Real or fake cries for help: A personality-based attempt to explain individual differences

in deception detection accuracy

Valentina Leona Zurhold

Student number: s1854038

Department of Psychology, University of Twente

BSc Thesis PSY 2021 – 2A; 202000384

Affiliation of Conflict, Risk & Safety

Peter Slijkhuis

Dr Steven Watson

July 22nd, 2022

Content

Abstract

This paper aims to uncover a possible connection between deception detection and human personality research. Therefore, we examined 20 undergraduate students in an online environment using a correlational study design. Our participants were presented with several video stimuli showing real people pleading for help in cases of missing loved ones. However, some of them were already identified to be deceivers. Consequently, we were able to use this material to determine our participants' ability to detect deception. In addition to that, we used a short Big Five measure to bring the field of personality research into play.

Our results indicated that there are certain negative correlations to be found between the Big Five subscales of Neuroticism or Agreeableness and the deception identification score of our subjects. In other words, once the participants' Agreeableness or Neuroticism score increases, the corresponding score of deception identification decreases. Therefore, some of the BFI dimensions seem to be of predictive value regarding the dependent variable.

Nevertheless, it remains questionable which of the Big Five dimensions are most strongly associated with the dependent variable of deception detection. This is a possible shortcoming of the relatively small sample size in combination with the correlations among the individual Big Five subscales. Thus, a comparable research design might bring about more valuable insights if the study's resources are not limited to the scope of a bachelor's assignment.

Keywords; *Big Five, Personality, Deception, Detection, Emotion*

Introduction

The ability to communicate is a key concept concerning the aspects that distinguish humans from other species, such as animals since it is an exceptional human capability (Berk, 2013, p. 359). In addition, communication is crucial in almost every domain of daily life (Berk, 2013, p. 360). Studies have shown that communication is becoming increasingly important e.g., in areas of excessive human functioning, such as medical sectors (Leonard, Graham & Bonacum, 2004). Considering the recent COVID-19 pandemic, it has been of vital importance to engage in risk communication, even political leaders have to communicate across countries to discuss subsequent steps of action in order to protect the public (Abrams & Greenhawt, 2020). However, communication is not infallible since it is a complex and developing process prone to bias and misinterpretation. These shortcomings are often a result of mismatching information through verbal and nonverbal communication channels (Kassin, Fein, Markus, McBain & Williams, 2019).

Nonverbal communication

The investigation of nonverbal communication dates back to the time of ancient Greece when a practice known as "physiognomy" - the art of judging character from facial characteristics -arose in an attempt to analyze the center of human expression and come up with claims about the individual person (Kassin et al., 2019). Despite several centuries of ever-evolving research and technological improvements, the original notion of physiognomy persists to this day (Kassin et al., 2019). As an example, studies have shown that people tend to form an often unconscious prejudgement about a person solely based on their facial features, without ever having interacted with them (Hassin & Trope, 2000). In addition, Berry

and Zebrowitz-McArthur (1986) found out, that adolescents with specific facial characteristics such as smooth skin or round eyes are more likely to be judged as kind-hearted and submissive, due to their baby-facedness, while people with more mature facial components e.g. low brows and wrinkled skin tend to be seen as more dominant and competent. These types of superficial cues can even go as far as influencing judicial settings or job interview outcomes (Kassin et al., 2019). Therefore, it is crucial to understand that each situation can be interpreted and evaluated on the basis of pre-determined schemes and judgements which are shaped through experience and observation (Kassin et al., 2019).

Furthermore, Trope (1986) argued that the context of a situation may be seen as another source of information when it comes down to the interpretation of nonverbal behavior cues. For example, multiple studies presented participants with a close-up picture of a person displaying an ambiguous facial expression (Trope, 1986). Subsequently, the subjects were told that the person in the photograph is either threatened by something, e.g., an aggressive dog or had just won a prize. As a result, the participants interpreted the given facial expression as two completely different emotions based on the context information they were presented with (Trope, 1986). Thus, an individual's behavior and physical characteristics may not only function as a starting point in terms of interpreting someone's actions but can also predict the emotional or psychological state of a person (Kassin et al., 2019).

