Lianne te Mebel

Deception Detection by Recognition of Deception Cues

Master Thesis Conflict Risk & Safety

Supervisor: Peter Slijkhuis

Second supervisor: Mariëlle Stel

24-08-2020

Keywords:

deception, deception cues, veracity judgement, gender, involvement, cognitive load

Abstract

To measure whether the findings of Bond, Levine and Hartwig (2015) are applicable during digital interactions, a study is implemented in which six deceptional cues were tested. Therefore, a 2 x 2 between-subject design research was implemented, with cognitive load (high cognitive load in the receiver condition and low cognitive load in the observer condition) and involvement (high involvement in the video condition and low involvement in the audio recording) as manipulations. The results show that the selected deceptional cues of Bond et al. (2015) are mainly supported in a technological setting. Significant findings are the deceiver being experienced as less cooperative, slightly harder thinking and hiding one's face. The seriousness of the crime is, as expected, non-significant. Lastly, the influence of gender remains inconclusive since the results support both sides. Future research can focus on micro expressions (visual and auditive) in online veracity judgements.

Introduction

Deception

Deception is a common problem. In the Netherlands, a total of 41730 crimes were registered in 2018, which belong under the case of deception (Centraal Bureau voor de Statistiek, 2019). Examples of those crimes are committing fraud and delusion. Due to the common presence of such crimes, it is relevant to keep researching this topic. Currently, different methods exist to detect deception. Examples of often-used methods are non-automated systems that visualize deception and deception detection by cue recognition. Research found different advantages and disadvantages regarding those methods. Still, relatively little research can be found which address those methods in a digital setting. Therefore, the following sections will explain which methods will be relevant to use in a digital setting.

Non-automated systems to detect deception

Several non-automated systems exist to achieve a more accurate means to detect and visualise deception. For instance, Functional Brain Imaging is shown to be a reliable method to detect deception and differentiate between deception and other activities in the Frontoparietal lobes (Yu, Tao, Zhang, Chan & Lee, 2019). Despite this possibility, Functional Brain Imaging is currently difficult to interpret due to an indirect form of measuring neural activity fluctuations (Buckner, Kriener & Yeo, 2013). Therefore, observations remain ambiguous without additional insights about their mechanisms. In addition to this, it is very difficult to perform a Functional Brain Image on people when doing research about deception, especially in digital interactions, due to the large size of the apparatus and when being applied, the reduction of contact with the user.

A second system to detect deception is eye tracking. Prior research shows that eye tracking should be able to reveal deception, by the interpretation of brief oculometric behaviours (Proudfoot, Jenkins, Burgoon & Nunamaker, 2016). Despite the drastic improvement of this technique in the last years, little research is currently implemented whether this method works (Lai et al., 2013). Furthermore, it can be difficult to detect deception when the receiver is not physically present, since the person to be interviewed need to be able to use the materials provided to track one's eyes. Also, differences in screen size and viewing distance will cause the receiver to be unlikely to understand where the participant looks at (Granka, Feusner & Lorigo, 2008). Thus, this method is difficult to use in online settings when the receiver is physically absent.

Deception detection by cue recognition

Despite several non-automated systems can be used to detect deception, they both still have limitations. Another method that can help to detect deception, has been researched and can be applied when the receiver is not physically present, is the use of cue recognition. Cues to recognize deception can be verbal or non-verbal. Findings are still differentiating whether those cues are reliable cues to detect deception. Regarding the verbal cues, it can be found that the use of more pauses when deceiving is a significant cue, whether the richness of details used is a non-significant one, in contrast with the current belief (Granhag & Strömwall, 2002). Also the pitch is shown to increase when a direct question is asked, and the answer is deceiving (Sondhi, Khan, Vijay & Salhan, 2016).

Despite some verbal cues are shown to be significant and others are disputable in their effectiveness to detect deception, people are more likely to conceal their verbal deception cues than their non-verbal ones (Caso, Vrij, Mann & De Leo, 2011). Therefore, a focus will be put on the visual cues in the following sections.

An author who did many research towards visual deception cues is DePaulo et al. (2003). DePaulo et al. (2003) tested 158 cues, of which 42 seemed to be significant. Those cues can be recognized when people try to deceive another person. Examples of those significant cues are that messages can be perceived as less cohesive and that people are perceived as less genuine, less emotionally involved and more tensed, compared with people who speak the truth (DePaulo, Epstein, & LeMay, 1990; DePaulo, Kashy, Kirkendol, Wyer & Epstein, 1996; Markus, 1977; Vrij, 2000).

Several cues can thus help to indicate whether a person deceives or not. Still, despite these findings, even people who are trained to detect deception are able to only detect 54% of the lies, compared with 50% of untrained judgers (Levine, 2010). The latter is equal to the level of change (50%). This is further supported by Driskell (2010), who states that the effect of a deception detection training is found to be positive and significant, but of a moderate magnitude. Despite the moderate magnitude, there are people who score significantly higher than the average (Frank, Menasco & O'Sullivan, 2008). An example can be found in gender: females are significantly better at detecting deception than males (Tilley, George, Marett, 2005), in real-life and video settings (Johnson et al., 2004). A higher commitment and a better self-awareness can be possible reasons why some people can detect deception better than others (Johnson et al. 2004).

Despite the moderately positive findings of visual cue detection, some researchers also disagree with cue recognition being a method to detect deception. In general, headshaking and negative facial expressions, which are brief and incomplete changes in expression (micro expressions), are the most common non-verbal behaviours during deception (Ganis, Kosslyn, Stose, Thompson & Yurgelun-Todd, 2003). Tsechpenakis et al. (2005) and Burgoon (2018), state that it is very difficult for humans to detect those non-verbal cues visually. Also Porter, ten Brinke and Wallace (2012) describes that people are relatively unable to distinguish micro

expressions from liars and truth tellers, since micro expressions of sadness, fear and disgust are relatively similar and thus difficult to differentiate. Furthermore, Buller and Burgoon (1996) argued that deceivers attempt to control their nonverbal behaviour in order to appear as credible. As can be noted, the outcomes of micro expressions are currently diverging and findings regarding this topic still need to be clarified.

Deceptional cues

When clarifying whether cue detection is an efficient method or not, it can help to demonstrate whether the effectiveness of deception recognition training can be improved somehow. Furthermore, deception can also occur in digital settings, in which other methods turn out to be difficult to implement and this method can be helpful to recognize one's veracity. Part of the unclarity that exist regarding the effectiveness of visual deception cues, can already be clarified by the meta-analysis of Bond et al. (2015). The authors tested the 158 cues of DePaulo et al. (2003). Forty-three of the cues were omitted because they were studied only once. Therefore, 115 cues were analysed and differentiated in effectiveness to recognise deception. In the following paragraphs, the two most efficient and the two least efficient will be described. In this way, it can be researched whether similar findings can be found when the interaction takes place in a digital setting. Furthermore, two additional cues will be described since those cues can possibly influence research outcomes.

