Did You Commit a Crime There? Investigating the Visual Exploration Patterns of Guilty, Innocent, Honest, and Dishonest Subjects When Viewing a Complex Mock Crime Scene

Marco Colasanti[1] · Giulia Melis[2] · Merylin Monaro[2] · Eleonora Ricci[3] · Francesca Bosco[4] · Michela Rossi[4] · Silvia Biondi[4] · Maria Cristina Verrocchio[1] · Alberto Di Domenico[1] · Cristina Mazza[3] · Paolo Roma[4]

Abstract

Previous studies with the eye-tracking technology have predominantly tracked eye parameters in response to a single simple stimulus, and have generated interesting - sometimes inconsistent - results in research on deceptive behavior. The present study analyzed visual patterns in response to a complex image, to investigate potential differences in eye fixation between guilty versus innocent, and honest versus dishonest participants.

One hundred and sixty participants were assigned to one of four experimental groups, defined by the parameters of honesty (dishonesty) and guilt (innocence), and asked to complete a computer-based task, looking at neutral and target images (i.e., images of the mock crime scene). RealEye software was used to capture participants' eye movements when viewing the images.

The findings revealed significant differences in eye movements between the four experimental groups in the pictures in which the area where the crime took place was clearly visible. Dishonest and guilty participants recorded fewer and shorter fixations in the area of the image where the crime took place than those who entered the crime scene but did not commit the crime. No differences between groups emerged in the visual patterns in response to neutral images, confirming that the number and duration of fixations in response to the target images may be attributed to the experimental condition.

Keywords Deception detection · Crime scene · Eye fixation · RealEye

Introduction

Over the past decades, research on verbal and non-verbal indicators of deception has increased considerably. The Concealed Information Test (CIT; formerly known as the Guilty Knowledge Test) is the most widely used scientifically validated interviewing test and psy-

chophysiological method in the field of criminal investigation to unveil concealed knowledge and deception during a recognition task (e.g., for reviews, see Ben-Shakhar and Elaad, 2003; Meijer et al., 2014; Suchotzki et al., 2017; Verschuere et al., 2011). Specifically, the CIT presents images of both crime-relevant information (i.e., target or key items, such as the murder weapon) and non-relevant information (i.e., non-target or non-key items), and measures interviewees' psychophysiological responses to the visuals. Consistently higher activation (e.g., higher heart rate or skin conductance) to the target item in comparison to the non-target items is considered to be predictive of deception (Krapohl et al., 2009).

Recently, another behavioral technique known as eye-tracking has attracted significant attention and produced interesting results in research on the detection of deceptive behavior. Eye-tracking is a well-known experimental method that was designed to measure and register visual search behavior, eye movement, and gaze location across time and tasks (for a review, see Singh & Singh, 2012). Prior research has outlined that eye movements can unveil cognitive processes (Zagermann et al., 2016) and serve as indirect measures of attention (Lee & Ahn, 2012; Pittarello et al., 2016; Tsai et al., 2012), memory and familiarity (e.g., Hannula et al., 2010), emotion (Lim et al., 2020; Perkhofer & Lehner, 2019), and deception (Cook et al., 2012; Proudfoot et al., 2013, 2016; Twyman et al., 2014). Several studies were carried out to analyze the ocular patterns and eye-movements associated with deception: the rationale for its use is that lying is a costly cognitive process and implies an increase in cognitive load, which may alter behavioral parameters, including eye movement and exploration (Monaro et al., 2017, 2018). For instance, research has shown that, when lying, pupils dilate (Proudfoot et al., 2016), saccade velocity increases (Vrij et al., 2015), blink rate and duration decreases (Leal & Vrij, 2010; Marchack, 2013), and fixation behavior increases (Zagermann et al., 2016). Despite the abundance of studies on this topic, they often employ research designs that are hardly comparable, both from a methodological and statistical point of view. Regardless, most of these studies suggest that variables related to eye fixations might be useful in the detection of concealed knowledge and deception.

Regarding concealed knowledge, for instance, Nahari et al. (2019), both through parallel and sequential displays, presented pictures of familiar and unfamiliar faces in two studies: in the first one, they instructed participants to naively conceal their familiarity; in the second, they instructed participants to conceal their familiarity by equally distributing their gaze between all faces. Interestingly, only when tackling the experimental task following their own naive strategies, participants directed their gaze more readily at familiar faces, as evidenced by an increase in the number and average duration of fixations. This result is in line with other studies (e.g., Ryan et al., 2007), which found that, when presented sequentially with both novel and known faces, participants tended to look at familiar faces for more time. Millen et al. (2020) examined eye fixations patterns of participants during honest and concealed recognition of both familiar, newly learned and unfamiliar faces and of familiar, newly learned and irrelevant scenes and objects, presented sequentially. No differences were found between honest and concealed recognition of familiar faces, characterized by fewer fixations and fewer looks compared to unfamiliar faces. Likewise, the recognition of familiar scenes and objects was characterized only by less fixations, in both the recognition conditions, compared to irrelevant scenes and objects. Moreover, in the concealed recognition, participants showed less fixations on newly learned faces, scenes and objects compared to unfamiliar faces and irrelevant scenes and objects. Derrick et al. (2011) asked participants to construct a mock improvised explosive device (IED) and then—when questioned during an

automated screening interview, during which an altered image of the IED was shown to all participants in both the experimental and the control group—instructed them to conceal the fact that they had done so. The findings showed that the participants who had constructed the IED had longer fixations on the altered portion of the image than participants in the control group.

Regarding deception detection, for instance, Pittarello et al. (2016) conducted a within-subjects study in which participants were asked to report whether or not one of two cards presented simultaneously on a screen was a joker. Participants could honestly report seeing the joker and lose money or falsely report not seeing it and keep the money. The results showed that, compared to when they were honest, dishonest participants obtained shorter and fewer fixations on the joker. In another study (van Hooft & Born, 2012), participants were asked to complete—honestly or deceptively—a personality inventory (i.e., the Five Factor Personality Inventory) and an integrity test in the context of personnel selection. The analysis of eye movement revealed that participants in the deceptive condition demonstrated, on average, almost one eye fixation less per item compared to participants in the honest condition. Cook et al. (2012) examined the eye movements of two experimental groups when responding to a questionnaire comprising items related to a mock-crime that half of the participants committed (i.e., guilty group) and half did not but was aware that it had been committed (i.e., innocent group). Findings showed that, compared to innocent participants, the guilty group reported shorter fixations on crime-related items, as well as less reading and re-reading of the crime-related items. Kim et al. (2016) asked participants to complete either a legal (i.e., innocent group) or an illegal (i.e., guilty group) task. After the task, they were shown pairs of items on a computer screen, in which one was neutral and the other was related to the illegal task. Guilty participants, compared to the innocent group, spent more time fixating the neutral stimuli and showed a pattern of avoidance of the crime-relevant items.

Overall, these studies suggest that, when being deceptive, people tend to avoid looking for too long at the incriminated stimulus, showing fewer and shorter fixations, but also that the type of stimulus and the familiarity with it plays an important role in the viewing pattern.

Therefore, the goal of the present study was to explore potential differences in eye fixation behavior between honest and dishonest participants who had (guilty) or had not (innocent) committed a mock crime, to analyze the effectiveness of eye-tracking technology in identifying deception from the visual exploration of a complex image. The present explorative study simulated a real crime scenario and exposed guilty and innocent participants to the entire crime scene. Participants' visual exploration of the crime scene was then analyzed to investigate potential differences in eye fixation between guilty and innocent, and honest and dishonest, participants. To facilitate the reader's understanding, the hypotheses have been detailed after describing the experimental design, stimuli and experimental procedure.

Materials and Methods

Participants

A total of 160 young adults voluntarily participated in the study. The only inclusion criterion was the ability to: read questions on a computer monitor, understand the meaning of those

questions, and answer the questions. The sample comprised students in a clinical psychology course who were given extra university credit for their participation and for recruiting friends and relatives ($n=56$) to participate. Three participants (1.8%) were excluded from the analysis due to technical computer problems that invalidated the procedure.

