to the possession of guilty knowledge, and thus, another indicator which could be useful, as a support of the eye-tracking, in police work. As expected, the results of the questionnaire show that the guilty condition did, indeed, feel more stressed by the interrogation than people of the innocent condition. Additionally, when putting the mean scores in the context of the 5-point Likert scale, participants of the guilty condition neither agreed, nor disagreed with the statements about stress. In contrast to that, participants of the innocent condition rather disagreed with the items. The mean of the witness condition lies somewhere in between these two options. This supports that

being guilty has an influence on the perception of stress when being under interrogation by the police and that participants of the guilty and witness condition feel more stressed than the ones from the innocent condition. Even though no difference was found for the witness group, the means still point in the right direction and show that the guilty condition perceived the most stress, followed by the witnesses and then the innocent, as it was assumed.

In line with the expectations, when looking at the variables 'warmth', 'coldness' and 'competence', it became clear that none of these had an effect on the experiment since the interviewer seemed to have been perceived by all participants in roughly the same way. The only exception was that the witness and innocent condition perceived the interviewer as differently competent. However, this could have also been a coincidence and due to the change of the interrogator based on the personal relationship with the participant.

Strong points and limitations of the study

Strong points of this study were that participants almost always stated that they liked the study and that it was nice to participate in. Hence, participants were probably highly motivated to fulfil their tasks. Another good point of the study was the variety of the sample. Because random people were asked to participate, students took part in it that were not only Psychology or Communication Science students, as that is usually the case. Moreover, some participants were students but also some of them were working, thus they had different backgrounds. The sample also consisted of a variety of nationalities.

Furthermore, Derrick and colleagues (2011) stated in their recommendations that the guilty knowledge effect should be examined with pictures that were related but not the same in further studies, to really determine the effect. This was done in the current study and can also be considered as a positive aspect.

Nevertheless, this study also had some limitations. For instance, it is really hard to simulate such a topic as committing a crime under laboratory settings. Participants would probably feel more involved if it was a real-life situation, especially in terms of their emotional responses. Studies also state that this plays a role in the motivational aspect of participants (MacLaren, 2001). To make it more realistic one could, for instance, hire actors and really play the scenarios that were watched during the experiment. Moreover, it should be considered to only use people that are unfamiliar with the participants. This way, it might be more realistic. However, to do this, many more

resources would have been needed than those that were available (e.g. actors, costumes, time, rooms *et cetera*).

Moreover, with the aforementioned diversity in the nationality of the participants also comes a problem, namely that English was not their native language. Also, for the people that were working, it could be that they are not that fluent in English due to the non-daily usage. Thus, it cannot be said with certainty if they have understood the scenarios, instructions and questions. This might have affected the results of the questionnaire in terms of its validity and reliability. If one would ask these questions to native speakers, there might be different results (reliability) and the questionnaire might also have measured something different than intended (validity) due to misunderstood questions. Apart from that, even for participants with sufficient English skills, it seemed to be hard to understand the instructions, especially for participants in the witness condition. This became clear when they agreed to the question if they had seen the weapon before even though they were instructed to claim they know nothing about the shooting. To detect hidden guilty knowledge was one of the main aspects of this study and by agreeing to that question, they did not possess hidden guilty knowledge anymore. Additionally, this might have had an effect on the questionnaire results since when they admitted they had seen the gun, they might not have felt very stressed during the interrogation since they did not hide anything. Moreover, this question was also difficult with participants that were playing a lot of shooting games at home, as already mentioned. After this problem became evident, it was, once more, repeated that the interrogation only refers to the shooting which has taken place and if they had seen the guns in connection to that.

Furthermore, due to the shortage of rooms, it was not possible to have a second room for the interrogation. This might have had an influence on the results since the police interrogator might have been seen by the participants beforehand. Additionally, they have watched the video of the mock crime in the same environment as where the interrogation took place. Hence, there was no physical distance to the committed crime which might have affected their emotions and also the simulation of a real-world scenario might have suffered from it.