Emotions

Speaking of emotions, the debate about the identification process of emotions dates back to the treatise of naturalist Charles Darwin (1872), who proposed an evolutionary approach to understanding emotions. In his work *The Expression of the Emotions in Man and Animals* he claimed that the human face presents the observer with a set of emotions that are intrinsic and universally understood across cultures (Darwin, 1872). This notion is

substantially supported by contemporary research to this day (Kassin et al., 2019). Furthermore, several researchers agreed on approximately six basic emotions namely: happiness, sadness, anger, fear, surprise and disgust (Ekman et al., 1987; Levenson, Ekman, Heider & Friesen, 1992). These emotions have been recognized in people and cultures all over the world (Ekman et al., 1987). In addition, Ekman (1993; 1999) assumed that our feelings are often accompanied by specific patterns of facial muscle activity in an attempt to produce a facial expression.

Nevertheless, many researchers, especially in the field of social constructivism are skeptical of Ekman's findings since he heavily relies on evolutionary perspectives, as well as biology (Ekman, 1993, 1999). For example, Ekman claims that our perception of a setting or event which results in the production of emotion is influenced by our developmental – ontogenetic - and our evolutionary – phylogenetic – past (Ekman, 1993, 1999). Further, he states that the onset of emotional display and expression already starts in infancy through the attachment style of toddlers (Ekman, 1993, 1999). Opposingly, Boiten (1996) proposed that emotions are socially learned, and people are taught to react in a certain way when presented with a related event by culturally distinctive social practices. Even though the debate about the actual evolution of emotions carries on until the present day, it is evident that understanding them is a vital process considering the social nature of human beings.

Another important facilitator of social interaction emerged in the fields of neurology and psychophysiology, also known as mimicry. Mimicking the facial expressions of others is an automatic and unconscious process starting in infancy (Meltzoff & Moore, 1977; Hess & Fischer 2013, 2014). According to Sato et al. (2004), facial mimicry can greatly facilitate social interaction by recognizing and understanding the opposite's emotions. As an example, toddlers tend to start crying as a response to another baby crying (Meltzoff & Moore, 1977). However, mimicking is not omnipresent by nature since it is not an inevitable process. Even

though it may not be possible for a person to decide whether they would like to mimic someone or not an individual is still capable of actively suppressing the facial muscle movements to be mimicked (Krumhuber, Likowski & Weyers, 2014). This occurrence is used as a critical instrument of distinction between real and fake emotions in criminology (Krumhuber, Likowski & Weyers, 2014).

Deception

As an immediate response to the 9/11 terrorist attacks, the detection of false emotions and malicious intent gained a tremendous amount of attention (Kassin et al., 2019). As an example, many public transportation companies made efforts to train their security staff in an attempt to observe the facial expressions of passengers and identify possible signs of hostility (Kassin et al., 2019). Nevertheless, the detection of lying behavior remains an extremely difficult process, which is why these efforts did not bring about as much effect as expected (Kassin et al., 2019). More specifically, although many years of fundamental research in this field have passed, the chance of uncovering a lie is still relatively low when comparing the performance of lay people to that of professionals (Ekman & O'Sullivan, 1991; Meissner & Kassin, 2002; Vrij, 2008). For example, Bond and DePaulo (2006) found out that the mean probability of detecting a lie, whether trained or lay personnel, is comparable to that of chance.

First attempts to detect and identify mechanisms related to deception are originally based on the workings of the founding father of psychoanalysis, Sigmund Freud, who claimed that "no mortal can keep a secret [...] (1905, p. 94). This notion has been further investigated by Paul Ekman who continued his research on the topic of emotions and deception (Ekman, 1993).

As a result, Ekman, Roper and Hager found out that specific facial muscles are activated when a person is experiencing a genuine emotion (1980). However, these muscle movements are hardly replicable if the same emotion is faked since it is more difficult to control something that is predominantly part of an automatic process (Ekman et al.., 1980; Meltzoff & Moore, 1977; Hess & Fischer 2013, 2014). Thus, Ekman (1985) coined the term "reliable" muscles to refer to facial muscle activity that is most likely to be absent when an emotion is expressed but not truly felt.