The final sample was composed of 157 participants, of whom 56 were male (35.7%) and 101 were female (64.3%). Participants ranged in age from 18 to 31 years (M=22.27, SD=2.38), and they were mostly students (N=141, 89.8%). The majority of the participants held a bachelor's degree (N=90, 57.3%), were Italian citizens (N=153, 97.5%), and lived in the center of Italy (N=114, 72.6%). A between-subjects experimental design was implemented. Participants were randomly assigned to one of four experimental conditions, defined by a combination of the manipulated variables: (a) visiting the target room vs. not visiting it, (b) stealing and photographing an exam paper hidden in the target room vs. not knowing that there was an exam paper, and (c) receiving the instruction to answer truthfully vs. deceptively when later questioned. In more detail, the four groups were as follows:

- *Guilty Honest* was comprised of 40 participants who visited the target room and were instructed to steal the exam paper and answer honestly, when later questioned;
- *Guilty Deceptive* was comprised of 38 participants who visited the target room and were instructed to steal the exam paper and answer deceptively when later questioned;
- *Innocent Honest - conceptually representing the control group of the present study -* was comprised of 40 participants who visited the target room, did not know that there was an exam paper, and were instructed to answer honestly when later questioned; and.
- *Naive - representing the group that is unfamiliar with the room -* was composed of 39 control participants who did not visit the target room, did not know that there was an exam paper, and were instructed to answer honestly when later questioned.

An a priori power analysis was run to determine the minimum size of the sample. It has been calculated that a sample size of 96 is sufficiently large to achieve statistical power of at least (1-β)=0.90 in a one-way ANOVA (test family=F test, statistical test=ANOVA: fixed effects, omnibus, one-way) involving four groups, given a significance level of 0.05 and a large effect size (0.40) (Faul et al., 2007).

Table 1 reports the descriptive statistics for each group, including all of the characteristics considered.

All participants provided informed consent prior to engaging in the study. The experimental procedure was approved by the local ethics committee (Board of the Department of Human Neuroscience, Sapienza University of Rome), according to the Declaration of Helsinki.

Experimental Procedure

The experimental procedure took place at the Department of Human Neuroscience, Sapienza University of Rome. Data were collected in December 2021. To ensure optimal lighting conditions, participants carried out the experimental task between 9:00 am and 3:00 pm. Following a joining phase, all participants completed an informed consent form and a sociodemographic questionnaire (see Sect. 2.3.1). Subsequently, they were randomly assigned to one of the four experimental groups. The experimental procedure lasted approx-

Table 1 Descriptive Statistics of the Sample and Differences Between Groups

		Guilty Honest N(%)	Guilty Deceptive N(%)	Innocent Honest N(%)	Naive N(%)	Total	F/χ^2	p
Age $M(SD)$		23.33 (3.16)	22.37 (1.48)	22.08 (1.89)	21.31 (2.24)	22.27 (2.38)	5.526	**0.002**
Sex							7.978	**0.046**
	Female	29 (72.5)	25 (65.8)	29 (72.5)	18 (46.2)	101 (64.3)		
	Male	11 (27.5)	13 (34.2)	11 (27.5)	21 (53.8)	56 (35.7)		
Educational level							23.833	**0.021**
	Middle school diploma	1 (2.5)	0 (0)	0 (0)	0 (0)	1 (0.6)		
	High school diploma	9 (22.5)	11 (28.9)	15 (37.5)	24 (61.5)	59 (37.6)		
	Bachelor's degree	26 (65.0)	27 (71.1)	24 (60.0)	13 (33.3)	90 (57.3)		
	Master's degree	3 (7.5)	0 (0)	1 (2.5)	2 (5.1)	6 (3.8)		
	Postgraduate	1 (2.5)	0 (0)	0 (0)	0 (0)	1 (0.6)		
Occupation							9.935	0.356
	Unemployed	1 (2.5)	1 (2.6)	1 (2.5)	1 (2.6)	4 (2.5)		
	Student	32 (80.0)	37 (97.4)	36 (90.0)	36 (92.3)	141 (89.8)		
	Employee	6 (15.0)	0 (0)	2 (5.0)	2 (5.1)	10 (6.4)		
	Freelancer	1 (2.5)	0 (0)	1 (2.5)	0 (0)	2 (1.3)		
Citizenship							3.951	0.267
	Italian	38 (95.0)	38 (100)	38 (95.0)	39 (100)	153 (97.5)		
	Other	2 (5.0)	0 (0)	2 (5.0)	0 (0)	4 (2.5)		
Vision impairment							1.601	0.659
	No	21 (52.5)	23 (60.5)	19 (47.5)	19 (48.7)	82 (52.2)		
	Yes	19 (47.5)	15 (39.5)	21 (52.5)	20 (51.3)	75 (47.8)		
Glasses/contact lenses (outside the experimental procedure)							0.869	0.833
	No	22 (55.0)	24 (63.2)	22 (55.0)	21 (53.8)	89 (56.7)		
	Yes	18 (45.0)	14 (36.8)	18 (45.0)	18 (46.2)	68 (43.3)		
Glasses during the experimental procedure							8.550	**0.036**
	No	35 (87.5)	37 (97.4)	39 (97.5)	39 (100)	150 (95.5)		
	Yes	5 (12.5)	1 (2.6)	1 (2.5)	0 (0)	7 (4.5)		

imately 15 min and, at the end of it, there was a debriefing with each participant to clarify their doubts and have feedback about their experience.

Guilty Honest

The first group was invited by the experimenter to enter the target room and told that there was an exam paper in the room that they had to photograph. Specifically, the instructions were as follows:

We will now take you into the professor's room. We ask that you observe it carefully for 2 minutes. You will find in a drawer the clinical psychology exam paper on which the correct answers are marked. You can take a photo of it but be careful as the professor

may return at any time and, if he discovers you, he will not award you the additional credit agreed upon.

Subsequently, participants were taken to another room where they were seated in front of a computer with a webcam connected to RealEye software (see Sect. 2.3). The webcam was positioned above the computer screen, approximately 50 cm from the participant. Participants were told that the professor had realized that someone had stolen the exam paper, and wanted to question the people who might have been in his room. Specifically, the instructions were as follows:

The professor has noticed that someone has stolen the clinical psychology exam paper and has decided to question those who might have been in his room.

At this point, the computer-based task began, and participants received further instructions through the computer display. The computer-based task comprised three parts:

- *Part 1—Neutral stimuli*: Participants were shown a series of six neutral pictures, each depicting a room they had never seen (Fig. 1). The inclusion of neutral images was implemented to exclude the possibility that the observed differences in visual exploration patterns were due to variations in spontaneous exploration behaviors among participant groups. They were asked to inspect the pictures carefully, as they would later be asked to recognize those pictures from a larger set of pictures (see, e.g., Fig. 1). Specifically, the instructions were as follows:

Fig. 1 Example Picture of a Neutral Room (Computer-Based Task Parts 1 and 2)

Now you will be shown six images that you have to remember. Look at them carefully. Each picture will be shown for 10 seconds. Later you will be shown more pictures and asked to recognize the ones you are being shown now.

- *Part 2—Recognition of neutral stimuli*: Participants were shown another series of six pictures. Three of these pictures belonged to the first set of pictures, while the remaining three were new. For each picture, participants were asked to answer the question, "Have you seen this image before?" by selecting "Yes" or "No" using the computer mouse. This was designed to record participants' eye-movement exploration patterns in response to previously viewed pictures with neutral valence. Specifically, the instructions were as follows:

Just now you were shown some pictures. Now you will be shown several pictures, each for 10 seconds. For each one you should answer YES if the picture was one you saw before, or NO if it was not.

- *Part 3—Target stimuli*: Finally, participants were shown six pictures of the professor's room captured from different angles (see Fig. 2). After every two pictures, they were asked to answer honestly if they had ever seen the room and, if so, how many times. This was designed to capture participants' eye-movement exploration patterns in response to the room in which they had photographed the exam paper. Specifically, the instructions were as follows:

Now you will be shown six pictures representing a room. Look at them carefully. Each picture will be shown for 10 seconds. Answer the questions that follow honestly.

Fig. 2 Target Pictures of the Professor's Room (Computer-Based Task Part 3)

Guilty Deceptive

In the second group, participants were instructed to enter the target room and photograph the exam paper. They received the same instructions as given to the *Guilty Honest* group. After the first part of the experiment, they were taken to another room, where they completed Parts 1 and 2 of the computer-based task described above. In Part 3 of the computer-based task, they were instructed to lie. Specifically, the instructions were as follows:

> *Now you will be shown six pictures representing a room. Look at them carefully. Each picture will be shown for 10 seconds. Answer the questions that follow, lying.*

Innocent Honest

In the third group, participants were instructed to enter the target room and observe it carefully for 2 min. Unlike the participants in the *Guilty Honest* and *Guilty Deceptive* groups, participants in this group were unaware that there was an exam paper hidden in the room. Specifically, the instructions were as follows:

> *We will now take you into the professor's room. We ask you to observe it carefully for 2 minutes.*

After this first part, participants followed the same procedure as employed for the *Guilty Honest* group. They were told that the professor had noticed that someone had photographed the clinical psychology exam in his room and he wanted to question the people who might have been there. Subsequently, they were instructed to complete the computer-based task, answering honestly.