In addition, the experiment was conducted with a laptop, whereas it would have been better to use a computer. Due to using the laptop, the view through the eye-tracker was not filled with the picture of the gun that was used. By using a bigger computer screen, which would have also stood

a little higher, it would have been easier to analyse the data and also, there would have probably been more data to analyse.

Suggestions for future studies

Many of the issues with this study could have been avoided if there had been pilot testing. Thus, it is advisable to first conduct a pilot study in order to eliminate misunderstandings early.

In terms of the questionnaire, many participants were confused by the question if there was a change in the pictures. This was just applicable to the guilty condition but also participants in that condition were confused by it since they did not recognise the blue dot. Therefore, a lot of answers were given that applied to the four guns and differences between these (e.g. that the murder weapon looked more used than the other three) instead of the difference between the gun in the video and how it looked during the interrogation. Due to that, it should be made clear in future studies that the difference between the weapon in the video and the one in the interrogation is meant with this question. This could be done by adding a simple question: Did you see a gun in the video? And if people answer yes to that question, a follow-up question appears on the screen: Was the picture in the interrogation different from what you saw in the video? What was different?

Moreover, it might be of interest to see in which way the eye fixation in the weapon of the incident differs from the other three images. For this, it might be handy to also look at the heatmaps and fixation counts of the other pictures. This could also be helpful regarding the problem with the gun familiarity of gamers. After analysing the other pictures, it can be differentiated if all of the objects looked familiar or really just the object under investigation. Moreover, this might also be handy when determining the influence of stress on eye movements. It could be that participants might try to control their eye movements during the interrogation. This could also be measured in a post-questionnaire.

Future studies should, if they want to study new methods for the detection of guilty knowledge, always include a witness condition. This may be really helpful for future detective work and also to avoid innocent people being punished. In the current study, there was no difference found for the witness condition, however, maybe in future studies, a cut-off score for guilty/not guilty could be determined regarding the fixation duration.

Therefore, further research is needed to determine if eyewitnesses may be distinguishable from guilty and innocent people and to support that eye-tracking may be a new, reliable form of measurement for detecting guilty knowledge.

Conclusion

Overall, the findings of this study support that guilty people differ in their eye-behaviour from innocent people. However, it was not possible to distinguish eyewitnesses in this process. Moreover, being guilty or possessing guilty knowledge seems to increase the perceived stress during an interrogation situation. Therefore, it can be concluded that eye-tracking may be used as a method to distinguish someone who committed a crime from someone who did not. Nevertheless, it was, in this case, not able to integrate people who are in the possession of guilty knowledge but are not guilty of the crime itself. Hence, future research should concentrate on how such people can be taken into account during the process of finding the truth. Being able to use such an advanced technique as eye-tracking to solve crimes might make life easier for many people.

References

Barnes, N. R. (2008). The polygraph and juveniles: Rehabilitation or overreaction? A case against the current use of polygraph examinations on juvenile offenders. *The University of Toledo Law Review, 39*(3), 669-692. Retrieved on March 22, 2019, from https://heinonline.org/HOL/LandingPage?handle=hein.journals/utol39&div=42&id=&page=

Ben-Shakhar, G., & Elaad, E. (2002). The Guilty Knowledge Test (GKT) as an application of psychophysiology: Future prospects and obstacles. In M. Kleiner (Ed.), *Handbook of Polygraph Testing* (pp. 87-102). New York, NY: Academic Press.

Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552-557. doi:10.7334/psicothema2016.383.

Bradley, M. T., & Warfield, J. F. (1984). Innocence, information, and the guilty knowledge test In the detection of deception. *Psychophysiology, 21*(6), 683–689. doi:10.1111/j.1469-8986.1984.tb00257.x

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and human behavior*, *28*(6), 687-706. doi:10.1007/s10979-004-0565-x

Den Besten, A. (2017). *Truth (or) lies within the eyes - Examining the effects of guilty knowledge with eye-tracking* (Unpublished master's thesis). University of Twente, Enschede, Netherlands.