Apart from Ekman and colleagues, many other researchers have investigated these signs of deception which led to the formation of the inhibition hypothesis. This hypothesis is based on the assumption that the real emotions of an individual are revealed at some point but quickly suppressed or inhibited (Ekman, 2003; Duchenne 1862/1990; Darwin, 1872). This phenomenon is also known as leakage or emotional leakage (Ekman, 2003).

In addition, Ekman and Friesen (1969) proposed a related theory by focusing on so-called micro-expressions, which are approximately 1/5th of duration in comparison to regular facial expressions (1969). Micro-expressions are described as brief accounts of the real emotional load breaking through the surface of the deceiver until they are swiftly suppressed (Ekman & Friesen, 1969). Contrary to the discovery of Ekman and Friesen, ten Brinke and Porter (2012) claimed that further investigation regarding the hypotheses of emotional leakage through micro-expressions would be necessary to support the notion since the vast complexity of facial expressiveness is still relatively unexplored.

According to ten Brinke et al. (2012), this critique also holds for the inhibition hypothesis because it is claimed to have major theoretical and applied implications. Secondly, ten Brinke et al. emphasize that the detection of micro-expressions is still relatively difficult, due to their abbreviated nature, which also calls for more extensive expertise on behalf of the observer (tenBrinke et al., 2012).

Deception and Personality

Interestingly, in a study by Levitan et al. (2015) it has been found that those individuals who are good at detecting lies are also better at deceiving others. These findings raise the question of whether there are indeed certain personality types to be identified as promising experts in the field of lie detection. This question has been the center of attention for many researchers since Maureen O'Sullivan and Paul Ekman published their paper *"The wizards of deception detection"* (2004). In their work, the two administered three different lie detection tasks to their participants in hopes of identifying individuals that were exceptionally good at detecting deception. As a result, their study revealed a small proportion of people with accuracy levels above 80 or even 90 percent dependent on the difficulty of the task (O'Sullivan & Ekman, 2004). Despite their extraordinary findings, it is yet to be investigated whether specific personality characteristics may be an underlying factor in this case.

Taking a closer look at personality research it is evident that the Big Five Personality Model is received extensive amounts of attention and support throughout the last two decades (Costa & McCrae, 1995; Goldberg, 1981; McCrae & John, 1992; Saucier & Goldberg, 1996). The Big Five Inventory (BFI) has originally been developed in the 1980s as a product of collaboration between several researchers (Costa & McCrae, 1995; Goldberg, 1981; McCrae & John, 1992; Saucier & Goldberg, 1996). Due to its nature of respectable replicability, the BFI is seen as a personality trait taxonomy that brought about a great degree of consensus (Goldberg, 1981, 1990; John, Naumann & Soto, 2008; McCrae & Costa, 2008).

Although it took a lot of effort as well as constant critique and reconsideration, a host of researchers agreed on five broad personality dimensions, namely: extraversion, agreeableness, conscientiousness, emotional stability and openness, that construct the core of what is now known as the BFI (Costa & McCrae, 1995; Goldberg, 1981; McCrae & John,

1992; Saucier & Goldberg, 1996, Block, 1995; McAdams, 1992). According to Goldberg (1990), these traits embody a variety of characteristics. For example, extraversion describes people that are talkative and sociable. Agreeableness holds for individuals that are sympathetic and kind. Conscientiousness is associated with being orderly and organized (Goldberg, 1990). Emotional stability is self-explanatory but may also be measured by its counter construct also known as "Neuroticism" which describes the absence of emotional stability e.g., insecurity and anxiety (Goldberg, 1990). And lastly, Openness describes someone's creativity and also correlates with intellect (Goldberg, 1990).

As an example of what a possible link between personality research and deception detection might look like, a study by Hirschberg et al. (2006) indicated the following results. In their study design, Hirschberg et al. (2006) compared the lie detection capabilities of humans to those of specialized machines and assessed their participants' personalities using a Big Five Inventory. As a result, they found that the BFI dimensions Openness and Agreeableness appear to be positively correlated with the accuracy of a participant's veracity judgement. In addition to that, Levitan et al. (2015) discovered a significant negative correlation between a subject's Neuroticism score and accurate lie detection.