Naive

Participants in the *Naive* group did not visit the target room. After completing the informed consent form and the sociodemographic questionnaire, they were asked to complete the computer-based task honestly, following the same instructions as provided to the *Innocent Honest* group.

Materials

Sociodemographic Questionnaire

An ad-hoc sociodemographic questionnaire was administered to all participants, to collect data on biological sex, age, education, occupational status, region of residence, and citizenship. Moreover, information was collected about visual impairments (e.g., myopia, astigmatism) and the habitual use of glasses or contact lenses.

Target Room

The target room (see Sect. 2.2 and Fig. 2) was located in the Department of Human Neuroscience, Sapienza University of Rome. The room was the actual office of the clinical psychology professor who granted extra university credit to participating students. A clinical psychology exam (with correct answers marked) was hidden inside the desk drawer near the wall. No participants had ever seen the room prior to the experimental procedure.

Computer-Based Task

The computer-based task was programmed using RealEye (RealEye sp. Z o.o., Poland), which is an online platform for screen-based webcam eye-tracking research. RealEye - and more in general webcam-based eye tracking systems, was proved to be a reliable low-cost alternative to remote eye tracking (Wisiecka et al., 2022). A study to measure the program's accuracy was conducted. In this experiment, participants were asked to click on 35 measuring points evenly distributed on the screen while eye measures were taken. The authors concluded that the average accuracy is 90 to 156 px (depending on the place on the screen), and the average accuracy for all measurements is 113 px (Lewandowska, 2020).

Prior to beginning the task, all participants underwent a RealEye calibration procedure (predetermined by the platform) that consisted of looking at and clicking on a moving dot. During the entire three-part computer-based task, RealEye captured participants' eye movements, gazes, and fixations through a webcam. Participants' privacy was guaranteed, as no images or sounds were recorded. The personal computer and webcam that were used met the necessary standards for the RealEye software. In particular, the computer operating system was Windows 10, the computer processor was an Intel Core i7 (6 core, 3.2 GHz), the computer RAM was 16 GB, and the RealEye software was run through an updated version of Google Chrome. An HD webcam capturing 1080p @ 60 frames per second was used.

Collected Measures

During the computer-based task, the RealEye software captured participants' gazes and fixations, to describe their patterns of exploring the visual stimuli. A gaze can be described as a point of visual focus, or a point of where a person was looking. In this context is expressed as a percent of the width (x-axis) and height (y-axis) of the visual object (i.e., picture), whereby the left edge is 0% and the right edge is 100%, and the top of the picture is 0% and the bottom is 100%. Based on the RealEye algorithm, the gazes' sampling frequency is 16 ms (i.e., at least one gaze is recorded every 16 ms). A fixation is defined as a series of gaze points (i.e., at least two gazes) that are very close in time and space, so that the user's gaze stops for long enough for a person to focus and process what he is seeing, it will be called fixation. RealEye automatically determines fixations according to certain parameters: minimum fixation duration (typically 100–300 ms) and speed. It automatically sets a "velocity limit" below which data are classified as fixations, and above which data are considered saccades. The program then applies a noise reduction filter to the collected data to reduce the influence of noise on the actual results. This replaces each point with the median of a certain number of points (called the noise reduction level). If the noise reduction level is set to 3, a median is calculated for three consecutive points.

Specifically, RealEye generates data for fixation and gaze features (e.g., x- and y- coordinates). Furthermore, in the present study, four additional features were computed for all participants and each image (e.g., total number of fixations). The comprehensive list of these features can be found in the Supplementary Information (S1 and S2).

As already described, these features were obtained for the full pictures displayed to participants. In addition, the total number and the sum duration of fixations were calculated for image areas of interest (AOI). Specifically, each picture was divided into the following AOIs:

- left: corresponding to the left part of the image (i.e., x-axis 0–33.3%);
- center: corresponding to the central part of the image (i.e., x-axis 33.3–66.6%); and.
- right: corresponding to the right part of the image (i.e., x-axis 66.6–99.9%).

Fig. 3 presents an example of a picture divided into AOIs.

The complete list of features entered in the statistical analysis is reported in the Supplementary Information (S3).

Research Hypothesis

Drawing upon the literature reviewed, we postulate that the *Guilty* groups, that enter the room and steal the exam, will exhibit a lower frequency and briefer duration of fixations on Areas of Interest that include the desk compared with both *Innocent Honest* and *Naive*. Conversely, we do not anticipate any disparities in AOIs that do not feature the critical stimulus.
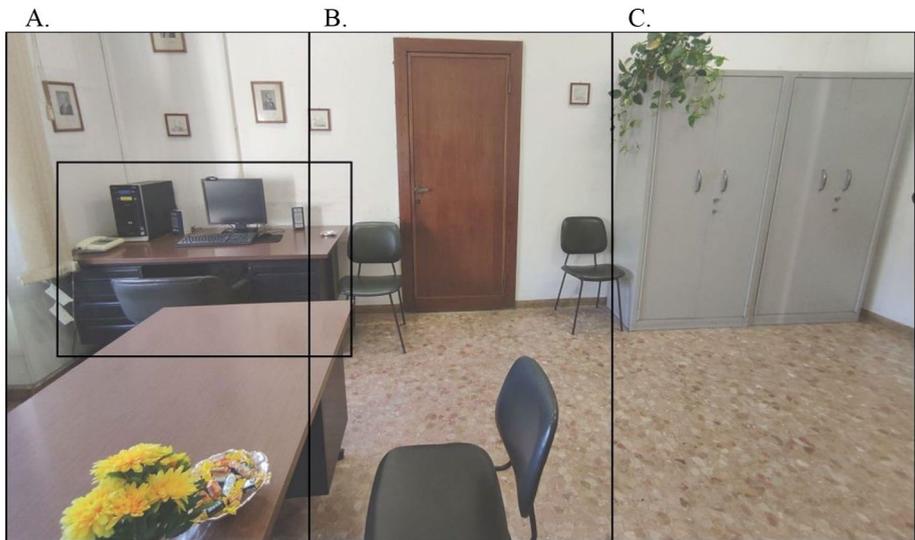

Fig. 3 Picture Target 1 Divided into AOIs
Note: Panel **A** represents the left AOI, panel **B** represents the central AOI, and panel **C** represents the right AOI. The desk that contained the exam paper that some participants photographed is visible in panel A (i.e., the left AOI).

Between the two guilty groups, we expect that the *Guilty Deceptive* group, in turn, will have fewer fixations and shorter durations compared to the *Guilty Honest* participants.

The *Naive* is expected to display an intermediate behavior relative to the other groups, demonstrating a more uniform pattern of visual exploration. Consequently, they will observe the stimuli without any noticeable distinctions, exhibiting in the AOIs including the desk more fixations than the *Guilty* groups but fewer fixations than the *Innocent Honest* group, who have already become acquainted with the room.

Moreover, we expect no differences between the four groups in image Target 2, where the desk is not present, and in neutral images.

Data Analysis

A one-way independent ANOVA was run for each feature to identify potential differences between the four experimental groups (i.e., *Guilty Honest*, *Guilty Deceptive*, *Innocent Honest*, *Naive*). The effect sizes of the score differences between groups were reported. As concerns magnitude, $\eta^2 = 0.01$ was considered indicative of a small effect size, $\eta^2 = 0.06$ was considered indicative of a medium effect size, and $\eta^2 = 0.14$ was considered indicative of a large effect size (Cohen, 1988). A Tukey test was performed as a post-hoc test to verify which groups accounted for the significant differences found by the ANOVA. All analyses were performed using JASP 0.14 (JASP, 2022).

Finally, we tested the possibility to differentiate between the four groups. For this purpose, we have trained and validated a number of logistic regression classification models through a 10-fold cross-validation procedure using WEKA 3.9 software (Frank et al., 2016). For each model, the following metrics are reported: ROC AUC, Accuracy, Precision, Recall, F-measure.

Results

Between-Groups Analysis of the Target Pictures

To identify an effect of the experimental condition on the visual exploration patterns of the whole target images, groups were compared for the five additional variables described above: number of fixations, duration of fixations, Euclidean distance between fixations, and time distance between fixations. To address the problem of multiple testing, a Bonferroni correction was applied, dividing the *p*-value by the number of tested variables (N=4) and setting the significance level to 0.0125 (Shaffer, 1995). Single AOIs were tested for the number and duration of fixations, with the significance level set to 0.025.

Number of Fixations

Considering the whole picture stimuli, no significant differences emerged between groups in the total number of fixations. Table S4a reports the results. However, some interesting differences were found concerning the AOIs (see Table 2).

