UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale

Corso di Laurea Magistrale in Psicologia Clinica

Tesi di Laurea Magistrale

Lie Detection: analysis of response latencies extracted from the audio of an interrogation

Relatore:

Prof. Giuseppe Sartori

Correlatrice:

Dott.ssa Giulia Melis

Laureanda: **Michela Dioni**

Matricola: **2050520**

Anno Accademico 2022/2023

INDEX

INTRODUCTION

Everyone lies.

What is certain, in fact, is the fact that each of us, sooner or later, will find ourselves lying. Although telling lies is seen in a purely negative light, it is not always so: for example, there are so-called white lies, which are small lies told for altruistic purposes.

Lying represents a central aspect of life, which is why scholars have long questioned what the correct definition of lying is and what are the most effective strategies for exposing it (Ekman, 1992). Over the years, the literature has been increasingly enriched with studies and experiments that have validated new techniques for lie detection: these methodologies include psychophysiological, verbal, and behavioral techniques.

The first chapter of this paper will define the meaning of lying, the different nuances it can acquire, the main characteristics and strategies used by liars, and, finally, the techniques so far validated for lie detection.

Next, within the second chapter of the paper, the concept of alibi, the fields of its application and the difficulties that come into play when it comes to constructing one will be presented. Considering past literature, several scenarios will be described in addition to the variables of time and preparation, which undoubtedly influence the performance of those intent on lying. In addition, the cognitive approach will be presented, which aims to detect the cognitive processes involved as well as the role that both different executive functions and memory play in the production of lies. Lie Detection methods based on this approach will be presented and, in more detail, the role of cognitive load and the technique based on response latency analysis, a reaction time analysis methodology

through which it is possible to discriminate between honest and dishonest subjects, will be described. Furthermore, the main techniques of conducting interrogation and alibi-unmasking methodologies that can be used as investigative techniques within the forensic field will be explained.

In conclusion, the third chapter will begin the second part of the paper, in which the conducted Lie Detection experiment will be presented. Its objectives, experimental paradigm, procedure, and methodology used, as well as the underlying research hypotheses will be reported. Finally, Chapters 4 and 5 will analyze and discuss the results of the statistical analysis conducted using a linear mixed model and Machine Learning models: of course, all the results will be discussed in light of the present and past literature, the limitations that emerged, and the future directions that could be taken.

CHAPTER 1: LYING AND LIE DETECTION TECHNIQUES

1.1 Lying.

Lying is a central aspect of our lives, and lying is a real fact of daily life (DePaulo et al., 1996). We all lie and each of us, sooner or later, will find ourselves lying. For this reason, scholars have long questioned what the correct definition of lying is and what are the most effective strategies to be able to expose it.

1.1.1 The definition of lying, its types, and the physiological rate of lying.

There is still no unambiguous definition of lying in the literature, however, the one most widely agreed upon within the scientific community is the one that presents it as a "psychological process by which an individual consciously and deliberately attempts to persuade another person to accept as true what the liar knows to be false, in order to gain some kind of gain or avoid a loss" (Abe, 2011). Notably implicit within this definition is how unconscious acts of deception cannot, on the other hand, be considered a form of lying. It is classifiable, therefore, as an erroneous response, despite the fact that not all of the latter are definable as lies: there is, in fact, a so-called "physiology of inaccuracy" including unintentional errors of recollection due to the vulnerabilities of the mnestic process (Sartori, 2021).

Summing up, then, there are two key concepts regarding the definition of lying: first, the fact that it is an intentional act "intended to promote in another person a belief that the liar regards as false" (Zuckerman et al., 1981); and second, the fact that it is always enacted with a definite purpose, whether it is the avoidance of loss or the obtaining of

potential gain (Burgoon et Buller, 1994). Lies are not all the same: there are a variety of types of different levels of entity and based on the type and complexity of the lie, the cognitive effort put into producing it also varies. Among the simplest to formulate we find, for example, those that consist in the reversal of the truthful answer - so-called truth reversals -, among the most complex, on the other hand, we can find Machiavellian lies, the example par excellence of which is represented within the Odyssey, when Odysseus tells Polyphemus that his name is "Nobody" (Sartori, 2021). A taxonomy of lies was proposed by DePaulo and colleagues (1996), who decided to make a tripartition: explicit lies, exaggerations or minimizations, and subtle lies, which include the intentional omission of details, were distinguished. The former are outright fabrications, in which the information conveyed is completely fabricated and different from what the sender knows to be the truth; they are common within the forensic field and are often described as self-serving and malicious. Exaggerations or minimizations, on the other hand, consist of the over- or underestimation of facts. Finally, subtle lies are often considered less negative than other forms of lying: digressing or reporting partial facts, in fact, is more easily morally justifiable than reporting a complete falsehood (Vrij, 2008).

Lying is cognitively more challenging than telling the truth.

The cognitive process underlying the production of lying is not a simple one, but rather quite so and consists of several stages (Ganis et Keenan, 2009): first, it is necessary to suppress the automatic response-that is, the truth with respect to the event that has occurred-while subsequently one must keep in mind the statements made previously in such a way as to convey information that is consistent, responding appropriately to the interlocutor's questions and reactions (Vrij, 2008). Thus, according to DePaulo and colleagues' (1996) classification reported earlier, explicit lies will prove to be the most

cognitively demanding, especially when compared with the simple act of omitting information observed when telling subtle lies.

Just as there are different types of lies there are also different reasons underlying them, based on which further possible categorization can be derived. In Beata's study (2015), the authors decided to divide the motivations behind lying into two groups: beneficial motivation and protective motivation. A beneficial lie is associated with the goal or benefit the individual wants to achieve; protective motivation, on the other hand, is to protect someone from the unpleasant effects resulting from the revelation of the truth. The cases of beneficial and protective motivation always refer to oneself or another individual. The conceptualization of the motivation underlying lying is, therefore, two-dimensional and derives from the combination of the two dimensions "beneficial-protective lies" and "lies oriented toward self-to others."

Within an even more recent study by Hart et al. (2020), different categories were identified based on the motivations of lies. First, there are so-called white lies, small social lies with benevolent intentions that are often told to protect others' feelings, thus avoiding uncomfortable or painful truths. The latter are perceived as less serious and more acceptable since they seem to be real protectors of social relationships. Anti-social or vindictive lies, on the other hand, are those formulated with the intent to create harm to another. Finally, personal lies are those whose purpose lies in obtaining an advantage or the avoidance of punishment. It is precisely these latter ones that are of interest in the forensic context, which is rich in intrinsic advantages of a different nature depending on the criminal or civil context. Regardless of the type of classification we consider, the motivations for why we are inclined to lie can be many but the categories of lying can be

summarized mainly as: selfish lies, formulated to bring gain to oneself, and altruistic lies, told for the good of others. Finally, it is interesting to note that in the study by Hart et al. (2020) it was found that the categories of lies are correlated with each other, suggesting that people who tell a certain type of lie tend to tell all the others as well.

Understanding the frequency of lying in daily life is not easy. Although most individuals associate the concept of lying with something extremely negative, people admit to lying on a daily basis, on average once or twice a day (DePaulo et al., 1996). Although it is very complex to have reliable data on the number of daily lies, most research being based on self-report systems, the results of all studies measuring the frequency of lying conclude that lying is a very frequent occurrence. For example, Tyler and colleagues reported an average of 2.18 lies per 10 minutes of conversation (2006). The spontaneous and "physiological" lying rate of the average subject was also studied by Ariely and coworkers (Mazar et al., 2008) through an experiment in which subjects were financially rewarded based on the number of problems solved: participants had to self-attribute the scores and the resulting reward. From the study, it emerged how individuals, put in the condition of not being caught, lie but "no more than a little": the explanation for this apparently irrational behavior would, according to the research, have to do with the self-esteem of the subjects, who would consider "socially permissible" lying a little, as opposed to lying too much, seen instead as reprehensible. Obviously, the motivation for such behavior would also lie in the gain from the lying narrative told. Finally, it has been shown how the swearing procedure also experimentally decreases, even in a context without punishment, the physiological lying rate.

Of particular interest is the fact that the oath taken at the beginning is followed by fewer lies than the oath taken at the end. In this respect, therefore, the procedure in use in criminal trials is one that empirically decreases the possibility of intentional alteration of recollection, as the witness is asked to take the oath before producing the statement (Sartori, 2021).

1.1.2 The characteristics and strategies of liars.

Within an investigative setting, it will be necessary to take into account every aspect, including psychological, that can influence the quality of the story. The characteristics of individual subjects play a key role in the frequency and types of lies that are told. There is compelling evidence that personality traits can be used to explain individual differences in lying patterns (Hart et al., 2020). Personality can, in fact, influence individual decisions (Cobb-Clark et Schurer, 2012) and some trait variables, e.g., locus of control, are able to predict unethical behavior (Street et Street, 2006). For example, the personality traits of Machiavellianism, narcissism, and psychopathy, which make up the so-called Dark Triad, have been found to have negative social implications related to behaviors of manipulation, emotional coldness, malice, and deception (Semrad et al., 2019) and consequently also to activities such as school cheating, sexual infidelity, interpersonal aggression, and crime (Muris et al., 2017). Not surprisingly, then, the same traits are also associated with lying (Azizli et al., 2016). The Big Five model personality factors (Costa et McCrae, 1992) also showed some association with lying. The study by Gylfason et al. (2016) found that people with higher trait extroversion were more likely to lie. In addition, lower levels of amicability and conscientiousness consistently predict lying, criminal offenses, and academic cheating (Williams et al., 2010; Hart et al., 2020), while higher

levels of neuroticism are associated with academic dishonesty and deceptive self-report (Eshet et al., 2014). Despite the many findings, the often also conflicting results suggest caution in using personality measures as predictors of lying, especially since each trait can only predict certain categories of lying.

Regarding gender differences, among individuals of adult age there does not seem to be a big difference in the frequency of lying, however, a picture seems to emerge for which men and women differ in the type of lies they tell: men would tend to tell more lies to gain benefit for themselves, while women to protect others (DePaulo et al., 1996; Vrij, 2008). This female tendency to tell altruistic lies can be seen as early as childhood (Saarni, 1984).

Anxiety also plays a key role in the frequency of lying: people with low levels of anxiety have been associated with high lying scores. In addition, individuals often report anxiety, guilt and increased cognitive load when they tell a lie (Buta et al., 2020; Caso et al., 2005).

In conclusion, since lying is essentially a process of social interaction, it has been shown that people who are more careful about their public impression and appearance are more likely to lie in order to maintain a socially desirable image. The findings of Buta et al. (2020) also showed how socially desirable responses are related to deceptive attitudes and higher frequency of lying, extending the results of previous research.

The convincing liar is not free to lie indiscriminately but must be able to select unverifiable topics about which he or she can tell falsehoods. Such a selection operation requires reasoning and, consequently, mental time. For this very reason, in the investigative phase, peremptory timing of interrogation is crucial, since it prevents the

suspect from performing the mental operations of verification and constructing a lie that is therefore solid and credible (Sartori, 2021).

Empirical research has made it possible to highlight the strategies that liars use in order not to be unmasked: these include under-reporting verifiable details, that is, information that can be checked a posteriori with external methodologies and feedback (e.g., having taken a plane, staying at a certain hotel, making a phone call at a certain time, and so on). On this type of data, within the investigative context, it is not possible to lie since, since they are easily ascertainable, they would lead to the identification of the lie. The experienced liar will, therefore, expose narrative rich in details but poor in verifiable information, on which he will be careful not to lie or contradict himself. It is for this reason that, within an investigative interview, it is essential to ask the suspect for a detailed account that is, through external and objective feedback, verifiable. The analysis and quantity of verifiable details is, therefore, a useful first clue to the subject's level of sincerity (Sartori, 2021). Another strategy commonly used by liars, which emerged from Strömwall et Willén's (2011) study, is the strategy called "close to the truth," and it is one of the many explanations behind the difficulty in distinguishing between honest and liars. Indeed, individuals who lie during police interrogation often get very close to the truth, omitting only a few details. One strategy provided by other participants, also previously presented in the study by Hines et al. (2010), was to plan the lie in advance: careful planning of the lie prior to the interview was found to be crucial in an attempt to convince the investigator of one's innocence. However, other offenders stated there was no need to think ahead of the interview or to have any lying strategy, both because of the risk of forgetting the previously planned lie and because an unprepared statement appears more credible and less contrived. Not having any strategy can, therefore, be considered

strategic in itself: one hopes for a spontaneous flow of words and natural behavior (Hartwig et al., 2010). Turning to perhaps the most interesting strategy found in Strömwall et Willén's (2011) study, some of the participants stated that they do not prepare for the interrogation and do not try to regulate their behavior; instead, they wait for information about the evidence having from the police directly from the investigating officer, and then act accordingly.

The great variability in deception strategies found reflects the different approaches to police interviews, as well as inevitable individual differences. Even with regard to a specific category of strategy such as the use of eye contact, contradictory ideas emerged: some individuals believe that maintaining eye contact with the interviewer is helpful in masking the lie, while others claim the opposite.

In conclusion, then, it is clear that offenders probably cannot be considered as a homogeneous group from which the same action can be expected in an investigative interview or interrogation (Strömwall et Willén, 2011). Not surprisingly, then, most studies conducted with people working in law enforcement and the justice system have found that the accuracy of lie detection is slightly higher than chance (Bond et Depaulo, 2008). Moreover, half a century of research on deception detection has established that reliable clues to be able to discriminate lies from truth are few and variable, resulting in the general belief that there are no valid and reliable indicators to be able to discriminate lies (DePaulo et al., 2003; Levine et McCornack, 2014).