Least effective cues

DePaulo et al. (2003) state that 'hiding one's face when deceiving' and 'seriousness of the crime' are helpful cues to detect deception. The meta-analysis of Bond et al. (2015) state that those cues are of little effectiveness to detect deception. Therefore, some additional research was assessed regarding those cues. Additional research related to 'hiding one's face when deceiving', shows that when people lie, they will look more often at the receiver, but for

shorter time intervals then when giving a truthful response (Jorgensen, 2015). This fixation avoidance is a response on the cognitive load or arousal that someone goes through during intentionally deceiving (Proudfoot et al., 2016).

Regarding the seriousness of the crime, DePaulo, Ansfield, Kirkendol and Boden (2002) state that the discovery of a lie can have severe consequences. The severity of the lie is expected to increase stress and therefore visualize cues (DePaulo et al., 2003). Still, next to source of DePaulo et al. (2003) no other research could currently be found that supports that this cue leads to recognition of deception.

Effective cues

The highest differentiation between liars and truth-tellers can be found in the person's cooperation and whether the person has to think hard about the answer. Therefore, those cues are stated to be the most effective cues to recognize deception (Bond et al., 2015). For those results, much support can be found in the literature. For the former cue, it is tested that when people tell the truth, they are more likely to cooperate than when lies are told (Mehrabian, 1972; Wiener & Mehrabian, 1968). This is supported by Gudjonsson (2003), who states that deception is mostly motivated by avoidance behaviour.

Also the cue that people have to think harder about the answer they are going to give, is supported by additional research. The meta-analysis of Suchotzki, Verschuere, Van Bockstaele, Ben-Shakhar and Crombez (2017) show that liars have a larger responding time than truth tellers. It is stated that, in contrast with people who tell the truth, liars need to make a decision what they are going to lie about and need to construct a lie (Walczyk, Mahoney, Doverspike & Griffith-Ross, 2009). This process causes the longer period of responding time. Several studies indicate that reducing the response time increases the perceived honesty of the people (Capraro, 2017; Lohse, Simon & Konrad, 2018).

Additional cues

Indifference and the amount of eye contact are relevant cues to consider, since those cues can predict lying, but also can cause attribution bias or guide exploring patterns. Whether the deceiver seems indifferent regarding the story, is the third most reliable cue that reveals deception (Bond, Levine & Hartwig, 2015). However, the cue of indifference can also be a factor prone to attribution bias. When research on deception is implemented in an unnatural setting, such as a lab setting, it is possible that participants perceive the interviewee as indifferent, which influence the perception of the other perceived cues. People try to make their own understanding of behaviour and create their own coherent perspective from the information (Tetlock & Levi, 1982). Therefore, inductive fallacy can occur: the outcome will be generalized to all cues, with a significant influence of the outcome in total (Walton, 1999).

Second, according to Bond et al., (2015), a moderate predicting cue is the amount of eye contact made. Comparing a moderate predicting cue with the high and low prediction cues can help to explore possible patterns. Other research shows different findings regarding this topic. When forcing a liar to make eye contact, higher cognitive demands are asked from the deceiver and in this way, lying can more easily be detected (Vrij, Mann, Leal & Fisher, 2010). Also, people should be able to recognize a reduction of eye contact when someone lies (Levine, Asada & Park, 2006). Still, other authors completely contrast this finding, by stating that liars deliberately seek eye contact to convince the interviewer that that person tells the truth (Mann et al., 2013). Sporer and Schwandt (2007) state that no evidence can be found that people avoid eye contact while lying, although gaze aversion is generally seen as the most important signal of deception. So, no consensus currently exist regarding this predicting cue and additional research is needed regarding this cue.

Involvement

To be able to compare the different deception cues described above, two different aspects should be considered that can influence the level of cue recognition. Difference in involvement, for example, can influence the accuracy of one's veracity judgement. The level of conversational involvement can range from high involvement by direct participation in the interview, to a lower level of involvement by observation of the conversation (Hartwig, Granhag, Strömwall & Vrij, 2004). Higher involved judgers use more judgement-relevant information, that the deceiver shows to them non-verbally (Reinhard, 2010). An example can be seen in practice. In interviews with an offender, one or more police officers are present while interrogating the suspect (Politie-verhoor[1], n.d.). No information can be found that an observer is present during those interviews. Jorgensen (2015) explains this by the perspective that direct participation in a research, instead of observation, makes the researcher more open to observe broader themes instead of testing a hypothesis. This will make it more likely for the researcher to notice cues, that he did not specifically focus at. In contrast, Musante & DeWalt (2010) claim that the addition of an observer enhances the quality of obtained data, the interpretation of data and improves the construction of research questions and hypotheses. Participation and observation thus change the manner that cues are recognized.

Another type of involvement can be found in the type of interview: direct (face-to-face) and indirect (through a medium) interviewing. The indirect communication in setting involvement can be subdivided in video and audio communication. Currently, a police interview is implemented by direct face-to-face communication (politie-verhoor[2], n.d.). This can be explained because research found differences in outcomes between direct and indirect interviews. Indirect communication is stated to obtain a lower feeling of involvement, self-disclosure, relationship building and increase the risk of confounding variables (Jorgensen, 2015; Mcconville, 1992; Lyyra, Myllyneva & Hietanen, 2018; Ruppel et al., 2016). Also, van

der Kleij, Schraagen, Werkhoven & de Dreu, 2009), state that senders in direct communication receive direct feedback about how their message is understood, in contrast to video meetings and that the use of specific gestures such as pointing are more easily interpreted when a face to face meeting takes place.

Despite those outcomes, other research contrast those findings or describe the positive aspects of the use of a medium. For example, O'Malley, Langton, Anderson, Doherty-Sneddon and Bruce (1996) show opposite findings of van der Kleij et al. (2009). These authors state that in both the direct and the indirect setting, the receiver uses visual cues to check the content of the message. Due to those visual cues, the mutual understanding remains similar in both settings (O'Malley et al., 1996). In addition to this, a video recording can be a valuable tool to support decisions in forensic judgements (Blandon-Gitlin & Mindthoff, 2018). In those videos, it can be helpful to obtain a higher accuracy in veracity judgements. Therefore, it can be relevant to test the accuracy of veracity judgements in indirect communication.

Despite the statement that indirect communication can be a valuable tool to detect deception, it is important to recognize the differences in the type of medium used. Deception detection via audio is shown to be significantly worse than communication by video (Horn, 2001). Still, the outcome can be caused by other aspects than the actual difference in involvement. For example, a change in medium can influence the quantity of the message provided (Qin, Burgoon, Blair & Nunamaker, 2005). More specifically, the change in medium changes the number of words used, verbs and sentences (Qin et al., 2005). This can be prevented using a similar message content between the two mediums. Lastly, technology made a rapid improvement in the last years, which can change the outcome of this study. Therefore, outcomes can possibly change when new research is implemented, with better technological quality and a similar message quantity.