Starting with image Target 1, the results showed a statistically significant difference in the left AOI, which displayed the desk with the exam paper (see Fig. 3). In particular, the

Table 2 Number of Fixations Between Groups in the Target Image AOIs

	Guilty Honest Avg. (SD)	Guilty Deceptive Avg. (SD)	Innocent Honest Avg. (SD)	Naive Avg. (SD)	F (df=3)	p	η²
Target 1 – left *	9.82 (4.22)	9.11 (5.22)	12.63 (5.18)	10.69 (4.05)	4.133	**0.008**	0.075
Target 1 – central	13.88 (5.33)	15.37 (6.51)	10.50 (4.52)	12.26 (4.67)	6.113	**<0.001**	0.107
Target 1 – right	4.05 (2.75)	3.68 (2.90)	4.95 (2.96)	4.10 (2.63)	1.426	0.237	0.027
Target 2 – left	6.22 (4.43)	6.05 (4.05)	6.83 (3.28)	7.15 (3.77)	0.674	0.569	0.013
Target 2 – central	16.77 (5.15)	17.97 (6.00)	14.32 (4.81)	15.69 (5.32)	3.329	**0.021**	0.061
Target 2 – right	5.25 (4.19)	5.82 (4.24)	7.72 (4.08)	6.15 (4.33)	2.530	0.059	0.047
Target 3 – left *	6.53 (5.06)	5.39 (3.37)	6.42 (5.12)	5.49 (2.36)	0.812	0.489	0.016
Target 3 – central	16.10 (6.03)	17.34 (5.38)	13.55 (4.88)	17.90 (4.33)	5.466	**0.001**	0.097
Target 3 – right	4.72 (3.63)	5.68 (4.29)	7.55 (4.38)	5.54 (4.59)	3.177	0.026	0.059
Target 4 – left *	7.38 (4.29)	5.13 (3.88)	7.97 (5.06)	4.92 (3.53)	5.268	**0.002**	0.094
Target 4 – central	14.53 (4.04)	14.18 (5.24)	12.25 (4.51)	15.33 (4.89)	3.095	0.029	0.057
Target 4 – right	5.50 (4.21)	7.18 (3.64)	8.03 (4.86)	7.56 (4.41)	2.605	0.054	0.049
Target 5 – left	6.00 (5.94)	5.16 (4.22)	7.50 (5.20)	6.18 (5.27)	1.359	0.257	0.026
Target 5 – central *	16.82 (6.44)	18.84 (5.67)	16.20 (5.30)	17.90 (5.48)	1.612	0.189	0.031
Target 5 – right	4.30 (3.16)	4.18 (3.67)	4.85 (3.94)	3.15 (3.00)	1.645	0.181	0.031
Target 6 – left	5.80 (5.41)	6.00 (5.20)	7.83 (5.98)	4.97 (3.44)	2.198	0.091	0.041
Target 6 – central	13.60 (5.56)	15.47 (4.93)	12.47 (5.38)	14.97 (5.72)	2.474	0.064	0.046
Target 6 – right *	6.22 (4.32)	5.11 (4.54)	6.75 (4.46)	5.79 (4.97)	0.899	0.443	0.017

Note. The first four columns report the average number of fixations and standard deviation of the four experimental groups with respect to each AOI in each target image. The last three columns report the results of the one-way independent ANOVA models. AOIs marked with * refer to the portion of the image in which the desk was visible. *p*-value significance was set to 0.025.
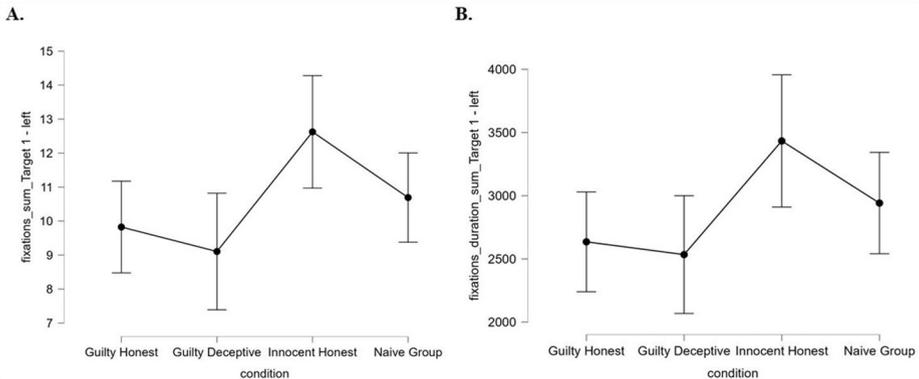

Fig. 4 Total Number of Fixations and the Sum Duration of Fixations Between Groups for the Left AOI in Image Target 1

Note: Between-group differences in image Target 1, left AOI. Panel **A** shows differences in the total number of fixations within the left AOI of image Target 1; Panel **B** shows differences in the cumulative duration of fixations within the left AOI of image Target 1.

post-hoc analysis revealed a statistically significant difference between *Guilty Honest* and *Innocent Honest* participants ($t = -2.668$, $p_{tukey} = 0.042$) and between *Guilty Deceptive* participants and *Innocent Honest* ($t = -3.310$, $p_{tukey} = 0.006$). Of note, participants who entered the room and photographed the exam (*Guilty Honest* and *Guilty Deceptive*) made fewer fixations in the area where the event took place than those who entered but did not commit the crime. This between-groups difference is demonstrated in Fig. 4, panel A. Consequently, there was also a difference in the central AOI, with *Guilty Honest* and *Guilty Deceptive* participants fixating on different locations relative to participants in the *Innocent Honest* group. In fact, the post-hoc analysis revealed a statistically significant difference between *Guilty Honest* and *Innocent Honest* groups ($t = 2.848$, $p_{tukey} = 0.026$) and between *Guilty Deceptive* and *Innocent Honest* groups ($t = 4.055$, $p_{tukey} < 0.001$). No differences were found between groups for the right AOI.

Figure 5 displays the heat maps of the visual exploration of image Target 1 of four participants who were prototypal of each of the four experimental conditions, respectively. Of note, while the *Innocent Honest* and *Naive* participants displayed a uniform exploration pattern, the *Guilty Honest* and *Guilty Deceptive* participants had few fixations in the left AOI (or any AOI that included the desk with the exam they had been instructed to photograph). Furthermore, all participants had few fixations in the right AOI, which contained very few observable elements.

As concerns image Target 2, no differences were found for the left and right AOIs. The only difference emerged for the central AOI, between *Guilty Deceptive* and *Innocent Honest* groups ($t = 3.024$, $p_{tukey} = 0.015$). Of note, this image did not include the desk with the exam.
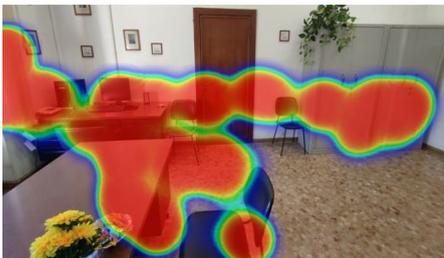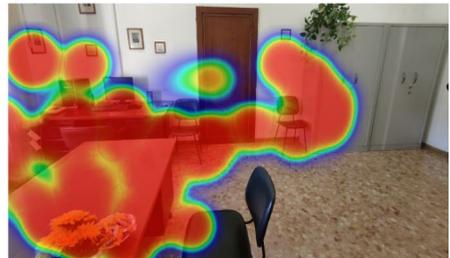
A.

B.
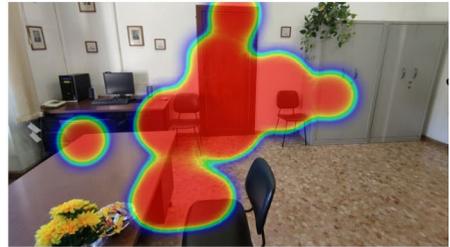
C.

D.

Fig. 5 Heat Maps
Note: The heatmaps show the prototypical exploration patterns of image Target 1 by participants assigned to *Guilty Honest* (panel **A**), *Guilty Deceptive* (panel **B**), *Innocent Honest* (panel **C**), and *Naive* (panel **D**). The desk that contained the exam paper that some participants photographed is visible in the left AOI.

In image Target 3, a significant difference between groups was found in the central AOI. Specifically, the post-hoc test highlighted a difference between *Guilty Deceptive* and *Innocent Honest* groups ($t=3.220$, $p_{tukey}=0.008$) and between *Innocent Honest* and *Naive* ($t = -3.716$, $p_{tukey}=0.002$).