1.2 Lie Detection Techniques.

As discussed in the previous section, several studies have shown how central lying is in the lives of all human beings and how difficult it is to identify. In fact, even though many people consider themselves to be excellent lie detectors - capable of picking up whether what others say is true or false - several research have proven that a person's accuracy in assessing someone else's credibility averages 54 percent, a value slightly higher than chance, thus poor. Therefore, the scientific community has become increasingly interested in devising and implementing scientific and as accurate methods to expose lies.

In order to develop such techniques, experts have taken up proven theories on the subject of the psychology of lying, including the important multiple factor theory of Zuckerman, De Paulo and Rosenthal (1981). The authors agreed on the three main aspects that may affect the signs of lying: the emotional aspect--for which an individual who lies would tend to be more nervous and agitated; the behavioral control aspect--that is, the attempt by the liar to control his or her own behavior in order to mask the lie and the signs resulting from it; and the cognitive load aspect--studies have shown how the act of lying requires increased cognitive efforts that result from having to imagine a lie, make it consistent with the context and what was previously said, and finally express it, while suppressing the truthful content. Over the years, much forensic research has focused on these aspects and the development of techniques that could observe and evaluate them accurately and precisely: Lie Detection techniques.

1.2.1 The psychophysiological Lie Detection techniques.

A study by DePaulo et. al (2003) showed that liars tend to appear more nervous and uncomfortable, either because of fear of being caught or because of guilt given by telling

a lie; these feelings would lead to a rapid increase in physiological activation. Because the feelings of discomfort and nervousness present in those who lie cause a physiological change, the scientific community has focused attention on observing and detecting these physiological indices.

The first machine made for the purpose of understanding whether a person was telling the truth or not was due to Vittorio Benussi (1914) and can in fact be considered an early version of the polygraph. He devised an instrument that recorded physiological changes such as skin conductance, blood pressure, heart rate and breathing pattern contextually to the test or interrogation to which the individual was subjected. As much as the polygraph gives an accurate measurement of the individual's arousal, it shows numerous limitations: in fact, although the correlation between stressful event and physiological hyperactivation is scientifically proven, it is not necessarily the case that the stressful event is necessarily having lied; a person could become physiologically activated precisely because he or she is being tested and not because he or she is telling lies. The polygraph, therefore, does not measure a direct correlation between lying and arousal but can only measure and detect the physiological activation of an individual who is subjected to a stressor; moreover, it is easily influenced by various subjective aspects that make it hardly generalizable. It should be noted that there are several interview protocols that are applied to the polygraph technique, each with limitations and strengths. One widely used protocol is the so-called "Relevant-Irrelevant Control Polygraph Test" (RIT): the individual is asked irrelevant questions-that is, control questions, which serve as a baseline-and relevant questions-that is, questions related to the crime being investigated. In this way, the irrelevant questions provide the baseline physiological activity, which is compared with the physiological readings recorded during the relevant questions: those who tell the truth should show the

same physiological activation for both types of questions; those who lie should show increased physiological activation in the relevant questions, i.e., those related to the crime in question, since - as we discussed earlier - nervousness and fear of being caught lead to intensified arousal. Another important protocol is the "Control Question Polygraph Test" (CQT) and is the most used polygraph implementation in the world. Unlike the RIT protocol, by employing different baseline questions concerning the ethics and humanity of the person being tested, truth tellers are expected to show greater physiological activation in control questions precisely because they are concerned about the interviewer's judgment. Conversely, those who lie will show greater arousal in response to relevant questions, as they correlate with the likelihood of being caught and thus blamed. Both protocols show important limitations; in the former, it is complex to be able to really distinguish between those who are telling the truth and those who are lying; moreover, those who tell the falsehood might be able to control their physiological activation during the testing phase. Even in CQT the rationale turns out to be weak in some places: just as truth tellers might show signs of hyperarousal due to fear of not being believed, liars do not necessarily show physiological hyperarousal in relevant questions. Otter-Henderson, Honts and Amato (2002) on the other hand have pointed out that polygraphs produce a number of false positives and false negatives, negatively affecting their accuracy. Because of this, the polygraph technique has often been criticized within the scientific community, so much so that it is scarcely used today, as well as prohibited in certain countries (Vrij, 2008).

With the advent of new technologies in the field of neuroscience, additional, increasingly precise and sophisticated techniques have been developed to find associations between brain components and lying responses. One of these techniques is thermography (TT), an

instrument that through infrared cameras scans and analyzes the subject's face: in this way, infrared radiation emitted by the individual can be detected, which is an index of the individual's temperature, and which changes as the subject tells the truth or lies. Another important technique is functional magnetic resonance imaging (fMRI), which allows blood flow in the brain to be recorded. It is based on the assumption that increased neural activity corresponds to increased blood flow, so it goes to see whether the areas where there has been significant activity are those used for lying or truth. Finally, an additional tool that can be used to identify the presence of a lie are event-related potentials (ERPs); through the placement of surface electrodes on the individual's scalp, these allow the recording of an electroencephalographic signal, or brain waves. Of all the detectable signals, the P300 wave is particularly important and has captured the interest of scholars; it, in fact, occurs in response to a stimulus that is considered by the subject to be relevant to him/herself, thus allowing one to understand whether for the individual being tested a certain interview question falls among the irrelevant or relevant stimuli. Through the discrepancy analysis of the magnitude of the evoked potential, the P300 methodology allows the identification of the subject who has the guilty knowledge, that is, the one who lies.

1.2.2 Behavioral Lie Detection techniques.

The behavioral techniques of lie detection are based on the measurement of reaction time (RT), that is, they calculate the time between the presentation of the stimulus and the individual's response. The motivation for such measurement lies in the fact that lying, as previously mentioned, requires considerable cognitive effort and this usually causes-at the behavioral level-a lengthening of response time in the individual (Walczyk et al.,

2003). Techniques using the reaction time analysis method are many, as numerous scientific studies have demonstrated its accuracy and validity (Debey et al., 2014). These include the Concealed Information Test (CIT) or Guilty Knowledge Test (GKT) - an interview protocol generally used in the use of polygraphs-which aims to differentiate between critical and irrelevant aspects of the crime in question, so that it is possible to discriminate between those who tell the truth-who will show indistinct physiological reactions-and those who lie-who will have different physiological activation in response to the relevant aspects related to the event. Another instrument that uses the reaction time technique is the Autobiographical Implicit Association Test (aIAT), a methodology constructed and validated in Italy by Giuseppe Sartori, a professor at the University of Padua, in 2008. This method is a modification of the well-known Implicit Association Test (IAT, Greenwald et al., 1998), a test that - by analyzing response times - tests whether there is an association between concepts. Similarly, the aIAT, relying on a principle called the "compatibility effect," tests for the existence of an autobiographical mnestic trace in the subject, i.e., in the presence of two conflicting hypotheses of truth - in the judicial context - it succeeds in identifying which version is the real one the subject has memory of. During the test, the subject must respond to sentences that appear on the computer regarding personal biographical events; he or she must mash the "true" or "false" keys depending on which version he or she believes is authentic. An algorithm analyzes the subject's reaction time and can determine whether the answer given is true or false.

The basic assumption is that the true memory has the fastest reaction times when matched with the authentic text. This tool has a very high accuracy rate, averaging 92 percent, is particularly flexible and capable of encoding even complex memories; moreover, it is an operator-independent tool, in that the results are not influenced by the examiner's

experience. However, there are some limitations, including the fact that for it to be used, memories must be sharp, not blurred, and negative phrases cannot be used.

1.2.3 The linguistic Lie Detection techniques.

In addition to psychophysiological and behavioral techniques, another way to assess the truthfulness of an account is to analyze its verbal content. Underlying this is the assumption that the account reported by a person who tells the truth is different from the account of a liar, both qualitatively and quantitatively. A widely validated tool in forensics that is based on this theory is the Criteria Based Content Analysis (CBCA), a methodology that assesses the quality of the subject's autobiographical recollection through 17 qualitative criteria that relate to the narrative. The evaluator examines the narrative, providing a 3-level rating: 0 = criterion absent, 1 = criterion partially present, 2 = criterion present. The sum of the scores provides an overall measure of how well the narrative exhibits the characteristics described above.

Studies from Sartori (2021) show that narratives from subjects who have had direct experience of the event tend to score higher on the CBCA, unlike those who "fabricate" the description (look for literature on Sarto's book). A method similar to the CBCA is Reality Monitoring (RM), devised by Johnson and Raye in 1981, which allows for highlighting structural differences between memories that result from direct or imagined perceptions and memories. The aspects that are analyzed to discriminate between real and false perceptions are: clarity of recollection, sensory aspects, spatial information, temporal information, emotional aspects, story reconstructability, realism, and cognitive operations. Both techniques can be regarded as systematizations of the evaluative criteriaology of intrinsic trustworthiness. Both RM and CBCA have a fairly high accuracy

rate-approximately 70 percent-but they are susceptible to countermeasures, e.g., those who are familiar with these methodologies can have their stories-while lying-rated as real.

CHAPTER 2: THE ALIBI AND ITS UNMASKING

2.1 Alibi: definition and areas of application.

In legal terms, an alibi is regarded as "a defense that puts the accused at the relevant time of the crime in a place other than at the relevant scene that makes it impossible for him to be the perpetrator" (Nolan, 1990). The alibi topic offers many interesting perspectives for psychological research, but scholars in the field have only recently become interested in the issue. When a questioned individual tells his or her version of the facts by uttering his or her alibi, the latter must then be subjected to a precise evaluation: first, the investigator must verify that the subject was in the space-time coordinates he or she indicated, and - subsequently - make sure that the individual could not have moved to the place of the crime in a certain time frame prior to the crime.

From a psychological perspective, it can be said that the process of establishing an alibi takes place within two domains, the domain of creation and the domain of credibility, which are in turn divided into two stages. The creation domain is concerned with the mnemonic processes of the subject and witnesses; in the first stage, the suspect's statement is collected, while in the second stage, details and information are sought that can confirm that alibi. The credibility domain refers to how individuals evaluate and ascertain the alibi. In the first phase-which takes place within the preliminary investigation-investigators and/or parties involved in the process make an initial assessment of the proposed alibi; the second phase involves the conclusion of the previous phase (Calabrò et al., 2010).

Confirming an alibi does not always follow a simple and straightforward process; sometimes it happens that-despite the many IT tools available-there is an inability to retrieve certain evidence and thus to ascertain the truthfulness of what the suspect has stated. Cooperation between different figures who can interact in a consistent and logical manner in order to use all available techniques and field resources, as well as always verify the reliability and accuracy of the tools used, is essential. In order that one can more easily discriminate between an actual alibi and a lying one, several techniques for unmasking the lie have been developed over time; in particular, interrogation and question administration techniques and response analysis techniques have been studied that were effective in ascertaining the true version of the facts. The study by Nieuwkamp et al. (2018) was instrumental in testing whether differences exist between false and real alibis and how accurate police investigators are in determining their veracity. The results that emerged showed how difficult it is to differentiate between true and false alibis and how the accuracy does not exceed 60%. In general, therefore, it can be said how the differentiation between true and false alibis is a complex issue.

2.2 Techniques for conducting interrogation to search for the truth.

Since it has been shown time and again how searching for the truth and categorizing an individual as an "honest" or "liar" is very complex and creates multiple problems, both practical and ethical, research-psychological and legal-has increasingly focused on the study of interrogation techniques, which can be considered as a means of deriving useful information (Caso et Palena, 2022). The goal to be pursued would be to consider both aspects - credibility and information gathering - in order to best seek the truth (Granhag, 2019). This direction would also be to be adopted in the various stages of investigation

and interrogation, where it is necessary to find as much information as possible in order to get a complete picture of the situation and what happened.

2.2.1 The SUE Technique.

The Strategic Use of Evidence (SUE) technique is a technique for conducting interrogation that goes in this direction, succeeding in meeting multiple needs, such as assessing the credibility and congruence between the evidence gathered and what the subject claims, finding as much information as possible, and obtaining a real confession from the suspect. It consists of two levels, one strategic-which is based on psychosocial and self-regulatory aspects, such as strategies and verbal responses of the interviewee- and one tactical-which is based on aspects of the specific event, such as interrogation planning, information gathering, and questions asked of the subject (Granhag et Hartwig, 2015). The SUE technique aims to detect congruencies and inconsistencies that may arise in the course of the interview, then leads the suspect to retract what-he believes the interviewers have as evidence against him. This end is pursued by showing the suspect the evidence against him that is present, so that the interviewers can receive more information and detect possible indications of lying through the discrepancies that have emerged.

Regarding this, the group led by Granhag experimented with three different ways of exhibiting evidence (Caso et Palena, 2022):

1. Late presentation: this is a method whereby the suspect is asked to make a statement with a free recall, that is, telling his or her version of events. Next, the interviewer starts asking him increasingly specific questions. Usually, the one who is lying would tend to use strategies of avoidance, i.e., omitting details, and

denial, e.g., denying that he or she was present at a certain place. Once both the free-recall narrative and the question-and-answer phase are concluded, the interviewer shows the suspect the evidence against him present.

2. Incremental presentation: a strategy designed with the goal of obtaining as much information as possible. Unlike the first mode, where evidence was presented only at the end of the interview, here it is shown to the suspect incrementally, that is, after a free recall followed by some questions, evidence is presented. Subsequently, additional questions and evidence are offered. In this way, the suspect may overestimate the amount of evidence against him or her that the authorities have, thus revealing information and facts that he or she thinks are already known to them, but in fact are not.
3. Tactical Use of Evidence (TUE): this method was proposed in 2011 by Dando and Bull and involves a brief free recall by the suspect, which is followed by specific questions from the interviewer. These questions focus on individual details of the case, and the interviewer with each response emphasizes the discrepancies present between the suspect's version and the evidence against him present.