When comparing the different types of involvement, it seems that direct communication is likely to be preferred above indirect communication. Still, detecting deception through indirect communication can still be important in specific situations, for example when face-to-face communication is not possible. An example is during the lockdowns during the COVID-19 breakout in 2020. In here, a lot of communication takes place in an online environment, in which it can be relevant to detect deception. Examples are job applications or annual interviews with employees. Furthermore, people who need to communicate from a large distance are not always able to meet each other in person. Therefore, a focus will be directed on the indirect forms of communication.

Cognitive load

Next to involvement, cognitive load is a factor that can have an influence on one's deception cues, with as consequence an increase in one's veracity judgement. Two theories will be discussed, which will explain this finding. First, according to the Cognitive Load theory, people have limited working memory capacity in their cognitive capacity (Skulmowski & Rey, 2017). An increased cognitive load is a consequence of a more demanding task, with performance reduction, stress and errors as a consequence (Nourbaksh, Chen, Wang & Calvo, 2017). This cognitive capacity can be measured by ones eye movement, by means of cue recognition or eye tracking, but also by the use of physiological data such as electro-dermal activity, heart rate and breathing pattern (Herten, Otto & Wolf, 2017; McDuff, Hernandez, Gontarek & Picard, 2016; Pouw, Mavilidy, van Gog & Paas, 2016; Zimasa et al., 2018;). Furthermore, it is shown that people increase the use of hand gestures to compensate for a high cognitive load (Pouw et al., 2016).

In contrast to the cognitive load theory, the dual-process theory states that people with a high cognitive load use a more intensive form to test veracity judgement (Reinhard & Sporer, 2008). People who are under a low level of cognitive load focus especially on the

non-verbal behaviour, while people with a high level of cognitive load focus at both visual and verbal behaviour. This finding is supported by the claims from Reinhard (2010) and Jorgensen (2015) in the involvement sections: more involved judgers use more judgement-relevant information that the deceiver shows to them non-verbally and direct participation in a research, instead of observation, makes the researcher more open to observe broader themes instead of testing a hypothesis.

Direct and indirect measurements

For now, it became clear that it can be relevant to measure whether deception cues can help improve one's accuracy in veracity judgement in an online environment. Still, due to this online setting, it can be difficult to objectively test the reliability. Therefore, it will be important to measure the consistency of the direct and the indirect measurements. People can, for example, state that they belief that the other person lies. However, when the other responses given show that people should think that the deceiver is telling the truth, the outcomes are inconsistent. Controlling for inconsistent statements increases the reliability of the research. To prevent this reduction in reliability, the consistency of the direct and indirect measurements is checked whether the outcomes of the persons are consistent or not.

Current research

In the previous sections, it is shown that the use of non-automated instruments have their own limitations and are not always possible to apply, such as in the COVID-19 breakout in 2020. Therefore, it can be relevant to be able to recognize deceptional cues as an additional measurement in forensic judgements. To test whether this can be a valuable method, a study will be implemented to test whether the findings of Bond et al. (2015) are applicable in an online environment. It will be tested whether the strongest deceptional cues can be recognized in practice by means of cue recognition accuracy, while considering the difference in involvement and cognitive load. The findings will shed light on whether deception cues can

be used in a broader context and improve decision making in forensic judgement. To find out whether this goal can be achieved, the following research question is stated: *"Can the focus on different deception cues, while differentiating in gender, involvement and the level of cognitive load influence the accuracy of one's veracity judgement?"*

In order to answer this research question, nine hypotheses are stated:

H01: *There is a positive correlation between the answers given at the direct and indirect measures.*

H02: *Female participants have a higher accuracy in veracity judgement than male participants in all conditions*

H03a: *A higher cognitive load increases a person's accuracy in veracity judgements*

H03b: *A higher cognitive load increases a person's attention at the deceiver's non-verbal behaviour*

H04: *A higher level of involvement increases a person's accuracy in veracity judgement*

H05: *People who have a higher accuracy in veracity judgement, perceive the deceivers as less cooperative*

H06: *People who have a higher accuracy in veracity judgement, perceive the deceiver to think hard about their answers*

H07: *People who have a higher accuracy in veracity judgement, perceive the deceiver being more indifferent than people with lower accuracy in veracity judgement.*

H08: *People who have a higher accuracy in veracity judgement, slightly perceive the deceiver to make less eye contact than people with lower accuracy in veracity judgement.*

H09: *No differences exist between people with high and low accuracy in veracity judgement, in the seriousness and the face shielding of the deceiver.*

Methods

Design

The research implemented is a 2 x 2 between-subject design. The dependent variable is the accuracy of veracity judgement. Independent variables are gender and the category that the participant is in, which varies in the level of involvement and the level of cognitive load.

Participants

In total, n=76 participants participated in this study. However, in the observer condition, 16 participants did not complete the study. n=13 Participants were excluded since too much data was missing. n=3 participants were kept in the dataset, since the answers seem to be filled in seriously and the amount of data was sufficient for some analyses. Therefore, a total of n=63 participants remain included in this study.

Participants from the receiver condition were students from the University of Twente, but also people from the network of the data collectors. The nationalities were Dutch (3.3%), German (76.7%) and other (20.0%). In the sample, age was ranging from 19 to 32 with an average of 20.5. In total, 36.7% of the participants was male and 63.3% was female. The participants indicated that they studied Psychology (50.0%), Communication Science (20.0%) and other (30.0%).

The participants in the observer condition were students from the university of Twente, but also people from the network of the data collectors. The nationalities were Dutch (54.5%), German (42.4%) and other (3.0%). In the sample, age was ranging from 20 to 55 with an average of 23. In total, 54.5% of the participants was male and 45.5% was female. The participants indicated that they studied Psychology (27.3%), Communication Science (9.1%) and other (63.6%).

Materials

When the participant subscribed for the study, an informed consent and scenario (see appendices A and B) were necessary to prepare the participant. Also, a list with questions asked during the interview was present (see appendix C). The questions were carefully selected: Open questions were used with an equal change of answering a 'yes', compared with getting a 'no', with predetermined answers. In this way, the possibility of biasing answers by the likelihood of estimators is reduced. Furthermore, an image was created, to help the participant indicate where the focus was directed to when answering the questions (see appendix D). Lastly, a post-questionnaire was used to measure the participants' perception and possible stereotypes. After the interview, the statistical software programme 'IBM SPSS' was used to analyse the outcomes of the study.