As regards image Target 4, differences were found in the left AOI, which depicted the desk. Specifically, the post-hoc analysis revealed a difference between *Guilty Deceptive* and *Innocent Honest* groups ($t = -2.964$, $p_{tukey}=0.018$) and between *Innocent Honest* and *Naive* ($t=3.203$, $p_{tukey}=0.009$). In other words, participants in the *Guilty Deceptive* group, who were instructed to steal the exam and lie, seemed to avoid the area of the picture that included the desk (i.e., the left AOI), and had fewer fixations in that AOI relative to participants in the *Innocent Honest* group, who did not commit the crime.

Finally, no differences were found for image Targets 5 or 6.

To summarize the aforementioned results, regarding image Target 1, participants who committed the crime (*Guilty Honest* and *Guilty Deceptive*) had fewer fixations in the area where the event took place compared to those who did not commit the crime (*Innocent Honest*). Conversely, these participants exhibit a tendency to fixate more on the central AOI, thereby significantly distinguishing themselves from *Innocent Honest* participants.

In image Target 2 (which does not contain the desk) and in Target 3 (which contains the desk on the left), a significant difference was found only in the central AOI, with *Guilty Deceptive* participants once again staring more at the central area compared to *Innocent Honest*.

Regarding image Target 4, differences were found in the left AOI (containing the desk). The *Guilty Deceptive* group, who was instructed to steal the exam and lie, tended to avoid fixating on the area of the picture that included the desk, with fewer fixations in that AOI compared to the *Innocent Honest* group. The latter, on the other hand, display a greater frequency of fixations within this AOI, even compared to the *Naive*.

No differences were found for image Targets 5 or 6.

Finally, considering the impact that spatial location appears to have in the previously presented results, additional analysis was implemented by calculating the Root Mean Squared Distance (RMSD) from the center of each AOI that showed significant results in the aforementioned analysis. Results are reported in the Supplementary Information (S5).

Sum Duration of Fixations

The ANOVA generated no significant results for the sum duration of fixations for any target image in its entirety (see Table S4b). However, some differences emerged for the AOIs (see Table 3). Regarding image Target 1, the ANOVA showed a significant result for the left AOI (where the desk was visible; see Fig. 3). Specifically, the post-hoc analysis revealed a statistically significant difference between the *Guilty Deceptive* and *Innocent Honest* groups ($t = -2.853$, $p_{tukey}=0.025$). In other words, participants who entered the room, photographed the exam, and then lied, demonstrated shorter fixations in the area where the event took place than those who entered the room but did not photograph the exam. This difference between groups is well represented in Fig. 4, Panel B. Consequently, there was also a difference in the central AOI: participants in the *Guilty Honest* and *Guilty Deceptive* groups showed greater fixation in this area, in contrast to participants in *Innocent Honest* and *Naive*. Specifically, the post-hoc analysis revealed a statistically significant difference between *Guilty*

Table 3 Sum Duration of Fixations Between Groups in the Target Images

	Guilty Honest	Guilty Deceptive	Innocent Honest	Naive	F	p	η²
	Avg. (SD)	Avg. (SD)	Avg. (SD)	Avg. (SD)	(df=3)		
Target 1 – left *	2634.87 (1234.81)	2533.95 (1417.61)	3433.35 (1636.26)	2941.74 (1237.30)	3.327	**0.021**	0.061
Target 1 – central	3306.68 (1488.50)	3895.71 (1707.43)	2450.93 (1281.29)	2835.82 (1284.16)	7.200	**<0.001**	0.124
Target 1 – right	988.68 (612.94)	850.58 (656.51)	1119.93 (801.88)	957.26 (689.76)	0.998	0.396	0.019
Target 2 – left	1381.45 (1022.52)	1475.89 (1174.50)	1628.50 (774.12)	1750.05 (975.09)	1.058	0.369	0.020
Target 2 – central	4139.15 (1453.10)	4266.40 (1521.33)	3318.48 (1127.64)	3787.39 (1475.18)	3.618	**0.015**	0.066
Target 2 – right	1360.97 (1136.13)	1477.97 (1251.86)	1866.05 (1099.82)	1635.26 (1231.12)	1.364	0.256	0.026
Target 3 – left *	1443.45 (1241.95)	1340.16 (920.46)	1435.38 (1056.78)	1186.33 (522.79)	0.592	0.621	0.011
Target 3 – central	4273.15 (1573.69)	4496.84 (1200.50)	3722.20 (1378.30)	4525.85 (1301.38)	2.909	0.036	0.054
Target 3 – right	1359.75 (1292.81)	1517.87 (1247.29)	2231.83 (1560.16)	1428.03 (1257.80)	3.574	**0.015**	0.065
Target 4 – left *	1871.60 (1191.94)	1389.66 (999.71)	2025.42 (1216.60)	1343.97 (983.66)	3.768	**0.012**	0.069
Target 4 – central	399.50 (1496.63)	3720.37 (1427.27)	2875.35 (1095.21)	3896.31 (1423.70)	5.128	**0.002**	0.091
Target 4 – right	1405.45 (1057.41)	1982.87 (1178.44)	2125.60 (1277.01)	1937.28 (1148.03)	2.970	0.037	0.054
Target 5 – left	1451.53 (1450.75)	1503.53 (1402.55)	1815.63 (1384.73)	1534.23 (1492.36)	0.517	0.671	0.010
Target 5 – central *	5001.23 (1863.33)	5574.16 (1729.33)	4465.43 (1563.58)	5204.87 (1797.77)	2.762	0.044	0.051
Target 5 – right	1166.45 (1084.39)	1169.42 (1329.18)	1294.05 (1163.95)	775.49 (867.42)	1.581	0.196	0.030
Target 6 – left	1554.10 (1408.93)	1466.79 (1305.74)	2180.68 (1915.73)	1420.77 (1387.97)	2.136	0.098	0.040
Target 6 – central	3932.25 (1731.60)	4619.00 (2053.35)	3397.95 (1763.85)	3964.13 (1826.96)	2.853	0.039	0.053
Target 6 – right *	1872.28 (1458.29)	1402.84 (1217.77)	1862.70 (1356.62)	1849.39 (1731.99)	0.957	0.415	0.018

Note. The first four columns report the average sum duration of fixations and standard deviation of the four experimental groups with respect to each AOI of each target image. The final three columns report the results of the one-way independent ANOVA models. AOIs marked with * refer to the portion of the image in which the desk was visible. *p*-value significance was set to 0.025

Honest and *Innocent Honest* ($t=2.642$, $p_{tukey}=0.045$), *Guilty Deceptive* and *Innocent Honest* ($t=4.403$, $p_{tukey}<0.001$), and *Guilty Deceptive* and *Naive* ($t=3.210$, $p_{tukey}=0.009$). No differences between groups were found for the right AOI.

As regards image Target 2, no differences were found with respect to the left or right AOIs. The only difference emerged with respect to the central AOI, between *Guilty Honest* and *Innocent Honest* groups ($t=2.620$, $p_{tukey}=0.047$) and between *Guilty Deceptive* and *Innocent Honest* groups ($t=2.987$, $p_{tukey}=0.017$). Of note, the desk with the exam was not present in this image, but located behind the observer/the point of observation of the image.

In image Target 3, the ANOVA showed a significant result for the right AOI, only. Specifically, the post-hoc analysis highlighted a significant difference between *Guilty Honest* and *Innocent Honest* ($t = -2.894$, $p_{tukey}=0.022$), and between *Innocent Honest* and *Naive* ($t=2.651$, $p_{tukey}=0.044$).

Also, the ANOVA provided significant results for image Target 4. Specifically, the post-hoc test showed statistically significant differences with respect to the left and central AOIs: in the left AOI between *Innocent Honest* and *Naive* ($t=2.740$, $p_{tukey}=0.034$); and in the central AOI between *Guilty Honest* participants and *Innocent Honest* ($t=3.380$, $p_{tukey}=0.005$), *Guilty Deceptive* participants and *Innocent Honest* ($t=2.726$, $p_{tukey}=0.036$), and *Innocent Honest* and *Naive* ($t =-3.315$, $p_{tukey}=0.006$).

Finally, no differences were found for image Targets 5 and 6 for the left, central, and right AOIs.

To summarize the aforementioned results, in Target 1, *Guilty Deceptive* participants, who were instructed to photograph the exam and lie when questioned, spent less time (in ms) looking at the AOI with the desk compared to participants in the *Innocent Honest* group. On the contrary, *Guilty Honest* and *Guilty Deceptive* participants appear to spend significantly more time observing the central AOI compared to both the *Innocent Honest* and *Naive* groups.

In image Target 2 (which does not contain the desk) a significant difference was found only in the central AOI, with both *Guilty Honest* and *Guilty Deceptive* participants making longer fixations in this area than *Innocent Honest* group.