Derrick, D. C., Moffitt, K., & Nunamaker Jr, J. F. (2011). Eye gaze behavior as a guilty knowledge test: Initial exploration for use in automated, kiosk-based screening. In *Hawaii International Conference on System Sciences*, 4-7. Retrieved on March 22, 2019, from https://www.erim.eur.nl/fileadmin/centre_content/future_energy_business/images/informs_award/2010/HICSS-eye-tracker-v4.pdf

Duchowski, A. T. (2007). *Eye tracking methodology. Theory and practice (2nd Ed.)*. London, England: Springer Science + Business Media

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473-489. doi:10.1111/0022-4537.00128

Groer, M., Murphy, R., Bunnell, W., Salomon, K., Van Eepoel, J., Rankin, B., ... & Bykowski, C. (2010). Salivary measures of stress and immunity in police officers engaged in simulated critical incident scenarios. *Journal of occupational and environmental medicine, 52*(6), 595-602. doi:10.1097/JOM.0b013e3181e129da

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly, 7*(4), 429-456. doi:10.1177/1098611104264748

Horvath, F. (1977). The effect of selected variables on interpretation of polygraph records. *Journal of Applied Psychology, 62*(2), 127-136. doi:10.1037/0021-9010.62.2.127

IBM Corp. Released in 2016. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Lombroso, C. (1911). *Criminal man*. New York, NY, London, UK: G. P. Putnam. The Knickerbocker Press.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*(6), 385-388. doi:10.1037/h0046060

MacLaren, V. (2001). A quantitative review of the Guilty Knowledge test. *Journal of Applied Psychology, 86*(4), 674-683. doi:10.1037/0021-9010.86.4.674

Meijer, E. H., Smulders, F. T., Johnston, J. E., & Merckelbach, H. L. (2007). Combining skin conductance and forced choice in the detection of concealed information. *Psychophysiology, 44*(5), 814-822. doi:10.1111/j.1469-8986.2007.00543.x#

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: evidence for unconscious alteration of judgments. *Journal of personality and social psychology, 35*(4), 250-256. doi:10.1037/0022-3514.35.4.250

Oatley, G., Ewart, B., & Zeleznikow, J. (2006). Decision support systems for police: Lessons from the application of data mining techniques to "soft" forensic evidence. *Artificial Intelligence and Law, 14*(1-2), 35-100. doi:10.1007/s10506-006-9023-z

Peth, J., Kim, J. S., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology, 88*(1), 96-103. doi:10.1016/j.ijpsycho.2013.03.003

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of consumer research, 21*(2), 274-290. doi:10.1086/209397

Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, *3*(3), 267-276. doi:10.3758/BF03212910

Sauerland, M., Raymaekers, L. H., Otgaar, H., Memon, A., Waltjen, T. T., Nivo, M., ... & Smeets, T. (2016). Stress, stress-induced cortisol responses, and eyewitness identification performance. *Behavioral sciences & the law*, *34*(4), 580-594. doi:10.1002/bsl.2249

Sheldon, W. H. (1942). *The varieties of temperament; a psychology of constitutional differences.* New York, NY: Harper & Row.

Smeets, T., Cornelisse, S., Quaedflieg, C. W., Meyer, T., Jelicic, M., & Merckelbach, H. (2012). Introducing the Maastricht Acute Stress Test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses. *Psychoneuroendocrinology*, *37*(12), 1998-2008. doi:10.1016/j.psyneuen.2012.04.012

Spain, M., & Perona, P. (2008). Some objects are more equal than others: Measuring and predicting importance. *Computer Vision–ECCV 2008*, 523-536. Heidelberg, Germany: Springer

Tyson, G. A. (1983). Stress and detention. *South African Medical Journal*, *64*(22), 858-859. Retrieved on April 26, 2019, from https://www.researchgate.net/profile/Graham_Tyson/publication/16560210_Stress_and_detention/links/02e7e52cdea7b5611b000000.pdf