2.2.2 The Scharff Technique.

An additional technique for conducting interrogation is the Scharff technique, conceptualized by Granhag's research group. Beginning with Hans Scharff's biography, in which the method by which he posed to the prisoners he was to interrogate is loosely described, the researchers set out to precisely delineate the interviewing techniques that emerged from these writings, with the aim of being able to study them experimentally.

After studying Scharff's biography in depth, researchers in Granhag's group highlighted the five central aspects of the Scharff approach:

1. Friendly approach: there is a need for a symmetrical relationship, in which the interviewer does not take a domineering role, but rather sets the conditions for the interviewee to feel accepted and free to speak;
2. Prevent strong pressure on the interviewee;
3. Behave in such a way as to appear to be already informed about everything: This way the interviewee will be more encouraged to speak up and reveal important details and information;
4. Refrain from asking direct questions: it is more optimal to seek confirmation or disconfirmation, whether verbal or paraverbal. In fact, the respondent will feel less guilt and perceive less guilt of having confessed if-instead of directly revealing facts-he only confirmed them;
5. Do not show interest in the additional information: the interviewer should not show interest in the new information provided by the interviewee, so as to make it appear that one already knows.

A recent meta-analysis conducted by Luke (2021) showed that the use of the Scharff technique is effective for multiple purposes, such as obtaining more information from the respondent and that the respondent, mistakenly thinking that the interviewer knows a lot, underestimates the extent of the new information he or she has confessed. Studies confirming its effectiveness are still limited, so it is good to interpret the results cautiously until more comprehensive demonstrations are available regarding the possible application of the Scharff technique (Caso et Palena, 2022).

2.3 The Cognitive Approach.

In light of the numerous research studies that have shown that the process of lying is extremely complex and requires significant cognitive effort on the part of the subject, scholars have begun to focus on the processes and executive components that come into play when a person wants to lie. Among the earliest studies related to the cognitive approach to lying must be mentioned the work published by Zuckerman, DePaulo and Rosenthal (1981), which showed that lying is considerably more strenuous than telling the truth. Indeed, there are numerous cognitive mechanisms that are activated when lying; these include attention, working memory, inhibition, and other executive functions that must be activated and integrated with each other for the lie to be consistent and believable (Gombos, 2006). Memory plays a relevant role in lying, both because the liar must remember all previous versions of his or her story and because-sometimes-memory errors can produce different accounts and be interpreted as a false narrative. A recent meta-analysis published by Vrij, Fisher, and Blank (2017) pointed out that the new aspect to be considered when discussing cognitive approaches to lying is the fact that investigators, who need to be able to discriminate between liars and nonliars, can amplify the differences - at the cognitive level - between them and, therefore, be better able to detect lying.

There are three main techniques that can highlight such differences: increasing cognitive load; encouraging the interviewee to provide additional information; and asking unexpected questions. It becomes clear that lying requires a particularly high cognitive effort compared to telling the truth. Having confirmed this hypothesis, the research was directed toward studying all the processes underlying lying and in particular set out to highlight the neural correlates involved in lying.

2.3.1 The cognitive processes involved in lying.

Lying involves a more cognitively demanding process than telling the truth (Christ et al., 2009). Indeed, when we lie, we do not simply retrieve an event from our memory as we do when we have to report a truthful fact. The cognitive process underlying the formulation of lies is not definable as a simple and unified process, but rather quite complex and consisting of several stages. According to Ganis and Keenan (2009) producing a deceptive response usually involves, first, making the decision to lie, after evaluating the entire context of the situation; then the individual will proceed with retrieving the information associated with honest and potentially deceptive responses, maintaining, and manipulating the relevant information retrieved from working memory. The response itself must then be maintained and encoded in memory to maintain its consistency over time, making it stable and durable in long-term memory. Finally, according to the authors, the liar will proceed with the inhibition of alternative deceptive responses and the honest response that is intended to be kept concealed. Should this inhibitory ability be lacking, this information could be retrieved and consequently betray the position of the liar, who would risk contradicting himself by letting some contradictory information slip out, if not even the whole truth. Gombos (2006) argues that during the procedure implemented to lie, therefore, two different processes are activated: the first one responsible for thought control, what we can and what we should not say, the second one for monitoring one's own credibility, analyzing the reactions of the interlocutor by modifying according to the latter one's behavior. Moreover, while lying, the individual will not only monitor the words used, but also pay attention to the whole sphere of nonverbal behavior, for example, gestures, facial expressions, posture and

intonation. Of course, all this makes the lying narrative more cognitively costly than the truthful narrative.

According to the cognitive approach, the various components of cognition and executive processes are crucial in the production and execution of lies that are also convincing. As executive processes we mean cognitive activities such as directed attention, metacognition, working memory, and finally inhibition of inappropriate information or responses in favor of convenient ones. These functions are crucial in planning and eliciting a lie (Walczyk et al., 2003). We can therefore conclude by saying that telling a lie, especially when it is to be made believable, requires numerous active cognitive mechanisms that must be successfully triggered and integrated.

2.3.2 The role of memory in lying.

As already stated, a crucial aspect in lying is memory.

Our memory is a limited storehouse of sources, information, facts, and experiences that we have lived - directly or indirectly - and consequently cannot be an exact snapshot of what actually happened. Indeed, when we recall a past event, it is highly likely that we include information that is not entirely accurate or inaccurate when recounting it, but this does not mean that we are lying. Often, inaccurate narratives or discordant versions are due precisely to mnemonic errors and not to an intention to lie. This is precisely why we need to be aware of the distortions that can result from memory errors; Schacter (2001) summarized the seven main "memory errors": lability, distraction, blockage, misattribution, suggestibility, distortion, and persistence. In addition to these, other important factors affecting memory are forgetting and false memories. Forgetfulness

refers to the process whereby, as the time elapsed since the crucial event increases, the memory fades or even disappears.

It was Ebbinghaus (1885) who explained - by means of a theory - the concept of forgetting: already in the minutes after the event, in the memorization phase, much information is lost; moreover, the more time elapses since the event, the more the accuracy of the memory decreases. This decay is initially very fast, then slows down as time progresses. Since forgetting occurs in any information storage process, whether of significant or non-significant events, it is important to be aware of this in order to better understand how discordance of narratives, or lack of details, is not necessarily a sign of falsehood, but may simply be the result of a natural process related to our memory. The term false memories refers to highly distorted memories of events that actually happened or even memories of events that never happened (Vannucci, 2008). Often these false memories appear to the subject's memory as real and vivid, even though they do not correspond to the actual extent of the events, or totally contrast with what happened. Numerous scientific researches have shown that that of false memories is a widespread phenomenon-although people are often convinced that they remember everything and are telling the truth-and it mostly affects autobiographical memory. In addition to forgetfulness and false memories, when evaluating a person's autobiographical recollection, one must consider many other aspects that can affect the subject's memory and that could lead-at first glance-to judging a truthful narrative as false.

The paradigm of post-event misleading information (Loftus et al., 1978) arose with the aim of testing how much post-event false information affected a testimony; it was seen that 60 percent of testimonial narratives were colored by wrong details after subjects had

received misleading information. Obviously, not all information has the same effect on testimony; there are two main variables that shape this phenomenon: the elapsed time interval-the effect of misleading information is higher the more time elapsed between it and the event (Loftus et al., 1978); and the source of the information-deception is more likely to occur if the information is rendered by a social source (Shaw et al., 1997).

Another important aspect is the so-called declarative contagion, also called memory conformity effect, which is the process whereby the subject reproduces--making it his own--a faulty memory produced by another subject who has previously recounted the same events. Studies have shown how the witness, exposed to misinformation from another witness, will tend to conform to what was said, producing an incorrect response to what happened. It is important to note that if a subject undergoes declarative contagion, he or she will then persist in the memory error (Sartori, 2021). All these memory errors have considerable importance when it comes to lying and must be taken into account and analyzed in order to be able to discriminate between genuine narrative and voluntary false narrative.

2.3.3 Cognitive Lie Detection Methods

We have already mentioned how it has been widely shown that liars exhibit, at baseline, a higher cognitive load. The cognitive approach to lie detection relies on the fact that investigators can amplify differences in the clues indicative of the cognitive load exhibited by truthers and liars through determinant cognitively demanding tasks. This approach encapsulates three techniques: imposing cognitive load, encouraging more information, and using unexpected questions (Vrij et al., 2017). In the first, experimenters mentally make the interview context more difficult: liars, who require more resources

than truth tellers, will have fewer cognitive resources at their disposal and may be less able than truth tellers to cope with these additional demands. Ways to increase cognitive load may be, for example, to ask respondents to tell their stories in reverse order (Evans et al., 2013) or to perform two tasks simultaneously (Visu-Petra et al., 2013). The second technique assumes that truthers provide more information, unlike liars, for whom it may be cognitively too difficult to add so much detail and, if they do, the additional information may be incorrect or resonate less plausibly. Moreover, liars might also be reluctant to add new information for fear of betraying themselves. Finally, the technique consisting of asking unexpected questions stems from the observation made through previous literature that has shown how liars use strategies to appear more truthful (Hartwig et al., 2007), including preparing answers to questions they are likely to expect to receive: planning does indeed make it easier to lie. However, such preparation carried out by liars has a limitation, which can be exploited by investigators. Vrij et al. (2009) were the first researchers to test the technique of unexpected questions by individually interviewing pairs of participants instructed to lie or tell the truth: from the experiments conducted, it could be found that up to 80 percent of the respondents managed to be correctly classified as liars or truthful.

When we talk about "unexpected question," we refer to spatial questions (e.g., "compared to the front door, where were you sitting?"), temporal questions (e.g., "which one of you finished before you ate?") and those concerning planning (e.g., "how far in advance did you plan your vacation?"). The difference between truth tellers and liars is elicited in the difficulty of answering such questions: the truth teller, in fact, should be able to easily retrieve the answer from his or her memory, unlike the liar, who is forced instead to have to create an ad hoc one, potentially also having difficulty remembering the answer given

in a later interview. Although liars may refuse to answer unexpected questions with statements such as "I don't know" or "I don't remember," such answers will create suspicion, especially if we are talking about simple questions or about central aspects of the event in question. When faced with unexpected questions, a liar has, therefore, little choice and is forced to improvise an answer that is not part of his original script, striving to make it plausible and mentally verifying that it cannot be unmasked. These complex mental operations result in an increased cognitive load, which is usually elicited in the lengthening of the response time, that is, the time between the end of the question and the beginning of the answer. In contrast, truth tellers experience similar levels of cognitive load while answering both expected and unexpected questions. Of course, the way in which truth tellers respond to unexpected questions may also depend on their familiarity with the topic on which they are asked to tell the falsehood: in fact, those who are more familiar with the topic of discussion may be able to handle unexpected questions better. This implies that the mere fact of being unexpected is not sufficient to make a question useful for lie detection. An unexpected question is effective when it produces an automatic response on the part of the truthful subject and a nonautomatic response on the part of the lying subject; the latter, moreover, should not be able to take refuge in non-remembering. For this to happen, all questions should address the central parts of the event rather than the peripheral parts (Vrij et al., 2017), and interviewers should be aware of both the response to expected and unexpected questions. The challenge of the unexpected question technique is, therefore, to be able to identify the right type of question to ask the liars in order to challenge them: for example, liars might be prepared to report on the details of the event about which they are lying, but not on what happened before or after that event (Warmelink et al., 2011).

2.3.4 The unexpected questions technique.

Over the past two decades, researchers have increasingly focused on studying additional aspects that could help in unmasking lies. Starting from the idea that lie planning makes the lying process easier, and at the same time gives interviewers fewer cues to be able to identify the lie (DePaulo et al., 2003), some studies have tried to use unexpected questions, i.e., questions that do not allow the respondent to mentally prepare an answer and his or her own version, because they are not predictable (Hartwig et al., 2007). Studies have shown how unexpected questions are often questions about the processes - and not so much the facts - of planning, thinking and organizing. The technique of unexpected questions is based on the fact that the truthful subject knows the unexpected information and retrieves it without particular difficulty, except for a physiological increase in latency in the response, whereas in the case of the lying subject the retrieval of this information must be quite difficult. In fact, it often happens that the liar responds - to unexpected questions - with expressions such as "I don't know," or "I don't remember." To remedy this problem, it is useful to ask the subject preliminary questions, related to the topic that the unexpected question will probe, so that one can retort if the respondent should later claim that he or she does not remember. Multiple experimental studies have shown that the technique of unexpected questions-focusing mostly on processes-are effective for the purpose of identifying a liar during an investigative interview. When implementing this technique, one must always verify that the questions thought to be unexpected are questions that the truthful subject can answer without great difficulty and, at the same time, require a very difficult response in the lying subject (Sartori, 2021).

Analyses conducted in a recent study (Monaro et al., 2017) discriminated, with 65-67 percent accuracy, honest people from liars who answered expected questions. In response

to unexpected questions, the discrimination accuracy between the two groups rose to 95 percent, a finding that reveals how unexpected questions can better expose those who lie. Two additional studies (Monaro et al., 2018, 2019) confirmed the effectiveness of unexpected questions in differentiating honest from lying subjects, in particular they showed that response times in liars are longer than response times from truth tellers.

2.4 Analysis of response latencies as an unmasking technique.

The term "response latency" or "reaction time" (RT) usually refers to the time interval between the onset of a stimulus and the corresponding response. Response times are often used to reveal processes that people may not be aware of. For this reason, analysis and measurement of response latencies is a potentially very reliable method for lie detection, since it exploits the information processing system and immediate stimulus-response behavior, which is very difficult to falsify.