In image Target 3 significant results were found only for the right AOI (which does not contain the desk). In particular, *Innocent Honest* group observes this AOI for a longer time than both *Naive* and *Guilty Honest* groups.

Regarding image Target 4, differences were found in the left AOI between *Innocent Honest* and *Naive* groups, with the first showing longer fixations than the second. Moreover, the *Guilty Honest*, *Guilty Deceptive* and *Naive* participants significantly differ from *Innocent Honest* group by exhibiting longer fixations in the central AOI.

No differences were found for image Targets 5 or 6.

Euclidean and Time Distance Between Fixations

In addition to the above-described variables, Euclidean distance between fixations and temporal distance between fixations were also calculated for the entire images. The ANOVA revealed no significant differences between groups for these variables. Tables S4c and S4d, report the results.

Classification Models

In order to test the possibility to differentiate between the four groups, a logistic regression classification model on four classes (*Guilty Honest* vs. *Guilty Deceptive* vs. *Innocent Honest* vs. *Naive*) was run. The model was trained and tested through a 10-fold cross-validation procedure. The number of fixations and the sum duration of fixations for any target image AOI were entered as predictors (number of predictors=36, see Supplementary Information - S6, for the complete list), while the Euclidean distance between fixations and gazes and the temporal distance between fixations were excluded as the statistical analysis did not show any significant results for these variables. The model showed a ROC AUC=0.63 (model Accuracy=40.13%, Precision=0.39, Recall=0.40, F-measure=0.40, number of *Guilty Honest* correctly classified=10/40, number of *Guilty Deceptive* correctly classified=15/38, number of *Innocent Honest* correctly classified=19/40, number of *Naive* correctly classified=19/39).

Moreover, running a number of logistic regression models on two groups, which were trained and validated through a 10-fold cross-validation procedure on the number of fixations and the sum duration of fixations for any target image AOI, we have checked the possibility to distinguish between the following groups:

- *Guilty Deceptive* vs. *Innocent Honest*: these two groups are the most interesting to distinguish, because they are the ones that most reflect investigative situations. In this case, people from both groups entered the room. However, some were guilty for the theft of the exam paper, some were not. In other words, some had something to hide (indeed, they deny having entered the room), some didn't. The model showed a ROC AUC=0.72 (model Accuracy=65.38%, Precision=0.65, Recall=0.65, F-measure=0.65, number of *Guilty Deceptive* correctly classified=26/38, number of *Innocent Honest* correctly classified=25/40).
- *Guilty Deceptive* vs. *Guilty Honest*: participants of these two groups are both guilty of stealing the exam paper. However, just *Guilty Deceptive* participants tried to hide the fact that they entered the professor's room. Running the model, we have obtained a ROC AUC=0.61 (model Accuracy=60.26%, Precision=0.60, Recall=0.60, F-measure=0.60, number of *Guilty Deceptive* correctly classified=23/38, number of *Guilty Honest* correctly classified=24/40).
- *Guilty Deceptive* vs. *Naive*: this case reflects the situation in which it has to be established whether a person was at the scene of a crime and is guilty of committing it, or whether he/she has never been at the scene of the crime and, therefore, cannot be guilty. Again, the goal is to detect who is guilty and who is not. The model revealed a ROC AUC=0.66 (model Accuracy=61.04%, Precision=0.61, Recall=0.61, F-measure=0.61, number of *Guilty Deceptive* correctly classified=24/38, number of *Naive* correctly classified=23/39).

Finally, to study which are the AOI that best differentiate between groups, for each group comparison (*Guilty Deceptive* vs. *Innocent Honest*, *Guilty Deceptive* vs. *Guilty Honest*, *Guilty Deceptive* vs. *Naive*) we have computed a number of point biserial correlations between the "group" variable and the predictors entered in the classification model (number of fixations and sum duration of fixations for any target image AOI). Results demonstrate

that *Guilty Deceptive* and *Innocent Honest* mainly differ for the number and duration of fixations in the central AOI of image Target 1. Also, the number of fixations in the left AOI of image Target 1, where the desk is located, has a moderate correlation with the "group" variable. The AOI that best differentiate between *Guilty Deceptive* and *Guilty Honest* is the left area of image Target 4, which corresponds to the desk location. Again, the central AOI of image Target 1 is the location that mainly discriminates between *Guilty Deceptive* and *Naive*. The correlational values are entirely reported in Supplementary Information (S7).

Between-Groups Analysis of the Neutral Recognized Pictures

To exclude the possibility that the effect found in the previous analysis was due to different patterns of exploring previously seen stimuli between groups, rather than the experimental condition, the number of fixations and the sum duration of fixations in the recognized neutral pictures in Part 2 of the computerized task were analyzed. A Bonferroni correction was applied, dividing the *p*-value by the number of tested variables (N=2) and setting the significance level to 0.025 (Shaffer, 1995).

Number of Fixations

The ANOVA showed no significant effect of group on the number of fixations. In particular, no effect was found for the overall images (see Table S4a) or AOIs (see Table 4).

Fixation Duration

With respect to the entire images, the ANOVA did not highlight any significant effect of group on fixation duration (see Table S4b). Regarding the AOIs, no differences between groups were detected in the central, left, or right AOIs in the three neutral images. The only detected differences were between *Guilty Deceptive* and *Innocent Honest* ($t=2.670$, $p_{tukey}=0.041$) and between *Innocent Honest* and *Naive* ($t = -2.921$, $p_{tukey}=0.021$) for the central AOI of the image Neutral R2. Table 5 reports the results.

Table 4 Number of Fixations Between Groups in the Neutral Images

	Guilty Honest	Guilty Deceptive	Innocent Honest	Naive	F	*p*	η²
	Avg. (*SD*)	Avg. (*SD*)	Avg. (*SD*)	Avg. (*SD*)	(*df*=3)		
Neutral_R 2 – left	3.15 (2.48)	2.42 (1.73)	3.38 (2.46)	2.56 (2.34)	1.582	0.196	0.030
Neutral_R 2 – central	9.58 (3.27)	10.29 (2.95)	8.80 (3.53)	10.31 (2.69)	2.058	0.108	0.039
Neutral_R 2 – right	2.85 (2.50)	3.13 (2.35)	3.80 (2.79)	2.36 (2.15)	2.587	0.055	0.048
Neutral_R 3 – left	3.18 (2.51)	3.82 (3.24)	3.23 (2.53)	2.90 (2.74)	0.750	0.524	0.014
Neutral_R 3 – central	8.60 (3.22)	8.00 (3.66)	8.30 (3.02)	9.46 (2.99)	1.477	0.223	0.028
Neutral_R 3 – right	3.28 (3.09)	3.37 (2.27)	3.28 (2.43)	2.44 (2.10)	1.188	0.316	0.023
Neutral_R 4 – left	3.83 (2.96)	2.82 (2.32)	3.10 (2.39)	3.18 (3.07)	0.974	0.407	0.019
Neutral_R 4 – central	8.83 (3.03)	9.68 (3.09)	8.93 (2.85)	9.56 (3.30)	0.795	0.498	0.015
Neutral_R 4 – right	1.85 (1.61)	2.26 (2.29)	2.40 (2.13)	1.33 (1.90)	2.265	0.083	0.043

Note. The first four columns report the average number of fixations and standard deviation of the four experimental groups with respect to each AOI of each neutral image during the recognition task. The final three columns report the results of the one-way independent ANOVA models

Table 5 Sum Duration of Fixations Between Groups in the Neutral Images

	Guilty Honest	Guilty Deceptive	Innocent Honest	Naive	F	p	η²
	Avg. (SD)	Avg. (SD)	Avg. (SD)	Avg. (SD)	(df=3)		
Neutral_R 2 – left	623.03 (468.14)	516.16 (393.52)	700.18 (549.44)	523.72 (583.88)	1.186	0.317	0.023
Neutral_R 2 – central	2194.45 (870.00)	2365.84 (730.42)	1888.1 (732.33)	2407.33 (815.90)	3.524	**0.017**	0.065
Neutral_R 2 – right	620.63 (472.00)	710.87 (536.46)	807.45 (637.94)	497.03 (473.28)	2.410	0.069	0.045
Neutral_R 3 – left	751.35 (628.80)	922.21 (737.25)	873.70 (798.46)	736.33 (724.67)	0.618	0.604	0.012
Neutral_R 3 – central	2091.55 (898.63)	1941.21 (1001.74)	2034.73 (890.47)	2389.56 (902.37)	1.707	0.168	0.032
Neutral_R 3 – right	836.33 (793.40)	815.11 (551.48)	763.73 (589.61)	613.41 (539.47)	1.001	0.394	0.019
Neutral_R 4 – left	896.20 (716.12)	734.37 (599.71)	777.65 (563.77)	785.92 (792.96)	0.411	0.745	0.008
Neutral_R 4 – central	2368.40 (881.12)	2586.47 (769.13)	2336.50 (773.70)	2683.33 (986.48)	1.517	0.212	0.029
Neutral_R 4 – right	414.50 (354.87)	556.53 (650.90)	545.95 (519.48)	297.44 (388.56)	2.430	0.067	0.045

Note. For each AOI, the first four columns report the average sum duration of fixations and standard deviation of the four experimental groups with respect to each neutral image during the recognition task. The final three columns report the results of the one-way independent ANOVA models.