Procedure

When the participants accepted to participate into the study, they were randomized into one of the four following conditions:

- **Video condition as receiver (vidR)** In this condition, the participant interacted with an actor through a video screen. The participant asked questions to the actor and decided whether the actor was genuine or not. Because the person was directly involved in the interaction, this condition is expected to be a condition with the highest experience of involvement due to a more direct form of communication. The highest amount of cognitive stress is expected because the task is a demanding one from the participant.
- **Video condition as observer (vidO)** In this condition, the participant observed an interaction between a receiver and the actor. The actor was visible through a video screen. The participant observed how questions are asked and whether he thinks that the actor was deceiving the receiver. In this condition, the second highest experience

of involvement is expected because the participant sees the participant and the second lowest level of cognitive stress was expected, because of the low demands of the task.

- **Audio condition as receiver (audR)** In this condition, the participant interacted with an actor through an audio format. In this condition, the participant asked questions to the actor and decided whether the story of the actor was genuine. Because the person did not directly see the participant, the second lowest level of involvement is expected. Still, the second highest level of cognitive stress is experienced because of the demands of the tasks.
- **Audio condition as observer (audO)** In this condition, the participant observed an interaction through an audio communication between a receiver and the deceiver. The participant observed how questions are asked and whether he thinks that the actor was deceiving the receiver. It was expected that the lowest amount of cognitive stress was experienced due to the passive role of the participant. Also, the lowest level of involvement was expected because the person does not see the face of the deceiver directly.

The research questions can be answered with a sufficient amount (n=30) of participants in the different conditions. Since very little participants were subscribing themselves for the study, it was decided to cancel the randomization. Also, the audio conditions were excluded since they only give some additional information about the level of involvement (H04). Thus, participants from the researchers' network were asked for participation and only two conditions with a high level of involvement were assessed.

When participants were allocated in one of the conditions, they were sent instructions how they can get online during the study. At the allocated moment, the participants were welcomed. The purpose of the study was told and they were asked to agree with the informed consent. Furthermore, demographic information had to be filled in on the attachment. When

the first part was completed, the participant was given instructions about the procedure of the task. When the condition was vidR or audR, one additional researcher was present who took the role of deceiver. In the video conditions, the actor was online visible through a screen or only by sound. In every condition, the same actor was used as deceiver and the questions were answered similarly.

When starting the experiment, the receiver obtained a scenario, which includes the name and study of the person, the situation that the student is accused of and the possible consequences of cheating. Then, the participants in the receiver condition started asking structured questions to the sender. The participants in the observer condition watched another person asking these questions. The first two questions were questions that ask the name and the study of the deceiver. The sender answers those by telling the truth. In this way, the participant could notice that those answers are true, since this information was provided in the scenario.

After the basic questions, the more loaded questions were asked one by one. The deceiver answers all of those by telling a lie. After each question, the participant had to fill in whether he beliefs the answer is the truth or lie. Furthermore, the participant was asked where this answer was based on. In both conditions, aspects of the deceiver's voice were mentioned (such as pitch, tone of voice etcetera), in which the participant can select multiple options. Furthermore, in the video conditions, a picture was present which showed the upper body of a person, in which bodily parts can be selected where the focus was directed at. This procedure goes on until all eight questions were completed. Then, the participant received the post-questionnaire, to measure what his perception regarding Laura is and whether he has stereotyped attitudes, which could influence the outcome. After filling in those questions, the participant was debriefed for his participation. In total, the procedure took around thirty minutes per participant.

Statistical analysis

After finishing the data collection, the data was categorized and analysed. First, it was tested whether the participant had a high or low consistency in their direct and indirect measurements. For H01, the participants' consistency was checked by means of a correlation check regarding the direct and indirect measurements and visualized by a graph. In this way, it could be seen whether participants gave similar answers to related questions and thus gave a representation of the understanding and the seriousness of the participants who filled in the questionnaire.

Next, the outcomes of the data were analysed. Before analysing the data of the following hypotheses, it was important to test whether the data is normally distributed by exploring the statistics in IBM SPSS. More specifically, it was measured whether the data is normally distributed according to the Shapiro-Wilk test, an analysis for small sample sizes. Then, for H02 and H03, it was tested whether a significant difference in accuracy between the groups (gender or condition) exist, in one of the eight questions that measure the participants' accuracy. To measure this, an independent samples t-test would be implemented when the data is normally distributed and a Mann-Whitney U test would be implemented when the data is significantly deviating from a normal distribution. For H03, another Mann-Whitney U test was implemented to test whether the claim of Reinhard and Sporer (2008) is correct: participants with a higher cognitive load should focus more at the non-verbal behaviour of the deceiver than participants with a lower cognitive load.

For H05 until H08, a new variable was made, which classifies the participants in moderate veracity judgers (0-4 correct) and expert veracity judgers (5-8 correct). Frequencies from those groups were calculated and different tests were implemented to test whether having a higher accuracy in veracity judgement, influences the perception to experience the deceiver as less cooperative, harder thinkers about their answers, being more indifferent and

make less eye contact than people with a moderate accuracy in veracity judger. How this expectation would be analysed, was tested depending on the distribution. Normally distributed items would be tested by means of an independent samples t-test, while items that deviate significantly from the normal distribution would be tested with the Mann-Whitney U test. Lastly, the perceived seriousness of the crime and the face shielding (H09) were tested in a similar procedure as the variables tested above.

Results

H01: *There is a positive correlation between the answers given at the direct and indirect measures.*

In general, a direct question: ('To which degree did you base your decision on visual behaviour?') and an indirect question ('Did you pay attention to the face of Laura?') are negatively, but mainly significant or close-to-significant correlated (see table 1). This is further supported by figure 1. When someone focuses more at visual behaviour, people in general state to pay more attention to the face.

Table 1

Correlation of direct measure 'To which degree did you base your decision on visual behaviour?' and indirect measure 'Did you pay attention to the face of Laura?'