Cognitive theory on lie detection highlights how it takes on average a longer time to formulate a deliberately false response than a truthful response: a previous meta-analysis (Suchotzki et al., 2017) showed, in fact, how the results of several studies converge that response times for false responses tend to be slightly longer than those for true responses, regardless of factors such as the complexity of the cognitive task or the response method (Seymour et al., 2000; Vendemia et al., 2005; Walczyk et al., 2009; Sheridan et Flowers, 2010). The differences found in response times can, therefore, be used to identify liars from truthful: this discrepancy can be explained since false responses include both the truthful response and an additional delay justified by reversing the original response or creating a new one. Thus, it can be said that lying adds a constant delay to true responses, although the reason why this occurs is still controversial. The most frequently advanced

notion is that lying increases the cognitive load of respondents: producing a statement as quickly as possible that is far from the truth inevitably requires cognitively demanding processes such as inhibiting the truthful response and deliberately constructing a new response that is consistent with existing facts and plausibly substitutes for the truth. In addition, all statements following the lie will have to be monitored to ensure the consistency of the story, and if more creative or complicated responses are required, the time taken will be progressively longer and reflect the processing load required.

Underlying the increased cognitive cost caused by the lying response, three executive functions have been proposed that could contribute to this effect: working memory, response inhibition, and task switching. In support of the hypothesis regarding working memory, it has been shown that, during lying, truth must be activated first, consequently elevating reaction times related to lying (Debey et al., 2014). Research has also shown that concurrent activation of the truth-telling and lying response results in a conflict between the two responses (Dong et al., 2010). Finally, the transition between telling the truth and telling falsehood, like the transition between different tasks, could be a cause of increased response latencies (Debey et al., 2015).

The results of studies based on reaction times have always been mixed, leading to both very large and small or nonsignificant effects. For example, in the study by Mapala et al. (2017), the results found were different from previous literature: liars did, in fact, show shorter latencies during lying responses, unlike truthful people, who responded equally quickly to both sets of questions. The decrease in response latencies in the liars could be the result of them responding more quickly to appear more confident in an attempt to deceive the questioner: this decrease in response latencies could also be an unconscious

strategy. Interestingly, most of the literature on lie detection has shown longer response latencies, but only when lying about arbitrary stimuli and using open-ended responses. These results highlight how context and question type make a difference.

Early attempts to distinguish between truthful and deliberately false responses on the basis of reaction time date back to the turn of the last century and made use of Jung's (1910) association-reaction method. Later, with the increased use of computerized measurements, the measurement of reaction times became increasingly easy, common and accurate. Verschuere et al. (2015) argue that a meaningful measurement of reaction time must meet certain criteria. First, the measurement must be accurate and reliable, that is, computerized. Secondly, participants must be able to respond immediately after the stimulus presentation and be instructed to respond as quickly as possible. For example, in the study by Ambach et al. (2011), participants were asked to respond to questions about a crime after a 4-second delay, which allowed subjects ample time to process the stimuli, prepare the required response, and strategically control task performance. Finally, to have a meaningful measurement, response times should be averaged over multiple measurements, with a proposed minimum of about 20 measurements per condition. In summary, then, the prerequisites for any meaningful measurement of response times as an indication of deception would seem to be computerized measurement, the instruction to respond as quickly as possible, and the use of multiple observations: the absence of these conditions in many studies of deception may explain why response latencies always fail to reliably distinguish liars from truthful.

CHAPTER 3: EXPERIMENTAL RESEARCH

3.1 Project description and research objectives.

The present study was conducted within the Department of General Psychology (D.P.G) of the University of Padua, in collaboration with the Department of Human Sciences of the University of Rome LUMSA.

The main objective of the research was to confirm and validate a lie detection technique, based specifically on the analysis of response latencies used to discriminate between lying and truthful subjects. In line with research conducted in the literature on lying (Suchotzki et al., 2017), one expects a lengthening of response times in subjects instructed to respond falsely. In fact, the lying subject, having no time to plan the lie, at the moment he or she improvises will have to inhibit the truthful response, replace it with the false response, and mentally verify that the lie cannot be easily exposed. These complex mental operations result in a high cognitive load that usually emerges leading to longer response times.

3.2 Materials and method.

This section will report the method, procedure and instruments used in the research.

3.2.1 Participants.

The total sample of participants in the study consisted of 60 subjects who were native speakers of Italian and resided mainly in Veneto, but also in the regions of Emilia-Romagna and Lombardy. The only inclusion criterion was to fall within an age range of

18 to 60 years; specifically, 15 males (25%) and 45 females (75%) aged between 18 and 60 years (M=29.23; SD=13.09) and of schooling between 8 and 17 years (M=14.58; SD=2.37) took part in the research. Participants were recruited through investigators' knowledge and word of mouth. In addition, subjects were randomly assigned to two different conditions by being divided into two groups: 30 participants belonged to the "truthful" condition, while the remaining 30 belonged to the "lying" condition. The participants, before starting the experiment, received assurances about the anonymity of their data, which were used only for the research purposes of the present study. Each of them viewed the informed consent and, by signing it, voluntarily decided to take part in the research.

3.2.2 Procedure and methodology.

After consenting to participate in the research by filling out an informed consent, participants were offered an interview of 34 open-ended questions - shown in the appendix - always administered in-person and lasting a total of about 10 minutes per subject. The interview was conducted in the same way for all participants.

First, before starting with the questions, the following delivery was read to each subject, "I will ask you questions about the last summer vacation you took." Next, if a given subject had been assigned the condition of "liar," it proceeded by requiring the participant to invent a vacation that never actually happened and to lie to all questions asked about it.

The first four questions, termed "control questions," are certain questions whose answer is known, and are the only ones that even those instructed to lie had to answer truthfully.

The next questions, on the other hand, referred to as "target" questions, are about the last vacation taken by the participant: among them, some are expected questions that can be predicted by the subject (e.g., "Where did you go on vacation?"), while others are unexpected questions (e.g., "How come you chose that destination?"), which mostly concern the planning of the vacation or the subjective experience the individual had. Finally, four questions in a 7-point Likert format were also included in the interview, with which we wanted to investigate the participants' perceived motivation, commitment, and difficulty.

The interview was audio-recorded, which allowed us to extract and analyze reaction times, that is, the latency times between the presentation of the stimulus (the experimenter's question) and the subject's response to it. The individual reaction times that emerged were coded using two different methodologies in order to explore the discriminative ability of both. On the one hand, the latency between the end of the experimenter's question and the participant's first verbalization, even in the absence of semantic meaning (e.g., "um"), was considered; on the other hand, the latency between the end of the experimenter's question and the participant's first word related to a sentence with semantic meaning was considered.

Finally, four types of comparisons were made:

A. The difference in latency times between honest and liars when comparing baseline and all target questions;
B. The difference in latency times between honest and liars when comparing baseline questions and "expected" target questions;

C. The difference in latency times between honest and liars when comparing baseline questions and "unexpected" target questions;
D. The difference in latency times between honest and liars when comparing "expected" and "unexpected" target questions.

3.2.3 Instruments.

In order to calculate the sample size needed to conduct the study, an analysis was conducted using GPower software. Two independent variables were considered, including one between (truthfulness: honest vs. dishonest) and one within (expected vs. unexpected questions), an α at 5% and a power at 0.95. Relative to effect size, two recent meta-analyses found a difference between honest and dishonest on reaction times of d = 1.45 (Verschuere et al., 2018) and d = 1.04 (Suchotzki et al., 2017), respectively. Nevertheless, a conservative choice was made, and an average expected effect size was set (Cohen's f = 0.25). These analyses returned a minimum sample size of 54 subjects.

As for the recordings, they were made through the "Voice Memos" application on the Apple brand cell phone. Subsequently, the study of reaction times was done by extracting and analyzing the coded response latencies using Audacity software.

3.3 Research Hypothesis.

From the research conducted in the literature on lying, a lengthening of reaction times-with both methods of encoding latencies-is expected in the responses of subjects instructed to respond falsely in all comparisons conducted. In addition, we hypothesize an interaction effect whereby within-subjects differences between expected and unexpected questions are greater for the liar than for the truth teller. In fact, the lying

subject, having no time to plan the lie, at the moment he or she improvises, must inhibit the true answer, replace it with the false one, and mentally verify that the lie cannot be easily exposed. These complex mental operations result in a cognitive load that usually lengthens the response time, and this is especially true for unexpected questions. This indicator is very efficient during the investigative phase, as opposed to the trial phase since the subject who is "trained" in lying produces a narrative in many respects indistinguishable from the truthful one. Also for this reason, it has been essential to include unexpected questions within the interview, for example regarding the subject's mental and decision-making processes: such questions, which the perpetrator has no way of foreseeing and for which he or she will not be able to prepare an answer in advance, take the liar by surprise, who will be led to respond with answers such as "I don't know" or greatly lengthening the time it takes to answer.

CHAPTER 4: DATA ANALYSIS AND RESULTS

4.1 Presentation of the statistical analysis conducted.

As described in section 3.2.1, the 60 participants were randomized into the two conditions of "honest" and "dishonest" such that there were 30 participants for each of the two experimental conditions.

Because the interviews and latency time extractions were conducted by two different experimenters, to reduce the risk of the analyses being biased by coder subjectivity, one of the two experimenters coded all 60 recordings, while the other coded only 50 percent of them. Next, an inter-rater agreement analysis could be conducted in order to assess whether the codings performed by the two different coders were comparable. Having demonstrated such comparability, the data were analyzed based on the codings of only one experimenter.

Through a linear mixed model, implemented through Jamovi software and specifically through the Gamlj package (Gallucci, 2019), two different analyses were conducted. The procedure of linear mixed models allows for the expansion of the general linear model, i.e., a linear regression analysis in which factors- called independent variables-are assumed to have a linear relationship with respect to the dependent variable. The type of mixed model provides flexibility in modeling the data, both in terms of their means, variances and covariances: it is characterized, therefore, as a flexible tool, capable of fitting other models that can be repurposed as linear mixed models (IBM, 2018).

4.2 First statistical analysis.

A first analysis investigated whether there were significant differences between response times to baseline and target questions, without differentiating between expected and unexpected target questions. An important finding emerges from Table 1 below, namely the R-squared Conditional, a statistical measure that represents the variance for a dependent variable explained by one or more independent variables in a regression model. In our case, this measure indicates that 28% of the variance of reaction times is explained by the independent variables.

Info	
Estimate	Linear mixed model fit by REML
Call	
AIC	10808.30
BIC	10895.01
LogLikel.	-5406.56
R-squared Marginal	0.12
R-squared Conditional	0.28
Converged	yes
Optimizer	bobyqa

Table 1: *Information regarding the implemented statistical model*

To understand which independent variables can explain the 28% of the variance in response times, we need to refer to Table 2 below. From it, it can be seen that truthfulness is associated with the variance of the dependent variable, i.e., dishonest people differ in response time from honest people significantly (p-value < 0.05). Similarly, the independent variable "question type" can be interpreted: as this effect is statistically significant (p-value < 0.05), we conclude that differentiation between baseline and target

questions can explain differences in response times. Predictably, the type of coding used to encode response latencies is also statistically significant, in that if semantically significant words are considered as the starting point of reaction time, the latency will surely be longer than in the condition in which verbalizations without semantic meaning (e.g., "um," "then") are also considered. The last relevant effect that emerges from the table below is the one representing the interaction between truthfulness and question type: in fact, considering first only honest subjects and then only those instructed to lie, significant differences between baseline and target questions are shown (p-value < 0.05).

	F	Num df	Den df	p
Truthfulness	4.32	1	67.61	0.042
Question type (1)	311.06	1	3534.00	< .001
Coding type	12.87	1	3534.00	< .001
Truthfulness * question type (1)	19.33	1	3534.00	< .001
Truthfulness * coding type	0.62	1	3534.00	0.432
Question type (1) * coding type	12.86	1	3534.00	< .001
Truthfulness * question type (1) * coding type	0.62	1	3534.00	0.431

Table 2: *Fixed Effects*

After understanding which independent variables have a significant effect and which have a linear relationship with the dependent variable, it is useful to go to see how and how much effect they have on reaction times. From Table 3 it is possible to obtain the unstandardized estimates, that is, the mean differences on reaction times, of the independent variables. In addition, it is possible to obtain-for these significant variables-the size of their effect on reaction times. Since there is no agreement in the literature on how to measure effect size in a mixed model such as ours, we adopted a conversion

formula that allows, given t and p, to obtain an estimate of the order of magnitude of an effect according to Cohen's classification (d)[1]. This classification considers an effect small when d = 0.2, medium when d = 0.5 and large when d = 0.8. Referring to the variable truthfulness and looking at the estimation column, we can see that the average difference between honest and dishonest is 0.29 seconds: although this seems like a small difference, the effect size associated with the variable tells us that we can actually consider this average effect (d = 0.51). Question type also has a significant effect on reaction time: estimates reveal that on average the difference between target and baseline questions is 0.92 seconds, and the effect is average (d = 0.59). Finally, the interaction analysis between truthfulness and question type produced a significant result, with an average difference of 0.46 seconds, but the effect size in this case is small (d = 0.15).