Discussion

The present study aimed at detecting potential differences in eye fixation patterns according to the participant conditions of being honest or dishonest and having committed (or not) a mock crime. To summarize, the analysis revealed a significant difference between the four experimental groups in the number of fixations and the sum duration of fixations. In particular, the analysis of the first target image (Target 1) revealed interesting differences in participants' exploration patterns of the portion of the image that included the desk with the exam that some participants photographed (i.e., the left AOI). Participants who entered the room and photographed the exam (*Guilty Honest* and *Guilty Deceptive* groups) demonstrated fewer fixations than the group who entered the room but did not commit the crime (*Innocent Honest*); moreover, participants in *Guilty Honest* and *Guilty Deceptive* groups focused their attention and fixated more on the center AOI, in contrast to *Innocent Honest* participants. This result may be interpreted as an attempt by *Guilty Honest* and *Guilty Deceptive* participants to avoid looking at the portion of the image where the crime took place, perhaps because they felt morally uncomfortable with their actions, or even shame and embarrassment. Furthermore, this effect may have been more pronounced for participants who were instructed to lie about their crime, in comparison with those who were instructed to respond honestly. However, further research is needed to verify this idea. Moreover, it should be noted that we conducted additional analyses to evaluate the root mean squared distance (RMSD) from the center of each AOI that showed significant results in the analysis of the number of fixations. Among these RMSD analysis, the only one that yielded statistically significant results was on the left AOI of image Target 1, revealing that the few fixations

made by guilty subjects are concentrated on the desk (see Supplementary Information, S5). In front of the larger scene, they tend to look away from the area containing the incriminated stimulus. However, analysing just the specific area containing the incriminated target, it emerges a tendency of subjects to make their - even few - fixations on it.

The findings related to the number of fixations are also supported by the results regarding the sum duration of fixations. Indeed, *Guilty Deceptive* participants, who were instructed to photograph the exam and lie when questioned, spent less time (in ms) looking at the AOI with the desk compared to *Innocent Honest* participants. This might suggest that participants in the *Guilty Deceptive* group consciously avoided inspecting the area in which the desk was displayed (i.e., the left AOI), relative to participants who did not commit the crime (*Innocent Honest* group), and instead tried to direct their gaze to the central part of the image.

These results are aligned with the work of Pittarello et al. (2016), who showed that, when participants were instructed to be dishonest, they allocated less attention (i.e., fewer and shorter fixations) to the target card when it was presented to them, relative to honest participants. Likewise, Cook et al. (2012) reported shorter durations of fixating on, reading, and rereading deceptive items by participants instructed to lie, relative to honest participants. Similarly, in a study conducted by van Hooft and Born (2012), participants in the deceptive group displayed, on average, one eye fixation less per item compared to participants in the honest condition. Also, the present results are aligned with the work of Kim et al. (2016), who found that deceptive participants who had committed a mock crime tended to avoid directing their gaze toward the crime-related stimulus and spent more time looking at neutral stimuli.

Concerning the other target images, a similar exploration pattern (relative to the number of fixations and the sum duration of fixations) emerged for image Target 4. Again, participants in the *Guilty Deceptive* group displayed fewer fixations than participants in the *Innocent Honest* group (who did not commit the crime) in the left AOI, suggesting an attempt to avoid the area of the image that included the desk. This avoidance was also reflected in the sum duration of fixations, as *Guilty Honest* and *Guilty Deceptive* participants had longer fixations in the central AOI.

On the other hand, in image Targets 2, 3, 5, and 6, no significant differences were found between groups in the AOI of interest (i.e., the AOI where the desk was visible). In more detail, the desk was not at all visible in image Target 2, while in image Target 3, the desk was only partly visible, and the drawer containing the exam was obscured. Finally, in image Targets 5 and 6, the desk was visible in only a small portion of the images and from a distant perspective; this may be why some participants did not focus their gaze on it. Moreover, it is worth noting that the results became less evident as the test progressed. This could potentially be attributed to a progressive fatigue or habituation effect, whereby participants may have given more attention to the first images and less attention to images presented later in the task.

Overall, the analysis of the visual exploration pattern of the target room allows to distinguish between guilty, innocent, deceptive and honest participants with a performance better than the chance level. The group of participants that are best distinguishable from *Guilty Deceptive* are the *Innocent Honest* (ROC AUC=0.72), revealing that the analysis of the visual exploration pattern may be a valuable aid to detect people who know to be guilty and are trying to hide something, compared to innocent people.

The analysis of the neutral images confirmed that the effect found for the target images with respect to the number and sum duration of fixations was not due to a difference in participants' spontaneous exploration patterns or the repeat exposure of the stimuli, but potentially attributable to the experimental condition.

To conclude, the findings suggest that eye-tracker parameters may provide interesting results for deception detection and contribute to the identification of honest versus dishonest and guilty versus innocent individuals through an analysis of eye fixations patterns (i.e., the number and sum duration of fixations) in response to a complex scene. Based on the results derived from our study, it would seem that the application of the eye-tracking technique confers notable benefits under specific conditions. Specifically, when the target of interest is unambiguously discernible, fully represented within the images, particularly in the initial frames, when participants are not influenced by the gradual onset of task fatigue or the attenuating effects of habituation. Furthermore, our findings support the notion that the disparities in eye-tracking patterns are more pronounced between the *Guilty Deceptive* and *Innocent Honest* groups, in line with our initial hypotheses.

Furthermore, concerning the *Naive*, we expected it to exhibit an intermediate behavior compared to the other groups. Specifically, we hypothesized that the *Naive* would display more fixations than the *Guilty* groups but fewer fixations than the *Innocent Honest* group, who had already familiarized themselves with the room, particularly in the areas of interest including the desk. This observed pattern is perfectly manifested in both the analyses conducted on the number of fixations and duration in Figure Target 1 (See Fig. 4, Panel A and Panel B). However, it is important to note that this difference is not statistically significant. One possible reason for this lack of significance could be attributed to certain limitations of the study, such as the relatively small sample size. Further investigations are warranted to provide additional insights and clarification on this matter.

Although our findings may be useful in orienting future research designs involving eye-tracker technologies, particularly those aimed at identifying deception, some limitations of the present study must be acknowledged. Primarily, the results are only generalizable with caution. While maximal effort was made to make the experimental situation as credible as possible (e.g., by threatening participants with the risk of losing extra university credit), participants' motivation to lie remained extrinsic. Furthermore, as data collection relied on a voluntary sampling of students, the data may have been distorted by selection bias and the fact that all participants were students of a similar age. A second limitation worth noting is associated with the sampling rate of the employed eye-tracker, RealEye, which operates at 60 Hz. This value may be relatively low compared to other high-end eye-tracking devices available on the market, offering higher sampling rates. Consequently, it is possible that certain fine-grained information regarding fixation patterns might not be captured or could be underestimated. Therefore, it is crucial to consider these limitations when evaluating the study findings.

Notwithstanding these limitations, we must emphasize the exploratory nature of the present research, which was the first study to have examined differences in eye patterns between honest, deceptive, guilty, and innocent participants in response to a complex stimulus. Thus, the findings may guide researchers in the trending field of eye-tracking research. Overall, the results are not only aligned with the current literature, but they also confirm that dishonest subjects have fewer and a shorter duration of fixations than honest subjects, even when looking at elaborate and complex stimuli. Eye movements, therefore, are a useful indicator

in deception detection. Future research should seek to confirm the present findings in eco-logical contexts.

Author Contributions MC, GM, MM, CM and PR contributed to the study conception and design. GM, MM and CM prepared the material. FB, MR and SB collected the data. GM, MM, CM performed the data analysis. MC, GM, MM, ER, FB and CM wrote the first draft of the manuscript. All authors contributed to the data discussion and interpretation. MCV, ADD, PR supervised the project. All authors commented on previous versions of the manuscript and read, revised and approved the final version of the manuscript.