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
r	-.262	-.302	-.227	.016	-.224	-.198	-.165	-.465
p	.038	.016	.076	.901	.082	.127	.204	.000

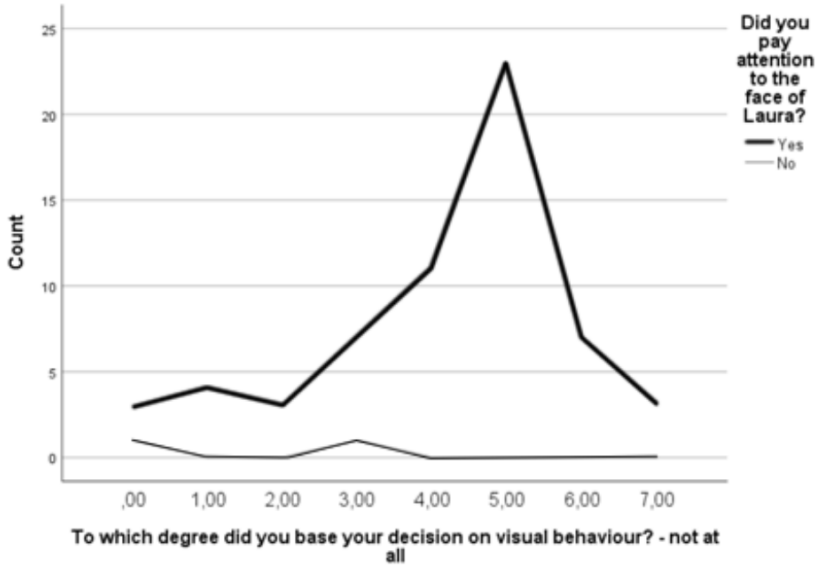

Figure 1: graphical representation of direct and indirect measure

H02: *Female participants have a higher accuracy in veracity judgement than male participants in all conditions*

When implementing an analysis for the normal distribution, it can be noted that the eight questions that assess the accuracy of the veracity judgement are significantly deviating from a normal distribution for both genders ($p<.001$). Therefore, the Mann-Whitney test was implemented to assess the difference in accuracy between men and women. In table 2, it can be seen that the first statement confirms the hypothesis. The other questions do not show a significant difference. Still, it is relevant to notice that the last question is close-to-significant, which contrasts with the expectation.

Table 2

Differences in accuracy of veracity judgement between gender conditions

Statement	Average men	Average women	U	Z	p
Q1	29.5	34.1	420.5	-2.135	.033
Q2	32.3	31.8	485.5	-.203	.839
Q3	31.9	31.2	465.0	-.180	.857
Q4	31.6	31.4	474.0	-.034	.973
Q5	32.7	29.6	414.0	-.954	.340
Q6	31.6	30.5	446.0	-.272	.786
Q7	32.1	30.1	431.0	-.530	.596
Q8	33.8	27.6	356.0	-1.746	.081

Since the outcomes contrast each other, a regression analysis was implemented to test for possible confounding variables. The findings show that age (p=.685), nationality (p=.061) and field of study (p=.173) and gender (p=.057) are no variables that confound the persons' veracity judgement. Still, it is relevant to note that the variables, especially gender, do influence the outcome somewhat.

H03: *A higher cognitive load increases a person's accuracy in veracity judgements*

The level of cognitive load was manipulated by the condition that the person was categorized in. The eight questions that assess the accuracy of the veracity judgement show a significant deviation from the normal distribution for both conditions (p<.001). Therefore, a Mann-Whitney U test was implemented. This test resulted in no significant finding between the two conditions in the eight different questions (see table 3).

Table 3

Differences in accuracy between conditions with a manipulation in cognitive load

Statement	**Average video receiver**	**Average video observer**	**U**	**Z**	**p**
Q1	33.7	30.5	444.0	-1.499	.134
Q2	32.2	31.9	490.5	-.122	.903
Q3	30.9	32.0	463.0	-.278	.781
Q4	30.8	32.1	460.0	-.340	.734
Q5	30.1	31.9	438.0	-.535	.593
Q6	31.7	30.3	443.5	-.364	.716
Q7	31.8	30.2	440.5	-.418	.676
Q8	28.0	33.0	375.0	-1.420	.155

In addition to this, an analysis was implemented to check the claim of Reinhard and Sporer (2008), who state that people with a higher cognitive load focus more on non-verbal behaviour than people with a lower cognitive load. Since 28 of the 32 items were significantly deviating from the normal distribution, the implemented analysis was a Mann-Whitney U test. It is shown that one item, which asked participants how much they looked at the visual aspects, significantly differentiated between the group with the higher cognitive load (μ=26.2) and the group with the lower cognitive load (μ=35.7) (U=321.0, Z=-2.119, p=.034).

H04: *A higher level of involvement increases a person's accuracy in veracity judgement*

Since very little participants were participating in the research, it is decided to remain this hypothesis unanswered.

H05: *People who have a higher accuracy (5-8 questions correct) in veracity judgement, perceive the deceivers as less cooperative*

In total, n=41 participants (68.3%) were categorized in the moderate condition and 19 (31.7%) participants were categorized in the expert condition. The Shapiro-Wilk test shows a normal distribution for the expert veracity judgers ($p=.095$), but a significant deviation from the normal distribution at the moderate veracity judgers ($p<.001$). Since the latter is far from a normal distribution, a Mann-Whitney U test was implemented. The findings show that the difference is not significant between the two classifications ($U=3195$, $Z=-1.144$, $p=.253$). Still, the direction is according to expectation. Participants from the expert classification ($\mu=26.8$) experience the deceiver as less cooperative than participants from the moderate classification ($\mu=32.2$)

H06: *People who have a higher accuracy (5-8 questions correct) in veracity judgement, perceive the deceiver to think hard about their answers*

Both the moderate veracity judgers ($p<.001$) and the expert veracity judgers ($p=.001$) show a significant deviation from the normal distribution. Therefore, the Mann-Whitney U test was implemented, with a close-to-significant finding as result ($U=278.0$, $Z=-1.879$, $p=.060$). The participants with a higher accuracy are shown to belief more that the deceiver has to think hard ($\mu=36.4$), compared to the moderate veracity judgers ($\mu=27.8$).

H07: *People who have a higher accuracy (5-8 questions correct) in veracity judgement, perceive the deceiver being more indifferent than people with lower accuracy in veracity judgement.*

The moderate veracity judgers ($p<.001$) and the expert veracity judgers ($p=.027$) both show a significant deviation. The Mann-Whitney U test shows that no significant difference ($U=359.0$, $Z=-.507$, $p=.612$) occurs between the moderate ($\mu=31.2$) and the expert ($\mu=28.9$) group.

H08: *People who have a higher accuracy (5-8 questions correct) in veracity judgement, slightly perceive the deceiver to make less eye contact than people with lower accuracy in veracity judgement.*

Both moderate ($p<.001$) and expert ($p=.01$) show a significant abnormal distribution. The Mann-Whitney U test finds $\mu=32.0$ for the moderate veracity judgers and $\mu=27.3$ for the expert veracity judgers. This difference is not significant ($U=329.5$, $Z=-1.002$, $p=.316$).

H09: *No differences exist between people with high (5-8 questions correct) and moderate/low (0-5 questions correct) accuracy in veracity judgement, in the seriousness and the face shielding of the deceiver.*

The face shielding had $p=.001$ in the expert condition and the other categories in the face shielding and seriousness of the crime had $p>.001$, which shows that all categories had a significant difference regarding the normal deviation. The Mann-Whitney U test shows no significant finding for the seriousness of the crime ($U=372.0$, $Z=-.296$, $p=.767$). However, whether the deceiver hides his face does show a significant difference ($U=266.0$, $Z=-2.161$, $p=.031$) between the moderate ($\mu=27.5$) and the expert ($\mu=37.0$) veracity judgers in the expected direction.