				95% Confidence Interval				
Names	**Effect**	**Estimate**	**SE**	**Lower**	**Upper**	**df**	**t**	**p**
(Intercept)	(Intercept)	0.84	0.07	0.71	0.96	67.61	11.96	< .001
Truthfulness1	dishonest-honest	0.29	0.14	0.01	0.56	67.61	2.08	0.042
Question type (1)1	target(unex+exp) - baseline	0.92	0.05	0.81	1.02	3534.00	17.64	< .001
Coding type1	fillers - semantic word	-0.19	0.05	-0.29	-0.09	3534.00	-3.59	< .001
Truthfulness * question type(1)1	bugiardo - onesto ✻ target(unex+exp) - baseline	0.46	0.10	0.25	0.66	3534.00	4.40	< .001
Truthfulness * coding type1	bugiardo - onesto ✻ fillers - semantic word	-0.08	0.10	-0.28	0.13	3534.00	-0.79	0.432
Question type(1)1 * coding type1	target(unex+exp) - baseline ✻ fillers - semantic word	-0.37	0.10	-0.55	-0.19	3534.00	-3.59	< .001
Truthfulness1 * question type (1)1 * coding type1	bugiardo - onesto ✻ target(unex+exp) - baseline ✻ fillers - semantic word	-0.16	0.21	-0.54	0.26	3534.00	-0.79	0.431

Table 3: *Fixed Effects and Estimates*

[1] The conversion formula and its results are given in the appendix.

Random effects were also included within the statistical analysis. The table below (Table 4) confirms for us the usefulness of considering these components: in fact, 25 percent of the variance in reaction times results from individual participants' subjectivity in response timing. The intraclass correlation coefficient (ICC = 0.18) – a reliability parameter that estimates the proportion of total variability due to between-subject variability – also justifies the use of the random effect.

Groups	Name	SD	Variance	ICC
Subject	(Intercept)	0.50	0.25	0.18
Residual		1.06	1.12	

Table 4: *Random Effect*

The next table (Table 5) shows the results from comparing target and baseline question types, again in relation to response times. The effect of these variables was shown to be significant (p-value < 0.05). Analyses showed an average difference in reaction times of 0.92 seconds: specifically, participants were faster – by an average of 0.92 seconds – in responding to baseline questions, while they showed longer reaction times to target questions. This result is also confirmed by a more than average effect size (d=0.59).

Comparison							
Question type(1)		Question type(1)	Difference	SE	t	df	$p_{bonferroni}$
baseline	-	target(unex+exp)	-0.92	0.05	-17.64	3534.00	< .001

Table 5: *Post Hoc Comparisons*

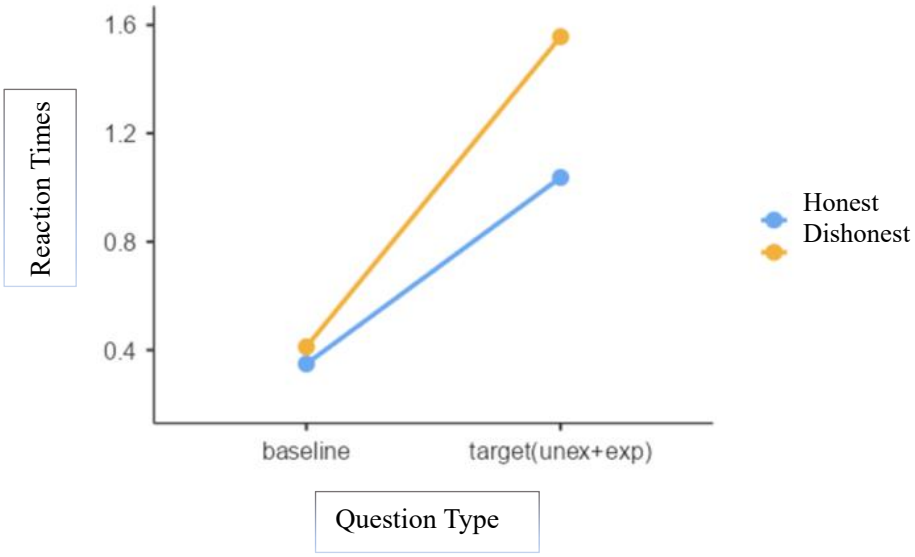
The last table (Table 6) reports what emerged in the comparison between honest and dishonest participants regarding differences in reaction times in response to baseline and target questions. In both groups, the difference in latency times is significant (p-value < 0.05), although looking at the unstandardized estimates it can be seen that the difference for the honest (0.69) is quite a bit smaller than that emerging in the dishonest (1.15). This is also confirmed by the effect size: in the case of the honest group the effect size is small (d = 0.32), while for those in the dishonest group it is medium (d = 0.52).

Moderator levels				95% Confidence Interval				
veridicità	contrast	Estimate	SE	Lower	Upper	df	t	p
Honest	target(unex+exp) - baseline	0.69	0.07	0.54	0.83	3534.00	9.36	< .001
Dishonest	target(unex+exp) - baseline	1.15	0.07	1.00	1.29	3534.00	15.58	< .001

***Table 6**: Estimates of the effect of question type*

To conclude, the graph below visually represents what we found through statistical analysis. In the baseline condition, the difference in the response times of honest and dishonest people is almost absent. In contrast, in the target condition, the response times are significantly farther apart, and it is precisely for this reason that this represents the most optimal condition for discriminating between honest and dishonest.

Figure 1: *Effects Plot*



4.2 Second statistical analysis.

In the second statistical analysis conducted, three levels were differentiated: it was investigated whether there were significant differences between the response times to the baseline and target questions, but in this case differentiating between expected and unexpected target questions. The comparison that was of most interest within the present study was that between baseline and expected target questions and that between expected and unexpected target questions.

The R-squared Conditional shows how the percentage of variance explained by the independent variables (29%) is very similar to that seen in the first statistical analysis conducted (28%).

Info	
Estimate	Linear mixed model fit by REML
Call	
AIC	10761.81
BIC	10886.62
LogLikel.	-5385.99
R-squared Marginal	0.13
R-squared Conditional	0.29
Converged	yes
Optimizer	bobyqa

Table 7: *Information regarding the implemented statistical model*

Table 8 below shows the independent variables that showed significant effects (p-value < 0.05) on reaction times.

	F	Num df	Den df	p
Truthfulness	7.42	1	60.85	0.008
Question(2)	183.54	2	3530.00	< .001
Coding type	38.48	1	3530.00	< .001
Truthfulness * question(2)	10.53	2	3530.00	< .001
Truthfulness * coding type	1.81	1	3530.00	0.179
Question(2) * coding type	7.39	2	3530.00	< .001
Truthfulness * question(2) * coding type	0.32	2	3530.00	0.723

Table 8: *Fixed Effects*

Taking in exam Table 9 next, the "main effects" of the different independent variables on reaction times can be studied by going to analyze the unstandardized estimates. Furthermore, as in the first statistical analysis, it was possible to derive-for the significant

variables-the effect size (effect size) that these variables exerted on reaction times according to Cohen's classification (d). This considers an effect small when d = 0.2, medium when d = 0.5 and large when d = 0.8. First, it can be seen from the table 9 that subject truthfulness is significantly and even slightly more associated with reaction time than the first analysis conducted, showing a medium-large effect size (d = 0.7). With regard to question type as well, concerning both the comparison between baseline and expected target question and between expected and unexpected target question, we note a statistically significant effect (p-value < 0.05) on reaction time, although the effect sizes found are not large (d = 0.4; d = 0.2). Predictably, the type of coding used-i.e., response onset given by semantically meaningful words or only semantically meaningless verbalizations-to encode response latencies also showed statistical significance. Finally, the last relevant effect that emerges is that representing the interaction between veracity and question type, considering the difference between baseline and expected targets (p-value < 0.05): the effect size in this case is small (d = 0.1). Finally, the interaction between the type of coding used and the truthfulness of the subjects was not shown to be significant: thus, regardless of the type of coding used in the analysis, the discrimination between honest and dishonest is similar.

				95% Confidence Interval				
Names	Effect	Estimate	SE	Lower	Upper	df	t	p
(intercept)	(intercept)	1.01	0.07	0.86	1.13	60.84	14.72	< .001
Truthfulness1	dishonest-honest	0.37	0.14	0.07	0.65	60.84	2.73	0.008
Question(2)1	baseline-expected	-0.80	0.05	-0.91	-0.69	3530.00	-14.81	< .001
Question(2)2	expected-unexpected	-0.27	0.04	-0.35	-0.20	3530.00	-7.18	< .001
Coding type1	fillers-semantic word	-0.25	0.04	-0.32	-0.17	3530.00	-6.20	< .001
Truthfulness1 * question(2)1	Dishonest-honest * baseline-expected	-0.42	0.11	-0.63	-0.21	3530.00	-3.87	< .001
Truthfulness1 * question(2)2	Dishonest-honest * expected-unexpected	-0.09	0.08	-0.23	0.06	3530.00	-1.20	0.229
Truthfulness * coding type1	Dishonest-honest * fillers-semantic word	-0.11	0.08	-0.28	0.06	3530.00	-1.35	0.178

Table 9: *Fixed Effects and Estimates*

In the following table (Table 10) we can observe on three levels the possible comparisons between question types in relation to response times, without discerning honest from dishonest. In all three cases the comparisons were shown to be significant (p-value < 0.05): there is, therefore, a difference in reaction time between both baseline and expected questions, between baseline and unexpected questions, and finally between expected and unexpected target questions. The average difference in reaction times that was found to be most pronounced was between expected and unexpected target questions (-0.27; d = 0.2): specifically, subjects were faster in responding to expected target questions, while they showed longer reaction times to unexpected target questions.

Comparison							
Question(2)		Question(2)	Difference	SE	t	df	$p_{bonferroni}$
baseline	-	expected	-0.80	0.05	-14.81	3530.00	< .001
baseline	-	unexpected	-1.07	0.06	-19.16	3530.00	< .001
expected	-	unexpected	-0.27	0.04	-7.18	3530.00	< .001

Table 10: *Post Hoc Comparisons*

Random effects were also included within the statistical analysis. Table 11 explains and confirms the usefulness of their use: in fact, 25% of the variance in reaction times results from individual participants' subjectivity in response timing. The intraclass correlation coefficient (ICC = 0.19) also justifies its use.

Groups	Name	SD	Variance	ICC
Subject	(Intercept)	0.50	0.25	0.19
Residual		1.05	1.11	

Table 11: *Random Effect*

Table 12 shows two different comparisons conducted both within the honest subjects' condition and within the dishonest subjects condition. In both groups, there was a significant difference (p-value < 0.05) between both the baseline and expected target questions, and between the expected and unexpected target questions. However, looking at the relative column estimates and comparing honest and dishonest, there seems to be a more pronounced discrepancy when considering the difference between baseline and expected target questions (honest = -0.59; dishonest = -1.01). Finally, the calculated effect sizes were not large in either those who were honest (d = 0.2; d = 0.1) or the group of liars (d =0.4; d= 0.1).

Moderator levels				95% Confidence Interval				
Truthfulness	contrast	Estimate	SE	Lower	Upper	df	t	p
Honest	baseline - expected	-0.59	0.08	-0.74	-0.44	3530.00	-7.74	< .001
	expected - unexpected	-0.23	0.05	-0.33	-0.12	3530.00	-4.23	< .001
Dishonest	baseline - expected	-1.01	0.08	-1.16	-0.86	3530.00	-13.21	< .001
	expected - unexpected	-0.32	0.05	-0.43	-0.21	3530.00	-5.93	< .001

Table 12: *Estimates of the effect of question type*

Taking Table 13 into consideration, it is possible to observe the difference between the two groups of participants by focusing separately on the different types of questions: with regard to the baseline questions, the difference between honest and dishonest was not significant (p-value > 0.05), and this result is understandable in light of the fact that in both experimental conditions participants were asked to answer the baseline questions truthfully. In contrast, for both expected and unexpected target questions, the difference between honest and dishonest was shown to be significant (p-value < 0.05). The effect size calculated when comparing the two groups for both expected and unexpected target questions was shown to be large in both cases (d = 0.84; d = 0.94). Specifically, moreover,

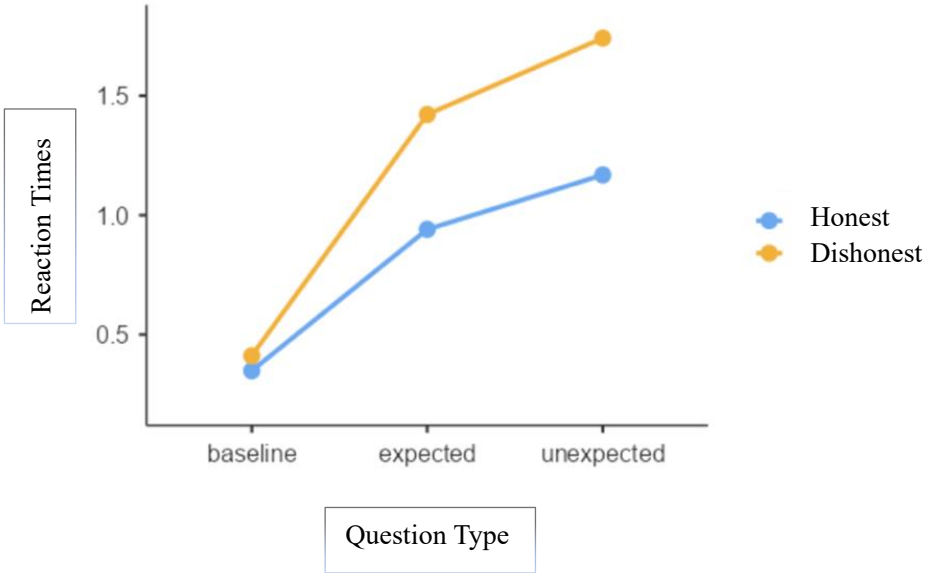
within the unexpected target questions condition, the difference seems to be slightly larger (0.57) than within the expected target questions condition (0.48). However, this comparison cannot be regarded as statistically certain, since looking at the confidence intervals of the two conditions "expected" and "unexpected" shows that they are, to some extent, overlapping.