Data Availability The data that support the findings of this study are available from the corresponding author, F.B., upon reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

Conflict of Interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, *88*(1), 131–151. https://doi.org/10.1037/0021-9010.88.1.131

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S. D., Woltz, D. J., & Kircher, J. C. (2012). Lyin' eyes: Ocular-motor measures of reading reveal deception. *Journal of Experimental Psychology: Applied*, *18*(3), 301–313. https://doi.org/10.1037/a0028307

Derrick, D. C., Moffitt, K., & Nunamaker, J. F. Jr. (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* (pp. 4–7).

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Frank, E., Hall, M., & Witten, I. (2016). *The WEKA Workbench. Online appendix for 'Data Mining: Practical machine learning tools and techniques'*. Morgan Kaufmann Publishers.

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: Using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 166. https://doi.org/10.3389/fnhum.2010.00166

JASP, version 0.14. Retrieved February 28 (2022) from https://jasp-stats.org/2020/10/14/introducing-jasp-0-14/

Kim, K., Kim, J., & Lee, J. H. (2016). Guilt, lying, and attentional avoidance of concealed information. *Social Behavior and Personality: An International Journal*, *44*(9), 1467–1476. https://doi.org/10.2224/sbp.2016.44.9.1467

Krapohl, D. J., McCloughan, J. B., & Senter, S. M. (2009). How to use the concealed information test. *Polygraph*, *38*(1), 34–49.

Leal, S., & Vrij, A. (2010). The occurrence of eye blinks during a guilty knowledge test. *Psychology Crime & Law*, *16*(4), 349–357. https://doi.org/10.1080/10683160902776843

Lee, J., & Ahn, J. H. (2012). Attention to banner ads and their effectiveness: An eye-tracking approach. *International Journal of Electronic Commerce*, *17*(1), 119–137. https://doi.org/10.2753/JEC1086-4415170105

Lewandowska, B. (2020, October). RealEye Eye-tracking system Technology Whitepaper [White Paper]. Realeye.io. https://www.realeye.io/whitepaper/

Lim, J. Z., Mountstephens, J., & Teo, J. (2020). Emotion recognition using eye-tracking: Taxonomy, review and current challenges. *Sensors, 20*(8), 2384. https://doi.org/10.3390/s20082384

Marchak, F. M. (2013). Detecting false intent using eye blink measures. *Frontiers in Psychology*, *4*, https://doi.org/10.3389/fpsyg.2013.00736

Meijer, E. H., Selle, N. K., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the concealed information test: A meta analysis of skin conductance, respiration, heart rate, and P300 data: CIT meta-analysis of SCR, respiration, HR, and P300. *Psychophysiology*, *51*(9), 879–904. https://doi.org/10.1111/psyp.12239

Millen, A. E., Hope, L., & Hillstrom, A. P. (2020). Eye spy a liar: Assessing the utility of eye fixations and confidence judgments for detecting concealed recognition of faces, scenes and objects. *Cognitive Research: Principles and Implications*, *5*(1), 1-18. https://doi.org/10.1186/s41235-020-00227-4

Monaro, M., Gamberini, L., & Sartori, G. (2017). Identity verification using a kinematic memory detection technique. In K. S. Hale & K. M. Stanney (Eds.), *Advances in neuroergonomics and cognitive engineering* (vol. 488, pp. 123–132). Springer International Publishing. https://doi.org/10.1007/978-3-319-41691-5_11

Monaro, M., Gamberini, L., & Sartori, G. (2018). Spotting faked identities via mouse dynamics using complex questions. In *Proceedings of the 32nd international BCS human computer interaction conference*. https://doi.org/10.14236/ewic/HCI2018.8

Nahari, T., Lancry-Dayan, O., Ben-Shakhar, G., & Pertzov, Y. (2019). Detecting concealed familiarity using eye movements: The role of task demands. *Cognitive Research: Principles and Implications*, *4*(1), 1–16. https://doi.org/10.1186/s41235-019-0162-7

Perkhofer, L., & Lehner, O. (2019). Using gaze behavior to measure cognitive load. In F. D. Davis, R. Riedl, J. vom Brocke, P.-M. Léger, & A. B. Randolph (Eds.), *Information systems and neuroscience* (vol. 29, pp. 73–83). Springer International Publishing. https://doi.org/10.1007/978-3-030-01087-4_9

Pittarello, A., Motro, D., Rubaltelli, E., & Pluchino, P. (2016). The relationship between attention allocation and cheating. *Psychonomic Bulletin & Review*, *23*(2), 609–616. https://doi.org/10.3758/s13423-015-0935-z

Proudfoot, J. G., Twyman, N. W., & Burgoon, J. K. (2013). Eye tracking and the CIT: Utilizing oculometric cues to identify familiarity with wanted persons. In *Proceedings of the rapid screening technologies, deception detection and credibility assessment symposium*. Hawaii International Conference on System Sciences (HICSS).

Proudfoot, J. G., Jenkins, J. L., Burgoon, J. K., & Nunamaker, J. F. (2016). More than meets the eye: How oculometric behaviors evolve over the course of automated deception detection interactions. *Journal of Management Information Systems*, *33*(2), 332–360. https://doi.org/10.1080/07421222.2016.1205929

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory, 15*(5), 508–525. https://doi.org/10.1080/09658210701391022

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, *46*(1), 561–584. https://www.annualreviews.org/doi/10.1146/annurev.ps.46.020195.003021

Singh, H., & Singh, J. (2012). Human eye tracking and related issues: A review. *International Journal of Scientific and Research Publications*, *2*(9), 1–9.

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin*, *143*(4), 428–453. https://doi.org/10.1037/bul0000087

Tsai, M. J., Hou, H. T., Lai, M. L., Liu, W. Y., & Yang, F. Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. *Computers & Education*, *58*(1), 375–385. https://doi.org/10.1016/j.compedu.2011.07.012

Twyman, N. W., Lowry, P. B., Burgoon, J. K., & Nunamaker, J. F. (2014). Autonomous scientifically controlled screening systems for detecting information purposely concealed by individuals. *Journal of Management Information Systems*, *31*(3), 106–137. https://doi.org/10.1080/07421222.2014.995535

van Hooft, E. A. J., & Born, M. P. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, *97*(2), 301–316. https://doi.org/10.1037/a0025711

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (Eds.). (2011). *Memory detection: Theory and application of the concealed information test*. Cambridge University Press.

Vrij, A., Oliveira, J., Hammond, A., & Ehrlichman, H. (2015). Saccadic eye movement rate as a cue to deceit. *Journal of Applied Research in Memory and Cognition*, *4*(1), 15–19. https://doi.org/10.1016/j.jarmac.2014.07.005

Wisiecka, K., Krejtz, K., Krejtz, I., Sromek, D., Cellary, A., Lewandowska, B., & Duchowski, A. (2022). Comparison of webcam and remote eye tracking. In *2022 Symposium on eye tracking research and applications* (pp. 1–7). https://doi.org/10.1145/3517031.3529615

Zagermann, J., Pfeil, U., & Reiterer, H. (2016). Measuring cognitive load using eye tracking technology in visual computing. In *Proceedings of the beyond time and errors on novel evaluation methods for visualization – BELIV '16* (pp. 78–85). https://doi.org/10.1145/2993901.2993908

Authors and Affiliations

**Marco Colasanti[1] · Giulia Melis[2] · Merylin Monaro[2] · Eleonora Ricci[3] ·
Francesca Bosco[4] · Michela Rossi[4] · Silvia Biondi[4] · Maria Cristina Verrocchio[1] ·
Alberto Di Domenico[1] · Cristina Mazza[3] · Paolo Roma[4]**

✉ Francesca Bosco
   francesca.bosco@uniroma1.it

Marco Colasanti
marco.colasanti@unich.it

Giulia Melis
giulia.melis@phd.unipd.it

Merylin Monaro
merylin.monaro@unipd.it

Eleonora Ricci
eleonora.ricci@unich.it

Michela Rossi
michelarossipsi@gmail.com

Silvia Biondi
silvia.biondi@uniroma1.it

Maria Cristina Verrocchio
mc.verrocchio@unich.it

Alberto Di Domenico
alberto.didomenico@unich.it

Cristina Mazza
cristina.mazza@unich.it

Paolo Roma
paolo.roma@uniroma1.it

[1]   Department of Psychological, Health and Territorial Sciences, University "G. d'Annunzio" of
      Chieti-Pescara, Chieti, Italy

[2]   Department of General Psychology, University of Padova, Padova, Italy

[3]   Department of Neuroscience, Imaging and Clinical Sciences, University "G. d'Annunzio" of
      Chieti-Pescara, Chieti, Italy

[4]   Department of Human Neuroscience, Sapienza University of Rome, Rome, Italy