Discussion

When looking at the results, different things are noticeable. First, a significant finding can be found, which supports the hypothesis that females are more accurate in making veracity judgements. However, a second question contrasts this finding, by the attainment of a close-to-significant outcome in the opposite direction. This statement implies that males are more accurate in making veracity judgements. Therefore, the findings are currently not informative about the direction of the hypothesis and whether an effect occurs. A possible explanation can be the sample size: since little participants participated in the study and the groups were not randomized, a possible unknown confounder could have caused one of the outcomes. The variables age, nationality, field of study and gender itself were already excluded as possible confounder.

Next to gender, it was expected that cognitive load influenced the level of veracity judgement. No significant effect was found regarding this relationship. However, a significant item was found about the expectation that that people focus more at the deceivers' non-verbal behaviour, when experiencing a high cognitive load. This significant item was in opposite of the expectation: people with a lower cognitive load focus more at the non-verbal behaviour of the deceiver. The former mentioned Cognitive Load theory can explain this finding: because too much is demanded from the participant, a reduction of performance occurs and therefore, the participant has a lower level of focus (Skulmowski & Rey, 2017).

The expectation that people with a higher accuracy perceive the deceiver as less cooperative than participants with a lower accuracy, was supported by the analyses. Also the hypothesis that expert veracity judgers experience the deceiver to think harder about the answer than moderate veracity judgers, was supported by a marginally significant finding. No support was found regarding the expert category perceiving the deceiver as being more indifferent, which is in contrast to the conclusion of Bond et al. (2015) that it is a moderate

cue. Still, the rank means of the group do show a direction that supports the hypothesis. Those findings are similar to the results of H08, which also has a positive but non-significant result that the experts expect the deceiver to make less eye contact. Still, the expectation from the latter variable was that the difference was only small and therefore, it will be difficult to get a significant result anyhow. Lastly, one of the items confirm H09: there is no difference in the perceived seriousness of the crime. The other item is shown to be significant: Expert veracity judgers show a higher recognition of the deceiver hiding his face than the moderate veracity judgers.

In total, several findings came out of the analyses, which help to support the research question: *"Can the focus on different deception cues, while differentiating in gender and the level of cognitive load influence the accuracy of one's veracity judgement?"*. It is shown that gender remains inconclusive, since the results support both sides. Other findings do explain the answer to the research question. The Cognitive Load theory is supported: people with a high cognitive load reduce in their performance. Furthermore, a significant effect is found that people with a higher accuracy perceive the deceiver as less cooperative and a slight effect was found that experts perceive the deceiver as thinking hard, which is in coherence with Bond et al. (2015) their findings. Furthermore, the seriousness of the crime is, as predicted, a non-significant cue. Lastly, hiding one's face seems to be conflicting with the findings of Bond et al. (2015), since a significant effect was found for this cue. In total, the outcomes of this research show overlapping findings with the research of Bond et al. (2015). Therefore, it can cautiously be said that those findings can possibly be applied when interaction takes place through a medium. This study also supports the finding that deception cues can possibly be recognized by people.

Different tests were analysed to test the reliability of the study. For example, the descriptive statistics show that the groups manipulated by cognitive load deviate somewhat in

the demographics, which is likely due to the manner of participant gathering. Furthermore, the often-significant deviation from a normal distribution can be related to the relatively small sample size. Despite these findings, the data seems to be relatively reliable. Significant deviations from the normal distribution and possible confounders are considered and participants responded in a similar manner to direct and indirect measures. No big deviations could be found and it seems that participants filled in the questionnaires seriously and understood the questions.

Several limitations could be found regarding this research. The first limitation is that randomization was not possible, since the lack of participants who were willing to participate, due to the COVID-19 breakout. Still, the results indicate that none of the measured variables confound the dependent and the independent variable. However, the presence of an unknown confounder cannot be excluded. Despite this finding, it is shown that participants reacted similarly at direct and indirect measurements and therefore, the data seems sufficient.

Second, the deceiver was an actor, who was not really in the stressful situation. Therefore, people can base their response on the acting instead of the actual deceptional cue. Still, the analyses show outcomes that were supporting the findings from Bond et al. (2015). Therefore, it is probable that this bias occurred, but it does not have an effect of what is researched currently.

Lastly, two participants did answer one of the last questions “I did fill in this questionnaire seriously” that they did not fill in the questionnaire seriously. However, when scanning over them manually, they did answer two questions that were known from the scenario correctly and their direct and indirect measures are similar. Therefore, the participants remained included. An explanation for this can be that they lost their interest at the end of the questionnaire or that they simply found the question itself irrelevant and therefore filled it in incorrectly.

The perspective that deception can be recognized by specific micro expressions remains a disputable item in research. This study gave some support that micro expressions do have an influence at a persons' veracity judgement. Significant findings are the deceiver being experienced as less cooperative, slightly harder thinking and hiding one's face. The seriousness of the crime is turns out to be non-significant as expected and the influence of gender remains inconclusive since the results support both sides. More evidence is needed to have sufficient support for the claim that deception detection by micro expressions is possible in a digital setting. Furthermore, this research focused at both the visual and the auditive aspects of the research. Future research can focus on the sole auditive version: are people able to recognize deception only by means of an auditive interaction? The level of cognitive load can be manipulated here by subcategorizing the sample in recorded and live audio interaction. Thus, more research needs to be done to complete the gaps regarding micro expressions (visual and auditive) in online veracity judgements.

References

Blandon-Gitlin, I., & Mindthoff, A. (2018). Do video recordings help jurors recognize coercive influences in interrogations. *Criminal juries in the 21st century: Psychological science and the law*, 195-219. Retrieved from https://books.google.nl/books?hl=nl&lr=&id=YB5pDwAAQBAJ&oi=fnd&pg=PA195&dq=Blandon-Gitlin+%26+Mindthoff,+2018&ots=jDlXGbh7dG&sig=g9vX4mZyyCmdkId7ahUa7_5UrNg&redir_esc=y#v=onepage&q&f=false

Bond, C. F., Levine, T., & Hartwig, M. (2015). New findings in non-verbal lie detection. *Deception detection: Current challenges and new directions*, 37-58. Doi: 10.1002/9781118510001.ch2

Buckner, R. L., Krienen, F. M., & Yeo, B. T. (2013). Opportunities and limitations of intrinsic functional connectivity MRI. *Nature neuroscience*, *16*(7), 832. Doi: 10.1038/nn.3423

Buller, D. B., and Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242. Doi: 10.1111/j.1468-2885.1996.tb00127.x