Moderator levels				95% Confidence Interval				
dom(2)	contrast	Estimate	SE	Lower	Upper	df	t	p
baseline	dishonest-honest	0.06	0.16	-0.26	0.38	119.75	0.39	0.699
expected	dishonest-honest	0.48	0.14	0.20	0.76	66.10	3.45	< .001
unexpected	dishonest-honest	0.57	0.14	0.29	0.86	72.31	4.02	< .001

Table 13: *estimates of the effect of truthfulness*

In conclusion, the graph perfectly represents how in the baseline condition the distance between honest and lying subjects is practically absent. In the other two conditions, however, both expected and unexpected target questions manage to significantly differentiate honest from dishonest subjects. Moreover, in the case of the lying subjects, the reaction time is always higher than that found in the honest participants. Finally, the slope of the two straight lines is also indicative of how the difference between reaction time in the baseline condition and the expected and unexpected target questions conditions is greater in the dishonest group.

Figure 2: *Effects Plot*



4.4 Manipulation Check

As mentioned in Section 3.2.2, the experimental interview ended with four 7-point Likert scale questions called "Manipulation Check" questions, which are questions designed to investigate the participants' perceived commitment, motivation, and difficulty. The main objective was to investigate whether the scores obtained by the honest ones resembled those of the dishonest ones, or whether instead there was a gap due to the different experimental condition. To make this comparison, the means, standard deviations and resulting effect sizes (effects size) of each score obtained on the four final questions were calculated, differentiating between the two groups of participants. Specifically, for the first question, it was to investigate the participants' motivation to be credible: the honest group obtained quite different scores (M=5.6; SD=1.18) than the dishonest ones (M=6.1; SD=0.91), reporting an effect size of d=0.35, i.e., quite small. The same result was found for the engagement variable: in this case the honest ones had slightly lower scores (M=3.7; SD=1.48) than the dishonest ones (M=4.3; SD=1.91), so the former group reported to have engaged on average less than the latter (d=0.4). Finally, the final two questions discriminated between the two experimental conditions in a better way. In fact, in the third question, which asked participants how much they had repeated and/or gone over their story in their heads, the honest ones scored very differently (M=3.9; SD=1.75) than the liars (M=2.3; SD=1.38) and this was confirmed by a "giant" effect size (d=1). The interview concluded by asking participants how difficult they found the questions: this question also discriminated the two experimental groups quite a bit. In fact, the honest ones reported lower mean scores (M=3.2; SD=1.74) than the dishonest ones (M=4.5; SD=1.81). This result was also corroborated by a "large" effect size (d=0.8).

4.5 Machine Learning

In the final analysis, the accuracy of correctly classifying truth tellers and liars was evaluated using Machine Learning (ML) models. This machine learning model allows the construction of algorithms capable of distinguishing between those who simulate and those who tell the truth. As demonstrated in the pilot study conducted by Sartori and colleagues (2017), this model achieved an accuracy rate of over 90%. Additionally, Machine Learning models provide further analysis by predicting trends not only at a group level but also at an individual level. In the context of this experiment, this analysis can assist in the decision-making process to determine whether an individual is telling the truth or lying.

To conduct this final investigation, three characteristics were considered: the average response times to each type of question - control questions, expected target questions, and unexpected target questions - calculated for each participant. Specifically: C: average response times to control questions; A: average response times to expected target questions; I: average response times to unexpected target questions.

The implemented Machine Learning procedure is the K-fold cross-validation (K=10), where the data sample is divided into 10 distinct blocks. The algorithm is trained on the first 9 blocks, known as the training groups, to learn from the data and classify between honest and dishonest individuals. Subsequently, the algorithm is tested on the tenth block to evaluate its learning capability. This process is repeated 10 times, with the algorithm being trained and tested on different blocks each time. At the end of the procedure, various performance values are calculated for each algorithm, including accuracy, precision, area under the curve, recall, and F1 score. These values indicate the goodness percentage of

the model's performance. Table 14 shows the performance values resulting from the implementation of the three best performing algorithms.

Model	*Accuracy*	*AUC*	*Recall*	*Precision*	*F1*
Logistic Regression	0,72	0,68	0,72	0,70	0,69
Ridge Classifier	0,70	0,00	0,67	0,68	0,66
K Neighbors Classifier	0,69	0,78	0,72	0,67	0,67

Table 14: *Values of the investigated performance indices for the top three algorithms.*

Notes: *The table shows the values of the performance indices for the three best implemented algorithms. The indices are shown in the columns, while the algorithms to which they were applied are in the rows.*

Later, a tuning of the best-performing algorithm was implemented, which in this case turned out to be the Logistic Regression (LR) model. This operation involves retraining the algorithm, which was previously trained on the available data, in order to autonomously adjust and optimize parameters for better performance. As can be observed in the table below (Table 15), the tuning of the best model (LR) resulted in higher performance indicators, indicating optimization and improved quality.

Model	*Accuracy*	*AUC*	*Recall*	*Precision*	*F1*
Logistic Regression	0,76	0,72	0,72	0,75	0,72

Table 15: *Performance indicators after tuning the best model.*

Finally, the Feature Importance graph (Figure 3) represents the influence of each variable - in this case the mean input - on the model, allowing for a better understanding of which factors impact the predictions and providing information about the relationships between variables. In this case, the feature "A" – that is the average response times to expected target questions - appears to be the one that most significantly aids in discriminating between honest and dishonest individuals.

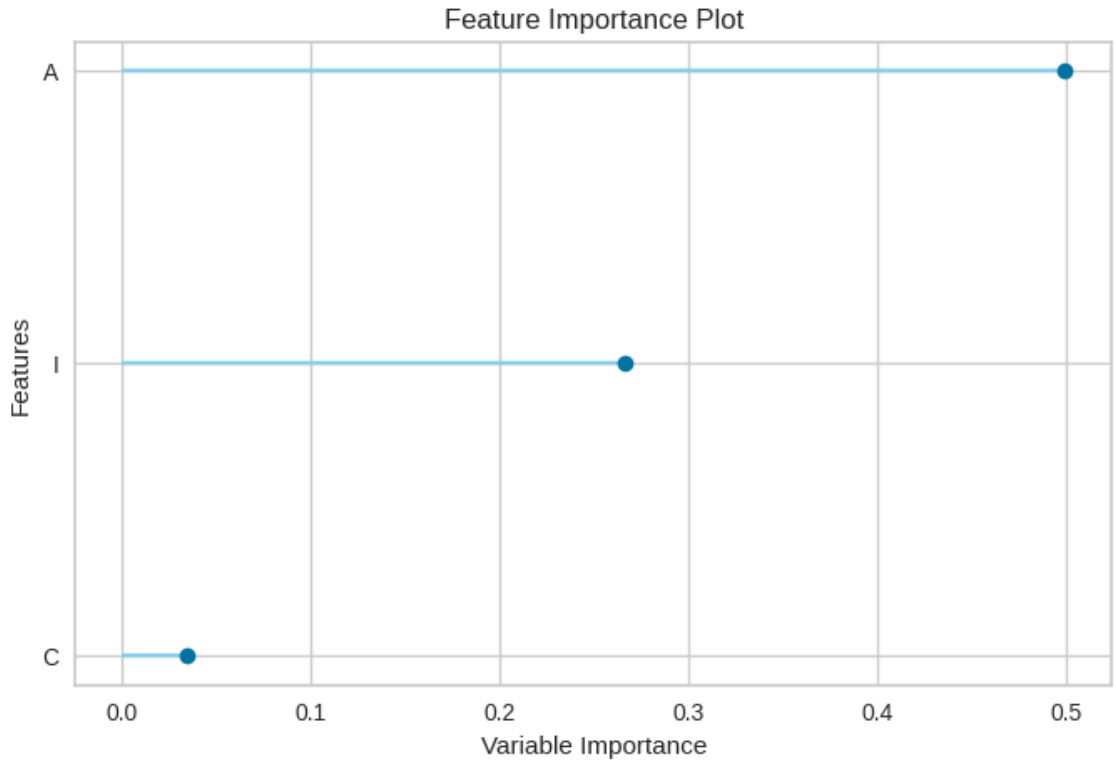


Figure 3: *Feature Importance graph*

CHAPTER 5: DISCUSSION OF RESULTS

In this last section, the objectives of the experimental paradigm, the hypotheses underlying the research and the methodologies used will be summarized; they will also be discussed in light of the results that emerged in Chapter 4, based on the present literature, the limitations that have emerged and the potential future directions that can be taken.

5.1 Research structure, objectives, and hypotheses.

The main objective of the research was to validate a lie detection technique based on the analysis of response latencies, the purpose of which was to discriminate between honest and dishonest individuals based on reaction times in response to questions posed by the interviewer. Consistent with research conducted in the literature in the area of lying (Suchotzki et al., 2017), the expectation is for longer reaction times in those instructed to lie. This delay is especially apparent in response to unexpected questions, i.e., those that the dishonest subject cannot foresee and for which he or she cannot prepare in advance a false answer that is consistent and plausible. Consequently, when faced with such unexpected questions, the liar will be forced to improvise an answer, inhibiting the truthful one and verifying-through complex mental operations-that the false one is consistent and difficult to expose. Such mental operations result in an increased cognitive load, which usually emerges with longer response times.

Given these assumptions, the experimental paradigm involved randomly dividing the 60 participants into two groups: the honest group and the dishonest group, the latter instructed to lie to the target-defined interview questions. The experiment included an interview investigating the last summer vacation taken by the subject: within it were 34 questions, specifically baseline questions, expected target questions, unexpected target questions and manipulation check questions. Participants assigned to the "honest" condition had to answer all the interview questions honestly, while the remaining participants, belonging to the "dishonest" condition, were asked to invent a fake vacation to answer the target questions, trying to come across as credible and consistent as possible. In line with previous literature, the hypotheses of the present experimental research were twofold: first, that for the dishonest there was a greater discrepancy in terms of reaction time between the baseline and target questions than for the honest individuals (H1), and second, that this discrepancy was more manifested by considering - within the questions defined as target - the unexpected ones (H2).

5.2 Discussion of the results.

The following subsections will discuss the results that emerged from both the statistical analysis conducted through a linear mixed model, the manipulation check questions conducted at the end of the interview, and finally the application of Machine Learning models.

5.2.1 Discussion of emerging results from the first statistical analysis.

Before proceeding to the first data analysis, as mentioned in Section 4.1, in order to reduce the probability that the results were compromised by a bias related to the subjectivity of

the two experimenters, an inter-rater agreement analysis was conducted. Having confirmed the comparability of the response time extractions by the two coders, the data from the sample intended as one was analyzed.

The implemented linear mixed model first investigated whether there were differences-between honest and dishonest-in response times to baseline and target questions. Given Table 2, which tells us which independent variables had a significant effect, i.e., those that have a linear reaction with our dependent variable "reaction times," it is useful to discuss each of them in light of the research hypotheses and literature already present. Table 3 (table of fixed effects and non-standardized estimates) shows that there is an average difference in reaction times between honest and liars; this finding confirms most of the results obtained from previous research (Suchotzki et al., 2017; Seymour et al., 2000; Vendemia et al., 2005; Walczyk et al., 2009; Sheridan et Flowers, 2010) which converge on the idea that reaction times for false responses tend to be slightly longer than those recorded for true responses. The significance of the independent variable "question type" shows us that there is an average difference of almost one second between reaction times in the two different groups of questions; it confirms, therefore, that it is necessary to look at the differences that exist, in terms of latency times, between baseline and target questions when trying to discriminate between honest and dishonest subjects. In fact, the control questions-which in this case consisted of questions regarding the participants' biographical data-act as a baseline and allow us to record the average time with which the subject responds to questions for which he or she knows the answer and for which he or she has no way of lying. This step is necessary to understand whether there is a difference in reaction times in the two different groups of questions and how it manifests itself in the honest and the liars. Also, in the same table (Table 3), another important finding

emerged, namely the significance of the interaction between truthfulness (honest-dishonest) and question type (baseline-target). The results indicate to us that there is an average difference of about half a second in the response latencies of honest and dishonest participants, specifically the honest participant group is on average faster to respond than the dishonest group. These data, in addition to confirming the present literature (Suchotzki et al., 2017; Seymour et al., 2000; Vendemia et al., 2005; Walczyk et al., 2009; Sheridan et Flowers, 2010), also confirm our first hypothesis (H1).