Van der Lans, R., Wedel, M., & Pieters, R. (2011). Defining eye-fixation sequences across individuals and tasks: The Binocular-Individual Threshold (BIT) algorithm. *Behavior Research Methods*, *43*(1), 239-257. doi:10.3758/s13428-010-0031-2

Verschuere, B., Crombez, G., & Koster, E. (2004). Orienting to guilty knowledge. *Cognition and Emotion*, *18*(2), 265-279. doi:10.1080/02699930341000095

Verschuere, B., Meijer, E., & De Clercq, A. (2011). Concealed information under stress: A test of the orienting theory in real-life police interrogations. *Legal and criminological psychology*, *16*(2), 348-356. doi:10.1348/135532510X521755

Appendix

Appendix A. Informed Consent

Informed Consent

‘I hereby declare that I have been informed in a manner which is clear to me about the nature and method of the research. My questions have been answered to my satisfaction. I agree with my own free will to participate in this research. I reserve the right to withdraw this consent without the need to give any reason and I am aware that I may withdraw from the experiment at any time. I am also aware of the fact that my personal data will be made completely anonymous. If I request further information about the research, now or in the future, I may contact Rebecca Amui or Clara Büttner (r.l.o.amui@student.utwente.nl; c.buttner@student.utwente.nl).’ If you have any complaints about this research, please direct them to the secretary of the Ethics Committee of the Faculty of Behavioural Sciences at the University of Twente, Drs. L. Kamphuis-Blikman P.O. Box 217, 7500 AE Enschede (NL), telephone: +31 (0)53 489 3399; email: l.j.m.blikman@utwente.nl

Signed in duplicate:

…… ……

Name subject Signature

I have provided explanatory notes about the research. I declare myself willing to answer to the best of my ability any questions which may still arise about the research.’

…… ……

Name researcher Signature

Appendix B. Scenarios and instructions

Scenario for the guilty condition:

You are at the University of Twente and you are responsible for a shooting that took place minutes ago at the Ravelijn building on the University Campus. You fled the crime scene and are currently in the Cubicus building trying to find a spot where you can hide your gun.

Instructions after the video and before the GKT/interrogation:

You were able to hide the gun but, on your way, out you run into the police. They received information that the suspected shooter ran into the Cubicus building. In order to catch the person responsible for the shooting, they want to question everybody in the building and each of you remains a suspect until proven innocent. Therefore, they take you into a room for interrogation. Of course, it is in your best interest to not say anything about the shooting and the gun. So, your intention is to **not let them show you know something about that** and just behave naturally. Is this clear?

Adapted from Bradley & Warfield (1984)

Scenario for eye-witness condition:

You are at the University of Twente and you are heading to your next tutorial in the Cubicus building on the University Campus. When you are walking down the hallway, you are being told that a shooting has taken place minutes ago at the Ravelijn building. In the next hallway, you unexpectedly spot someone carrying a gun. In a split second, you decide to follow the suspected shooter from a close distance.

Instructions after the video and before the GKT/interrogation:

You were able to run away without the shooter noticing you and, on your way out, you run into the police. They received information that the suspected shooter ran into the Cubicus building. In order to catch the person responsible for the shooting, they want to question everybody in the building and each of you remains a suspect until proven innocent. Therefore, they take you into a room for interrogation. Of course, it is in your best interest to let them know you have nothing to do with this. So, your intention is to let them know you **don’t know anything about that** and just behave naturally. Is this clear?

Adapted from Bradley & Warfield (1984)

Scenario for innocent condition:

You are at the University of Twente and you are heading to your next tutorial in the Cubicus building on the University Campus. When you are walking down the hallway, you come across the new study areas. You stop and decide to take a look insight.

Instructions after the video and before the GKT/interrogation:

When you are walking out of the study area, you are being told that a shooting has taken place minutes ago at the Ravelijn building and that police is outside the Cubicus building. They received information that the suspected shooter ran into the Cubicus building. In order to catch the person responsible for the shooting, they want to question everybody in the building and each of you remains a suspect until proven innocent. Therefore, they take you into a room for interrogation. Of course, it is in your best interest to let them know you have nothing to do with this. So, your intention is to let them know you don't know anything about that and just behave naturally. Is this clear?