Burgoon, J. K. (2018). Microexpressions are not the best way to catch a liar. *Frontiers in Psychology.* Doi: https://doi.org/10.3389/fpsyg.2018.01672

Capraro, V. (2017). Does the truth come naturally? Time pressure increases honesty in one-shot deception games. *Economics Letters*, *158*, 54-57. Doi: https://doi.org/10.1016/j.econlet.2017.06.015

Caso, L., Vrij, A., Mann, S., & De Leo, G. (2006). Deceptive responses: The impact of verbal and non-verbal countermeasures. *Legal and Criminological Psychology*, *11*(1), 99-111. Doi: https://doi.org/10.1348/135532505X49936

Centraal Bureau van de Statistiek. (2019). *Geregistreerde criminaliteit; soort misdrijf, regio.* Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83648NED/table?fromstatweb

DePaulo, B. M., Epstein, J. A., & LeMay, C. S. (1990). Responses of the socially anxious to the prospect of interpersonal evaluation. *Journal of Personality*, *58*(4), 623-640. Doi: https://doi.org/10.1111/j.1467-6494.1990.tb00247.x

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*(5), 979–995. Doi: https://doi.org/10.1037/0022-3514.70.5.979

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. Doi: https://doi.org/10.1037/0033-2909.129.1.74

Driskell, J. E. (2012). Effectiveness of deception detection training: A meta-analysis. *Psychology, Crime & Law*, *18*(8), 713-731. Doi: https://doi.org/10.1080/1068316X.2010.535820

Frank, M. G., Menasco, M. A., & O'Sullivan, M. (2008). Human behavior and deception detection. *Wiley handbook of science and technology for homeland security*, *7*(2), 1. Doi: https://doi.org/10.1002/9780470087923.hhs299

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L. & Yurgelun-Todd, D. A. (2003). Neural Correlates of Different Types of Deception: An fMRI Investigation. *Cerebral Cortex*, *13*(8), 830–836. Doi: https://doi.org/10.1093/cercor/13.8.830

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16*(3), 243–257. Doi: https://doi.org/10.1002/acp.784

Granka, L., Feusner, M., & Lorigo, L. (2008). Eyetracking in online search. *Passive eye monitoring*, 283-304. Doi: https://doi.org/10.1007/978-3-540-75412-1_16

Gudjonsson, G. H. (2003). *The psychology of interviews and confessions: a handbook (chapter 1).* Hoboken: John Wiley and Sons.

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

Herten, N., Otto, T., & Wolf, O. T. (2017). The role of eye fixation in memory enhancement under stress–An eye tracking study. *Neurobiology of learning and memory*, *140*, 134-144. Doi: https://doi.org/10.1016/j.nlm.2017.02.016

Horn, D. B. (2001, March). Is seeing believing? Detecting deception in technologically mediated communication. In *CHI'01 Extended Abstracts on Human Factors in Computing Systems,* 297-298. Doi: https://doi.org/10.1145/634067.634243

Johnson, A. K., Barnacz, A., Constantino, P., Triano, J., Shackelford, T. K., & Keenan, J. P. (2004). Female deception detection as a function of commitment and self-awareness. *Personality and Individual Differences*, *37*(7), 1417-1424. Doi: https://doi.org/10.1016/j.paid.2004.01.011

Jorgensen, D. L. (2015). Participant observation. *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, 1-15. Doi: https://doi.org/10.1002/9781118900772.etrds0247

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational research review*, *10*, 90-115. Doi: https://doi.org/10.1016/j.edurev.2013.10.001

Levine, T. R. (2010). A few transparent liars explaining 54% accuracy in deception detection experiments. *Annals of the International Communication Association*, *34*(1), 41-61. Doi: https://doi.org/10.1080/23808985.2010.11679095

Levine, T. R., Asada, K. J. K., & Park, H. S. (2006). The lying chicken and the gaze avoidant egg: Eye contact, deception, and causal order. *Southern Communication Journal*, *71*(4), 401-411. Doi: https://doi.org/10.1080/10417940601000576

Lohse, T., Simon, S. A., & Konrad, K. A. (2018). Deception under time pressure: Conscious decision or a problem of awareness?. *Journal of Economic Behavior & Organization*, *146*, 31-42. Doi: https://doi.org/10.1016/j.jebo.2017.11.026

Lyyra, P., Myllyneva, A., & Hietanen, J. K. (2018). Mentalizing eye contact with a face on a video: gaze direction does not influence autonomic arousal. *Scandinavian journal of psychology*, *59*(4), 360-367. Doi: https://doi.org/10.1111/sjop.12452

Mann, S., Ewens, S., Shaw, D., Vrij, A., Leal, S., & Hillman, J. (2013). Lying eyes: Why liars seek deliberate eye contact. *Psychiatry, Psychology and Law*, *20*(3), 452-461. Doi: https://doi.org/10.1080/13218719.2013.791218

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35*(2), 63–78. Doi: https://doi.org/10.1037/0022-3514.35.2.63

McDuff, D. J., Hernandez, J., Gontarek, S., & Picard, R. W. (2016). Cogcam: Contact-free measurement of cognitive stress during computer tasks with a digital camera. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4000-4004. Retrieved from https://affect.media.mit.edu/pdfs/16.McDuff-et-al-CHI.pdf

Mehrabian, A. (1972). *Nonverbal communication*. Chicago: Aldine Atherton.

Musante, K., & DeWalt, B. R. (2010). *Participant observation: A guide for fieldworkers*. Rowman Altamira: United Kingdom.

Nourbakhsh, N., Chen, F., Wang, Y., & Calvo, R. A. (2017). Detecting users' cognitive load by galvanic skin response with affective interference. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *7*(3), 1-20. Doi: https://doi.org/10.1145/2960413

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G. & Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. *Interacting with Computers, 8*(2), 177-192. Doi: https://doi.org/10.1016/0953-5438(96)01027-2

Politie-verhoor[1]. (n.d.). *Algemeen.* Retrieved from https://politie-verhoor.nl/politieverhoor-algemeen/

Politie-verhoor[2]. (n.d.). *Hoe verloopt een politie verhoor?.* Retrieved from https://politie-verhoor.nl/hoe-verloopt-politieverhoor/

Porter, S., Ten Brinke, L., & Wallace, B. (2012). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior*, *36*(1), 23-37. Doi: 10.1007/s10919-0110120-7

Pouw, W. T., Mavilidi, M. F., van Gog, T., & Paas, F. (2016). Gesturing during mental problem solving reduces eye movements, especially for individuals with lower visual working memory capacity. *Cognitive Processing*, *17*(3), 269-277. Doi: https://doi.org/10.1007/s10339-016-0757-6

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

Qin, T., Burgoon, J. K., Blair, J. P., & Nunamaker, J. F. (2005). Modality effects in deception detection and applications in automatic-deception-detection. *IEEE.* Doi: 10.1109/HICSS.2005.436