5.2.2 Discussion of emerging results from the second statistical analysis.

The second linear mixed model implemented investigated whether there were significant differences in response times between honest and dishonest respondents by using a three-level comparison: in fact, the target questions were in this case differentiated between expected target questions-those that were thought to have been predicted by the liars-and unexpected target questions, mostly concerning the mental and decision-making processes of the individual subject. The comparisons that were of most interest within the present study were, first, between baseline and expected target questions and, subsequently, between expected and unexpected target questions. Table 8 present in Section 4.3 shows the independent variables that had a significant effect on the dependent variable "reaction times": from here, it is possible to discuss each of these components in light of the research hypotheses and results that have already emerged within the past literature. First, it appears from Table 9 (Table of fixed effects and unstandardized estimates) that there is a significant mean difference between the reaction times of honest and dishonest people: this finding goes to confirm most of the results obtained from previous research in the literature (Suchotzki et al., 2017; Seymour et al., 2000; Vendemia

et al., 2005; Walczyk et al., 2009; Sheridan et Flowers, 2010), which agree that lying inevitably leads to longer response times. A significant effect on response time can also be inferred for the independent variable "question type"-regarding both the comparison between baseline and expected target questions and between expected and unexpected target questions-despite the fact that the effect sizes found are not very large (d = 0.4; d = 0.2). Specifically, also taking into consideration Table 10 (Table of Post Hoc Comparisons), it can be seen that within both experimental groups the most pronounced mean difference in reaction time was between expected and unexpected target questions (Difference = -0.27; d = 0.2): subjects, therefore, were faster in responding to expected target questions than to unexpected target questions. This result is shown to be in line with previous literature, in which the effectiveness of unexpected questions in discriminating between honest and dishonest has been confirmed (Monaro et al., 2017; 2018; 2019). The independent variable regarding the type of coding used to encode response latencies - the onset of responses between semantically meaningful words and verbalizations lacking semantic meaning were, in fact, discriminated - also showed statistical significance: this result could be considered predictable, since the response latency will surely be shorter in the condition in which verbalizations lacking semantic meaning (e.g., "um" or "then"), which are often uttered while the subject is mentally formulating a semantically meaningful response, are also considered. Within Table 9, another important finding is evident, namely the significance of the interaction between veracity (honest-disonest) and question type (baseline-target). Specifically, the analysis of Table 12 (Table of estimates of the effect of question type) allows us to study the different comparisons made within both the honest and dishonest subjects' condition. In both groups, a significant difference was shown between both baseline and expected

target questions, and between expected and unexpected target questions. However, looking at the relative column estimates and comparing honest and dishonest, there seems to be a more pronounced discrepancy when considering the difference between baseline and expected target questions (honest = -0.59; dishonest = -1.01). From these outcomes, therefore, it would seem more useful to compare baseline and expected target questions to effectively discern honest from dishonest: this result proved unexpected in comparison to the present literature (Monaro, 2018). Finally, regarding the interaction between the type of coding used and the truthfulness of the subjects, it was not found to be significant (Table 9): this result suggests that regardless of the type of coding used in the reaction time analysis, the discrimination between honest and dishonest is similar.

In conclusion, it is interesting to consider Table 13, in which the difference between the two groups of participants can be observed by focusing separately on the different types of questions. With regard to the response times to the baseline questions, it can be seen that the difference between honest and dishonest was not significant (p-value > 0.05): this result is understandable in light of the fact that in both experimental conditions the participants were asked to answer truthfully these types of questions whose answer was known to the experimenter (e.g., "What city are you currently in?"). In contrast, with regard to the response time to both expected and unexpected target questions, the difference between honest and dishonest emerged as significant (p-value < 0.05). In particular, within the unexpected target questions condition, the difference seems to be slightly more pronounced (0.57) than in the expected target questions condition (0.48), although this result cannot be considered statistically certain. The fact that it is precisely the unexpected target questions that are able to bring out the difference between truthful

and liars more effectively is absolutely in line with recent studies and findings in the literature (Monaro, 2018, 2019).

5.2.3 Discussion of emerging results from the Manipulation Check questions.

The experimental interview concluded with four 7-point Likert scale questions called "Manipulation Check" questions, which are questions designed to investigate the participants' perceived engagement, motivation, and difficulty. By calculating the averages, standard deviations and consequent effect sizes of each score obtained on the four questions, the two conditions of honest and dishonest could be differentiated and reflections could be made on them.

Having asked the dishonest experimental group to lie/simulating the tale of a vacation that never happened--without making promises to them or giving incentives of any kind, it was hypothesized that they might not have much motivation in answering the questions. However, with regard to the variable "motivation," the dishonest group scored higher on average (M=6.1) than the sincere group (M=5.6). The same result was also found for the variable "commitment": in this case, the honest group had slightly lower scores (M=3.7) than the dishonest group (M=4.3). The final two questions, however, discriminated between the two experimental conditions more effectively. The third question regarding how much the participants had repeated and/or gone over their story in their heads showed different results than what might have been expected: honest people scored higher on average (M=3.9) than liars (M=2.3). Finally, the interview concluded by asking participants how difficult they found the questions: this result was in line with what might have been expected and discriminated the two experimental groups quite a bit. In fact, the honest ones reported lower mean scores (M=3.2) than the dishonest ones (M=4.5),

making it clear that the greatest difficulty in dealing with the interview was experienced by the group of liars.

5.2.4 Discussion of emerging results from the Machine Learning analysis.

Lastly, to enhance the discrimination ability between honest and dishonest individuals, a Machine Learning analysis was employed as an additional tool. This analysis aimed to predict the individual's class (honest or dishonest) rather than solely considering the group level. The K-fold Cross validation procedure, consisting of 10 repetitions, was utilized to evaluate the performance of the best algorithm. Ultimately, the Logistic Regression (LR) model exhibited superior performance, yielding notable accuracy and precision indices (0.68-0.72), thereby demonstrating its proficiency in discerning between the two classes. To further optimize the model's performance, a tuning process was conducted, involving the retraining of the LR algorithm using the available data. This refinement resulted in improved performance metrics (0.72-0.76), particularly in terms of accuracy and precision indices. These findings reinforce the effectiveness of the model in effectively distinguishing participants in the "honest" and "dishonest" conditions.

Subsequently, the Feature Importance graph was generated, illustrating the discriminative capabilities of the various variables under investigation. Notably, the analysis revealed that the "A" variable exhibited the strongest discriminatory power between subjects in the honest and dishonest conditions. Surprisingly, this finding contradicted both most of the existing literature and the experimental paradigm's second hypothesis (H2), which had been partially supported by previous analyses. H2 suggested that unexpected questions would be more effective in distinguishing between honest and dishonest individuals. However, the graph demonstrated that a more fruitful and effective differentiation could

be achieved by considering the comparison between control questions and expected target questions.

5.3 Emerging limitations and future directions

Despite the overall positive results of the study, which largely confirmed the initial hypotheses, it is important to acknowledge the limitations and shortcomings encountered. Regarding the sample composition, it is noteworthy that there was a significant overrepresentation of female participants (75%). A future direction could involve selecting a more balanced sample, allowing for more generalized discussions of the results, and potentially exploring gender differences.

However, the majority of limitations identified in this study pertain to the interview administration phase. For instance, it was observed that participants in the honest experimental group took considerable time to recall details related to their vacation, often failing to retrieve them immediately. This extended response time for the sincere group exceeded expectations based on existing literature. Conversely, participants in the dishonest condition frequently provided generic or inaccurate responses, mentioning incorrect airports or time zones, and providing inconsistent and illogical answers. It is important to note, however, that participants in the dishonest group were compelled to fabricate a vacation that did not occur. It can be hypothesized that some of them lacked sufficient motivation to provide precise and accurate responses, as they had no incentive to do so.

Regarding the limitations observed in the interview administration for the dishonest participants, future studies could consider requesting more precise, detailed, and in-depth

responses. Additionally, it would be beneficial to implement an objective verification of the answers provided to ensure their reliability and consistency. Furthermore, a potential avenue for future research could involve providing compensation to dishonest participants, thereby creating a genuine motivational drive to lie consistently and believably. In real forensic contexts, individuals often resort to deception for various purposes, such as avoiding criminal charges, seeking monetary compensation, or preventing sole custody of their child. These goals motivate individuals to invest effort in constructing a coherent, logical, and credible story.

In addition to the mentioned limitations encountered during the administration phase, further challenges were identified upon analyzing the results of the conducted statistical analyses. Consistent with prior literature, the data analysis revealed that dishonest individuals generally took longer to respond to target questions compared to honest individuals. This increased response latency was particularly pronounced for unexpected questions, suggesting that unexpected questions were more effective in discriminating between the "honest" and "dishonest" experimental conditions. However, a subsequent analysis unexpectedly revealed that the greatest discrepancy in response time was observed when comparing control questions with expected target questions. Simultaneously, the machine learning analyses confirmed previous studies in the literature, highlighting the algorithm's effectiveness in effectively discriminating between honest and dishonest participants after being trained and tested on the provided experimental data. The Feature Importance analysis revealed that the variable representing the average response times to expected target questions exhibited the strongest discriminatory power between the honest and dishonest conditions.

In conclusion, it would be both interesting and beneficial to implement future modifications to the experimental paradigm, optimizing its application and outcome. This could involve expanding the participant sample, designing question types more rigorously, and exploring the motivations underlying the minor discrepancies observed in the results of the conducted statistical and machine learning analyses.

CONCLUSIONS

The current study aimed to evaluate the efficacy of a Lie Detection technique based on analyzing reaction times. Specifically, the study examined response latencies within an interview to determine whether they could effectively discriminate between honest and dishonest individuals. Analyzing reaction time as a means to detect deception has been widely utilized in previous research due to its reliance on the information processing system and immediate stimulus-response behavior. This technique serves as a difficult-to-falsify index and enables precise identification of dishonest subjects. Numerous studies in the literature (Suchotzki et al., 2017; Seymour et al., 2000; Vendemia et al., 2005; Walczyk et al., 2009; Sheridan & Flowers, 2010) have confirmed the efficacy of this technique, demonstrating that lying individuals generally exhibit longer response times compared to honest individuals. Subsequent studies (Monaro et al., 2017; Sartori, 2021) have further specified that this increase in reaction time among liars is particularly evident when they are posed unexpected questions, i.e., questions that are unpredictable and do not allow for premeditated fabrication that appears consistent and logical. Building on this existing literature, the present study formulated two main experimental hypotheses: firstly, that there would be a discrepancy in response times between honest and dishonest subjects when comparing baseline and target questions, and secondly, that this discrepancy would be more pronounced when comparing control and unexpected target questions.

The study involved a sample of 60 participants who were randomly assigned to two experimental conditions: the honest condition and the dishonest condition. The goal was

to ensure an equal distribution of 30 participants in each group. During the study, participants underwent an interview focused on their most recent summer vacation. The interview consisted of a total of 34 questions, including baseline questions, expected target questions, unexpected target questions, and manipulation check questions. In the honest condition, participants were instructed to provide truthful responses to all the interview questions. In contrast, participants in the dishonest condition were instructed to fabricate fictional vacations and provide answers to the target questions that appeared believable and consistent. Subsequently, the collected data were analyzed using a linear mixed statistical model, specifically a linear regression analysis. This statistical approach assumes a linear relationship between the independent variables, which in this case are factors related to the type of question, and the dependent variable, which is the participants' reaction times.

The results acquired showed that-consistent with most of the published literature-there is an average difference in reaction times between honest and liars; moreover, the significance of the independent variable "question type" shows that there is an average difference of almost one second between the reaction times of the two different groups of questions. This last result that emerged, therefore, underscores the need to focus on the differences that exist, in terms of response latency, between baseline and target questions when trying to discriminate between honest and dishonest subjects, and thus confirms the first hypothesis (H1) of the experimental paradigm. Further analyses conducted returned that in both experimental groups - honest and dishonest - the most pronounced mean difference in reaction time was between expected and unexpected target questions: subjects were faster in response to expected target questions than to unexpected target questions. This outcome was also in line with previous literature - within which the

effectiveness of unexpected questions in discriminating between honest and dishonest is confirmed - and partially confirms the second research hypothesis (H2). This second research hypothesis, however, was disconfirmed through the latest implemented analyses conducted using Machine Learning models. The best-performing algorithm, after considering all three variables - namely the average response times to control, expected, and unexpected questions - indicated that the Feature Importance, which represents the variable that best discriminates between the "honest" and "dishonest" groups, was variable "A." This variable corresponds to the average response times to expected questions. Despite expected questions belonging to the category of target questions, which generally provide better discrimination between honest and dishonest subjects compared to control questions, this result contradicts both previous literature and the results of the previously conducted statistical analyses.

In conclusion, although there were some limitations in the implementation and interpretation of the present experimental paradigm, it has yielded numerous important findings for expanding the literature on Lie Detection in forensic contexts. The ability to discriminate more effectively between truth-tellers and liars, for example, during interrogations, is an important objective to pursue, which appears increasingly attainable thanks to new technologies. As potential future directions, it would be interesting to collect a larger sample of participants and further refine the structure of both expected and unexpected questions to assess their discriminatory power more comprehensively.

BIBLIOGRAPHY

Abe, N. (2011). How the Brain Shapes Deception: An Integrated Review of the Literature. *The Neuroscientist*, *17*(5), 560–574.

Ambach, W., Dummel, S., Lüer, T., & Vaitl, D. (2011). Physiological responses in a Concealed Information Test are determined interactively by encoding procedure and questioning format. *International Journal of Psychophysiology*, *81*(3), 275–282.

Azizli, N., Atkinson, B. E., Baughman, H. M., Chin, K., Vernon, P. A., Harris, E., & Veselka, L. (2016). Lies and crimes: Dark Triad, misconduct, and high-stakes deception. *Personality and Individual Differences*, *89*, 34–39.

Beata, A., Cantarero, K., & Soroko, E. (2015). Motivation and Consequences of Lying. A Qualitative Analysis of Everyday Lying. *Forum Qualitative Sozialforschung*, *16*, preprint 31.

Bond Jr., C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*, 477–492.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior*, *18*(2), 155–184.

Buta, M., Visu-Petra, G., Koller, S. H., & Visu-Petra, L. (2020). A Little Lie Never Hurt Anyone: Attitudes toward Various Types of Lies over the Lifespan. *Psychology in Russia: State of the Art*, *13*(1), 70–81.

Calabrò, V., Costabile, G., Fratepietro, S., Ianulardo, M., Nicosia, G., - "L'alibi informatico. Aspetti tecnici e giuridici". Chapter in IISFA Memberbook (2010).

Caso, L., Gnisci, A., Vrij, A., & Mann, S. (2005). Processes Underlying Deception: An Empirical Analysis of Truth and Lies when Manipulating the Stakes. *Journal of Investigative Psychology and Offender Profiling*, *2*, 195–202.