Adapted from Bradley & Warfield (1984)

Appendix C. Interrogation Questions

1. Please state your full name.
2. Please state your date of birth.
3. What is your country and place of birth?
4. How long have you been living in the Netherlands?
5. When was the last time you have travelled abroad?
6. Where did you go to?
7. What did you do last night as of 8 p.m.?

Appendix D. Questionnaire

Start of Block: Default Question Block

Q1.1 Participant number (filled in by researcher)

_____

Q1.2 What is your age?

_____

Q1.3 What is your gender?

O Male

O Female

O Other

Q1.4 What is your nationality?

O Dutch

O German

O Other, namely: _____

Q1.5 What is your highest level of education?

- ☐ MAVO
- ☐ HAVO
- ☐ VWO
- ☐ German Abitur
- ☐ Bachelor
- ☐ Master
- ☐ PhD
- ☐ Other, namely: _____

Q1.6 What is your occupation?

O Working

O Student

O Other

End of Block: Default Question Block

Start of Block: Block 1

Q2.1 The following questions will focus on the **first part** of the experiment: **the video** *(first-person perspective)*

Q2.2 Clearly describe what you did/saw in the video:

_____

_____

_____

_____

_____

Q2.3 Please indicate to what extent you agree/disagree with the following statements

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
When I think back to the video, I feel guilty	O	O	O	O	O
When I think back to the video, I feel tense	O	O	O	O	O
When I think back to the video, I feel regret	O	O	O	O	O
When I think back to the video, I feel bad	O	O	O	O	O
When I think back to the video, I want to apologize	O	O	O	O	O

End of Block: Block 1

Start of Block: Block 2

Q3.1 The following questions will focus on the **second part** of the experiment: **the interrogation**

Q3.2 Please indicate to what extent you agree/disagree with the following statements

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
When I think back to the interrogation, I feel stressed	☐	☐	☐	☐	☐
When I think back to the interrogation, I feel nervous	☐	☐	☐	☐	☐
When I think back to the interrogation, I feel aroused/tensed	☐	☐	☐	☐	☐
When I think back to the interrogation, I feel pressured	☐	☐	☐	☐	☐

During the interrogation, I felt an increased heartbeat	☐	☐	☐	☐	☐
During the interrogation, I felt I could cope with all the things going on	☐	☐	☐	☐	☐
During the experiment, I felt confident with my ability to handle my instruction successfully	☐	☐	☐	☐	☐

End of Block: Block 2

Start of Block: Block 3

Q4.1 During the interrogation, I found the interviewer to be...

0	10	20	30	40	50	60	70	80	90	100

Warm	
Friendly	
Confident	
Dominant	
Cold	
Competent	
Intimidating	
Kind	
Likeable	
Intelligent	

End of Block: Block 3

Start of Block: Block 4

Q5.1 During the interrogation, four images were presented:

Image 1: Kalashnikov (long, brown)
Image 2: Hunting Rifle (long, black)
Image 3: Handgun (black)
Image 4: Revolver (silver)

Q5.2 Did you notice some sort of change in one of these pictures?

O Yes

O No

End of Block: Block 4

Start of Block: Block 5

Display This Question:

If Did you notice some sort of change in one of these pictures? = Yes

Q6.1 Reminder on four images:

Image 1: Kalashnikov (long, brown)

Image 2: Hunting Rifle (long, black)

Image 3: Handgun (black)

Image 4: Revolver (silver)

Display This Question:

If Did you notice some sort of change in one of these pictures? = Yes

Q6.2 In which image did you notice a change?

O Image 1

O Image 2

O Image 3

O Image 4

Display This Question:

If Did you notice some sort of change in one of these pictures? = Yes

Q6.3 What exactly caught your attention in this picture?

_____

_____

_____

_____

_____