Reinhard, M. A. (2010). Need for cognition and the process of lie detection. *Journal of Experimental Social Psychology*, *46*(6), 961-971. Doi: https://doi.org/10.1016/j.jesp.2010.06.002

Reinhard, M. A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology*, *44*(3), 477-488. Doi: https://doi.org/10.1016/j.jesp.2007.07.012

Ruppel, E. K., Gross, C., Stoll, A., Peck, B. S., Allen, M., & Kim, S. Y. (2016). Reflecting on connecting: Meta-analysis of differences between computer-mediated and face-to-face self-disclosure. *Journal of Computer-Mediated Communication*, *22*(1), 18-34. Doi: https://doi.org/10.1111/jcc4.12179

Skulmowski, A., & Rey, G. D. (2017). Measuring cognitive load in embodied learning settings. *Frontiers in psychology*, *8*, 1191. Doi: https://doi.org/10.3389/fpsyg.2017.01191

Sondhi, S., Vijay, R., Khan, M., & Salhan, A. K. (2016). Voice analysis for detection of deception. *11th International Conference on Knowledge, Information and Creativity Support Systems,* 1-6. Doi: 10.1109/KICSS.2016.7951455

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13*(1), 1–34. https://doi.org/10.1037/1076-8971.13.1.1

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

Tetlock, P. E., & Levi, A. (1982). Attribution bias: On the inconclusiveness of the cognition-motivation debate. *Journal of Experimental Social Psychology*, *18*(1), 68-88. Doi: https://doi.org/10.1016/0022-1031(82)90082-8

Tilley, P., George, J. F., & Marett, K. (2005). Gender differences in deception and its detection under varying electronic media conditions. Institute of Electrical and Electronics Engineers: New York City

Tsechpenakis, G., Metaxas, D., Adkins, M., Kruse, J. Burgoon, J., Jensen, M. L. … Nunamaker, J. *(2005).* HMM-Based Deception Recognition from Visual Cues. *IEEE International Conference on Multimedia and Expo,* 824-827. Doi: 10.1109/ICME.2005.1521550

Van der Kleij, R., Maarten Schraagen, J., Werkhoven, P., & De Dreu, C. K. (2009). How conversations change over time in face-to-face and video-mediated communication. *Small Group Research*, *40*(4), 355-381. Doi: https://doi.org/10.1177/1046496409333724

Vrij, A. (2000). Detecting lies and deceit. Chichester, England: Wiley

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). 'Look into my eyes': Can an instruction to maintain eye contact facilitate lie detection?. *Psychology, Crime & Law*, *16*(4), 327-348. Doi: https://doi.org/10.1080/10683160902740633

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. Doi: https://doi.org/10.1007/s10869-009-9090-8

Walton, D. (1999). Rethinking the fallacy of hasty generalization. *Argumentation*, *13*(2), 161-182. Doi: https://doi.org/10.1023/A:1026497207240

Wiener, M., & Mehrabian, A. (1968). *Language within language: Immediacy, a channel in verbal communication.* New York: AppletonCentury-Crofts.

Yu, J., Tao, Q., Zhang, R., Chan, C. C., & Lee, T. M. (2019). Can fMRI discriminate between deception and false memory? A meta-analytic comparison between deception and

false memory studies. *Neuroscience & Biobehavioral Reviews*, *104*, 43-55. Doi: https://doi.org/10.1016/j.neubiorev.2019.06.027

Zimasa, T., Jamson, S., Henson, B., Tomlinson, A., Donkor, R., & Skrypchuk, L. (2018). The Effect of Mood Valence and Arousal on Car Following; Evidence from Driving Behaviour and Eye Tracking. *KEER2018: Go Green with Emotion, 7*(146), 199-212. Retrieved from http://www.ep.liu.se/ecp/146/022/ecp18146022.pdf

Appendices

Appendix A: Informed consent

Several rights are present for you as participant:

- If you have questions after the study, a mail can be sent to l.temebel@student.utwente.nl. You will be answered as soon as possible.

- You have the right to stop the study whenever you want, without providing a reason why you want this.

- Until the data analysis has started, you have the right to receive your data and to delete data if you prefer to, without giving arguments why.

- The results will be anonymized completely. After processing the data, the anonymized data has to be stored at the University of Twente for 10 years. However, nobody except for the researchers are allowed to restore this data, which will only be done in specific situations.

- If you are interested in the end-product, you can mail l.temebel@student.utwente.nl. The final paper will be sent to you

- I am sufficiently informed about my rights and agree with the procedure of data collection and data processing described above (option that has to be selected to continue with the study).

Appendix B: Scenario

We would like to ask you to imagine the following situation:

You are a member of the examination board at the University of Twente. The psychology students had an exam two weeks ago. A teacher filed a complaint with the examination board about a student, Laura Smit, that allegedly cheated during the test. He claims to have seen her looking around during the test. In the examination board, the procedure for alleged fraud cases is that the student is not guilty, until the opposite is proven. It is known that the student was absent a week prior to the test, due to sickness. This is the first time the student has been accused of committing any type of fraud.

A week ago, the student received a letter with the request to answer some questions today. The questions asked to the student will be used to investigate whether the student cheated on the exam or not. Also, the consequences of cheating during a test are made clear to the student: she can be expelled from this examination, part of the examinations or all the examinations for a year. This expulsion applies for each university. With repeated fraud, the student's enrolment can be reversed definitively. The conversation was held via an online video platform because the exam commission is currently very busy with upcoming events. After your colleague ends the conversation, you are requested to revise the conversation between the student and your colleague.

Appendix C: questions asked and answers given to the participant

1: What is your name and study?

I am Laura/Dennis Smit and I study Psychology.

2: Do you know why you are here today?

Yes, I am here because I am accused of cheating on my exam.

3: Have you ever been suspected of exam fraud before?

No, I have not cheated on examinations before. I find it important to pass the tests with my own knowledge and competencies, so not by cheating.

4: Why do you think you have been accused of cheating?

Well, I was looking to see whether there were other people that finished the test. I think this has been misinterpreted by the examiners. I looked at the people, not at the papers.

5: Did you study for the test?

Yes, I did study, but had less time for this because I was suffering a fever. In my opinion, I learned everything, but in less detail than normal.

6: Can you tell me more about that?

As said before, I want to pass the tests myself instead of by cheating. My past grades have always been good and I have never been accused before, why would I do that now?

7: Did you perform the fraud that you are being accused of?

No, of course I did not cheat (this answer was given with the accentuation on 'of course' and by a visual 'offensive' expression).

8: Do you have anything else to add?

I hope that the truth will get out soon. I really did not do it and this situation is pretty stressful to me.

Appendix D: Image of the participant