Caso, L., & Palena, N. (2022). *Interrogare. Metodi e strategie per la raccolta delle informazioni e la valutazione della credibilità*. Il Mulino.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(7), 1557–1566.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, *115*(1), 11–15.

Costa, P., & Mccrae, R. (1992). Neo PI-R professional manual. *Psychological Assessment Resources*, *396*.

Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police officers to interview tactically. *Journal of Investigative Psychology and Offender Profiling, 8*(2), 189–202.

Debey, E., De Houwer, J., & Verschuere, B. (2014). Lying relies on the truth. *Cognition*, *132*(3), 324–334.

Debey, E., Liefooghe, B., De Houwer, J., & Verschuere, B. (2015). Lie, truth, lie: The role of task switching in a deception context. *Psychological Research*, *79*(3), 478–488.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*, 979–995.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

Dong, G., Hu, Y., Lu, Q., & Wu, H. (2010). The presentation order of cue and target matters in deception study. *Behavioral and Brain Functions*, *6*(1), 63.

Ebbinghaus, H. (1885). Memory: A contribution to experimental psychology (translated by HA ruger & CE bussenues, 1913). *New York: Teachers College, Columbia University*.

Ekman, P., & Friesen, W. V. (1969). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica*, *1*(1), 49–98.

Ekman, P., Rolls, E. T., Perrett, D. I., Ellis, H. D., Bruce, V., Cowey, A., Ellis, A. W., & Perrett, D. I. (1992). Facial expressions of emotion: An old controversy and new findings. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *335*(1273), 63–69.

Eshet, Y., Grinautsky, K., Peled, Y., & Barczyk, C. (2014). No More Excuses—Personality Traits and Academic Dishonesty in Online Courses. *Journal of Statistical Science and Application*, *2*.

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a Psychologically Based Credibility Assessment Tool. *Journal of Applied Research in Memory and Cognition*, *2*(1), 33–41.

Gallucci M. (2019). *GAMLj: General analyses for the linear model in jamovi* [Jamovi module].

Ganis, G., & Keenan, J. P. (2009). The cognitive neuroscience of deception. *Social Neuroscience*, *4*(6), 465–472.

Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, Social, and General Psychology Monographs*, *132*, 197–214.

Granhag, P. A., & Hartwig, M. (2015). The strategic use of evidence technique: A conceptual overview. In *Detecting deception: Current challenges and cognitive approaches* (pp. 231–251). Wiley-Blackwell.

Granhag, P. A., Hartwig, M., & Granhag, P. A. (2019). *Reading Lies: Nonverbal Communication and Deception* (SSRN Scholarly Paper Fasc. 3318207).

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (s.d.). *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*.

Gylfason, H. F., Halldorsson, F., & Kristinsson, K. (2016). Personality in Gneezy's cheap talk game: The interaction between Honesty-Humility and Extraversion in predicting deceptive behavior. *Personality and Individual Differences*, *96*, 222–226.

Hart, C. L., Lemon, R., Curtis, D. A., & Griffith, J. D. (2020). Personality Traits Associated with Various Forms of Lying. *Psychological Studies*, *65*(3), 239–246.

Hartwig, M., Anders Granhag, P., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*(2), 213–227.

Hartwig, M., Granhag, P. A., Stromwall, L. A., & Doering, N. (2010). Impression and Information Management: On the Strategic Self- Regulation of Innocent and Guilty Suspects. *The Open Criminology Journal*, *3*(1).

Hines, A., Colwell, K., Hiscock-Anisman, C., Garrett, E., Ansarra, R., & Montalvo, L. (s.d.). Impression management strategies of deceivers and honest reporters in an investigative interview. *European Journal of Psychology Applied to Legal Context*, *2*(1). Recuperato 15 marzo 2023, da https://journals.copmadrid.org/ejpalc/art/b096577e264d1ebd6b41041f392eec23

IBM Documentation. (2022, settembre 13). https://www.ibm.com/docs/it/spss-statistics/saas?topic=models-linear-mixed-model

Jung, C. G. (1910). The Association Method. *The American Journal of Psychology*, *21*(2), 219–269.

Levine, T. R., & McCornack, S. A. (2014). Theorizing about deception. *Journal of Language and Social Psychology*, *33*(4), 431–440.

Loftus, E., Miller, D., & Burns, H. (1978). Semantic Integration of Verbal Information into a Visual Memory. *Journal of experimental psychology. Human learning and memory*, *4*, 19–31.

Luke, T. J. (2021). A meta-analytic review of experimental tests of the interrogation technique of Hanns Joachim Scharff. *Applied Cognitive Psychology*, *35*(2), 360–373.

Mapala, T., Warmelink, L., & Linkenauger, S. (2017). Jumping the gun: Faster response latencies to deceptive questions in a realistic scenario. *Psychonomic Bulletin & Review*, *24*.

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, *45*(6), 633–644.

Mazza, C., Monaro, M., Orrù, G., Burla, F., Colasanti, M., Ferracuti, S., & Roma, P. (2019). Introducing machine learning to detect personality faking-good in a male sample: A new model based on Minnesota Multiphasic Personality Inventory-2 Restructured Form scales and reaction times. *Frontiers in Psychiatry, 10,* Article 389.

Monaro, M., Gamberini, L., & Sartori, G. (2017). The detection of faked identity using unexpected questions and mouse dynamics. *PLOS ONE*, *12*(5), e0177851.

Monaro, M., Toncini, A., Ferracuti, S., Tessari, G., Vaccaro, M. G., De Fazio, P., Pigato, G., Meneghel, T., Scarpazza, C., & Sartori, G. (2018). The Detection of Malingering: A New Tool to Identify Made-Up Depression. *Frontiers in Psychiatry*, *9*.

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The Malevolent Side of Human Nature: A Meta-Analysis and Critical Review of the Literature on the Dark Triad (Narcissism, Machiavellianism, and Psychopathy). *Perspectives on Psychological Science*, *12*, 183–204.

Nieuwkamp, R., Horselenberg, R., & van Koppen, P. (2018). True and false alibis among prisoners and their detection by police detectives. *Psychiatry, Psychology and Law*, *25*, 902–921.

Nolan. (1990). *Black's law dictionary: Definitions of the terms and phrases of American and English jurisprudence, ancient and modern* (6th ed. / by the publisher's editorial staff; coauthors, Joseph R. Nolan [and others], Centennial ed. (1891-1991).). West Pub. Co.

Otter-Henderson, K., Honts, C., & Amato, S. (2002). Spontaneous Countermeasures During Polygraph Examinations: An Apparent Exercise in Futility. *Polygraph*.

Saarni, C. (1984). An observational study of child-ren's attempts to monitor their expressive behavior. *Child Development*, *55*, 1504–1513.

Sartori, G., Zangrossi, A., Orrù, G., & Monaro, M. (2017). Detection of malingering in psychic
damage ascertainment. In P5 medicine and justice (pp. 330-341). Springer, Cham.

Sartori, G. (2021). La memoria del testimone: Dati scientifici utili a magistrati, avvocati e consulenti. Giuffrè Francis Lefebvre.

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers* (pp. x, 272). Houghton, Mifflin and Company.

Semrad, M., Scott-Parker, B., & Nagel, M. (2019). Personality traits of a good liar: A systematic review of the literature. *Personality and Individual Differences*, *147*, 306–316.

Seymour, T., Seifert, C., Shafto, M., & Mosmann, A. (2000). Using response time measures to assess "guilty knowledge." *The Journal of Applied Psychology*, *85*, 30–37.

Shaw, J. S., Garven, S., & Wood, J. M. (1997). Co-witness information can have immediate effects on eyewitness memory reports. *Law and Human Behavior*, *21*(5), 503–523.

Sheridan, M. R., & Flowers, K. A. (2010). Reaction Times and Deception—The Lying Constant. *International Journal of Psychological Studies*, *2*(2), p41.

Street, M., & Street, V. L. (2006). The Effects of Escalating Commitment on Ethical Decision-Making. *Journal of Business Ethics*, *64*(4), 343–356.

Strömwall, L., & Willén, R. (2011). Inside Criminal Minds: Offenders' Strategies when Lying. *Journal of Investigative Psychology and Offender Profiling*, *8*, 271–281.

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*, *in press*.

Tyler, J. M., Feldman, R. S., & Reichert, A. (2006). The price of deceptive behavior: Disliking and lying to people who lie to us. *Journal of Experimental Social Psychology*, *42*(1), 69–77.

Vannucci, M. (2008). *Quando la memoria ci inganna. La psicologia delle false memorie*. Carocci.

Vendemia, J. M. C., Buzan, R. F., & Simon-Dack, S. L. (2005). Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. *Behavioural Neurology*, *16*(1), 25–36.

Verschuere, B., Suchotzki, K., & Debey, E. (2015). Detecting deception through reaction times. In *Detecting deception: Current challenges and cognitive approaches* (pp. 269–291). Wiley-Blackwell.

Visu-Petra, G., Varga, M., Miclea, M., & Visu-Petra, L. (2013). When interference helps: Increasing executive load to facilitate deception detection in the Concealed Information Test. *Frontiers in Psychology*, *4*.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Wiley.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the Liars: The Benefit of Asking Unanticipated Questions. *Law and Human Behavior*, *33*(2), 159–166.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, *22*, 1–21.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, *17*(7), 755–774.

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive Lie Detection: Response Time and Consistency of Answers as Cues to Deception. *Journal of Business and Psychology*, *24*(1), 33–49.

Warmelink, L., Vrij, A., Mann, S., Leal, S., & Poletiek, F. (2011). The Effects of Unexpected Questions on Detecting Familiar and Unfamiliar Lies. *Psychiatry, Psychology and Law*, *20*, 1–7.

Williams, K. M., Nathanson, C., & Paulhus, D. L. (2010). Identifying and profiling scholastic cheaters: Their personality, cognitive ability, and motivation. *Journal of Experimental Psychology: Applied*, *16*, 293–307.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and Nonverbal Communication of Deception. In *Advances in Experimental Social Psychology* (Vol. 14, pp. 1–59). Elsevier.

Zuckerman, M., Driver, R., & Guadagno, N. S. (1985). Effects of segmentation patterns on the perception of deception. *Journal of Nonverbal Behavior*, *9*, 160–168.

APPENDIX

A – Interview

Task delivery: "I'm going to ask you some questions about the last vacation you took."

Follow-up questions:

1. What is your name?
2. How old are you?
3. Where were you born?
4. What city are you currently in?

Target questions:

5. Where did you go on vacation?
6. Who were you with?
7. What year and month did you leave, and if you also remember the day?
8. How far in advance did you arrange and book your vacation?
9. How come you chose that destination?
10. Approximately what time slot did you leave?
11. What was the weather when you left?
12. What luggage did you take with you?
13. What means of transportation did you use?
14. Transportation variants:

    IF car/bus: which exit did you take to get to your destination?

IF train: what station did you arrive at?

IF airplane: to which airport did you arrive?

IF ferry: to which port did you arrive?

15. How long did it take you to get to your destination?

16. When you arrived at your destination what was the temperature?

17. What was the time zone of your destination?

18. Where did you stay in your accommodation?

19. Accommodation variants:

IF in hotel/b&b: how was the room you slept in furnished?

IF in apartment/house: how was the room you slept in furnished?

IF in camping: describe the camper/tent you slept in

20. How many days did you stay?

21. What activities did you do during your vacation?

22. Variant activities:

IF culture vacation: what is the most beautiful thing you visited?

IF sea/mountain/lake...: what is the most beautiful natural landscape you saw?

23. What is the best memory you have of the vacation?

24. What is the thing you would change that did not go as you planned?

25. What emotions and feelings did the vacation arouse in you?

26. Describe a meal that you particularly remember from the vacation

27. Approximately how much budget did you spend on this vacation?

28. Is it a destination you would recommend and why?

29. Once the vacation was over what time slot did you return home?

30. What did you do immediately after returning?

Likert questions with 7-point scale:

At the end of the interview, "Well, thank you, the vacation interview is over.

We will now ask you a few more questions, which you must answer honestly."

1) How motivated were you to be credible, on a scale of 1 (not at all motivated) to 7 (very motivated)?
2) In the minutes you had to prepare for the interview, how much effort did you put into preparing your story, on a scale ranging from 1 (I did not put in any effort at all) to 7 (I put in a lot of effort)?
3) In the minutes you had to prepare for the interview, how much did you repeat and go over your story in your head, on a scale from 1 (not at all) to 7 (very much)?
4) How difficult did you find the interview, on a scale from 1 (not at all difficult) to 7 (very difficult)?

B - Formula and conversion table for effect size (d)

$$d = 2 * t/\sqrt{df_{error}}$$

	A	B	C
1	t	df_error	Cohen d
2	0,39	119,75	0,07127822
3	3,45	66,1	0,84868859
4	4,02	72,31	0,94548984
5	2,08	67,61	0,50592701
6	17,64	3534	0,59346526
7	3,59	3534	0,12077893
8	4,4	3534	0,14802988
9	17,64	3534	0,59346526
10	9,36	3534	0,31489994
11	15,58	3534	0,52416036
12	2,73	60,84	0,7
13	14,81	3530	0,49853735
14	7,18	3530	0,24169468
15	6,2	3530	0,20870571
16	3,87	3530	0,13027276
17	3,05	3530	0,10266974
18	7,74	3530	0,26054552
19	4,23	3530	0,14239115
20	13,21	3530	0,44467781
21	5,93	3530	0,19961691
22	3,11	60,61	0,7989475
23	7,33	3526	0,24688393
24	14,82	3526	0,49915686
25	2,97	3526	0,10003346
26	2,97	3526	0,10003346
27	3,87	3526	0,13034663
28	3,06	3526	0,10306478