Current research

In an attempt to add something to the ever-evolving debate about human emotion, this paper aims at combining personality research as illustrated by the Big Five personality model with the fundamental workings regarding deceptiveness and lying behavior. Inspired by the *"Cry Me a River"*-paper by tenBrinke and Porter (2012), we strive to assess our participant's lie detection capability with the help of video stimuli. The video material shows real people displaying either fake or real emotions and should function as a medium that combines the

different communication channels. Therefore, we intend to capture the full complexity of lying behaviour.

In addition to that, based on the research of Hirschberg et al. (2006), it is hypothesized that there is a significant relationship between the dimension of Openness, as measured by the Big Five model, and the accuracy of deception identification. This may further be supported by the correlations of intelligence and Openness according to Goldberg (1990). Furthermore, we speculate that there may be a negative association between accurate lie detection and Neuroticism, as depicted by Levitan et al. (2015). Lastly, we assume that participants who score high on Agreeableness tend to be worse in terms of correctly identifying lying behaviour. This hypothesis may be explained by the fact that most people tend to perform poorly when asked to look out for lies because they are biased in the sense of acknowledging what others are saying at face value (Bond & DePaulo, 2006; Vrij, 2008). Therefore, further investigation is needed to examine whether people that are highly agreeable and often tend to avoid conflict will indeed perform worse on deception detection tasks.

Thus, the central question of this research paper reads as follows; "To what extent do individual differences in specific personality traits based on the Big Five personality dimensions influence the probability of detecting deception?".

Methods

Participants

The current study focuses on the target group of undergraduate students due to their status of accessibility. Therefore, we used the method of convenience sampling. The research was approved by the Ethics Committee of the University of Twente and an online consent

form has been included. All of the 20 participants (*M*age=24 , *SE*age=.9, 13 [*65%]* female, 7 [*35%]* male,) were provided with an extensive online information sheet prior to giving or declining their consent. The students were compensated for their effort by receiving sona system credits regardless of their state of consent.

Inclusion and exclusion criteria

Since our sample focused on undergraduate students it has been an explicit requirement to be employed as a student. In addition to that, participants who abandoned the study at a certain point or leave out several answers were excluded from the dataset. Therefore, a total of 17 of the 37 total participants needed to be excluded. We suppose that this particularly high amount of drop-outs is not solely caused by abandonment but rather by a problem with the internet browser since the study did not work for Safari-users.

Procedure

The research took place in an online environment. In the beginning, the participants were provided with an information sheet about the topic and its risks as well as a consent form. If they would choose to decline consent the study would automatically redirect them towards the end.

After receiving consent, the participants were asked to fill out several demographic questions related to their age, gender and profession. We included the profession-related question because we originally planned to focus on two different populations. On the one hand, the students which would make up the "lay personnel" and on the other hand police officers or other judicial staff members. This would have enabled us to compare the deception identification scores among the two groups. However, we did not include this condition in our

final set-up since the data collection brought about errors that lead to high percentages of participant drop-out.

Subsequently, the subjects were redirected to a platform called "iMotions" which would video-record their faces for the next sequence of the study. Therefore, a working web camera has been set as a requirement. It is however not known whether some students were excluded because of this prerequisite. The original goal of using this technology was to scan the participants' facial activity while being presented with specific video footage. However, the facial expression analysis has not been completed due to the given problems in the data collection process and will therefore not be further investigated in this paper.

As a next step, the participants were presented with specific video footage and provided with three questions for each video. During the process of answering these questions, the participants had the opportunity to refresh their memory and rewatch the video that they were evaluating at the time.

Subsequently, we included a short Big Five Inventory to scan for individual differences in terms of specific character traits. After completing this questionnaire, the participants were offered the option to be debriefed to fully understand what this study is about. If they finished reading through the information or chose not to receive any, their data was uploaded to the iMotions server and the procedure has been completed.

Instruments

The Big Five Inventory we selected is known as the BFI-10 (Rammstedt & John, 2007). It measures the five core dimensions of the personality model. Therefore, it consists of the following subscales: Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness. Each of these constructs is assessed by two short items on a five-point Likert scale.

As an example, the first item of the Extraversion scale reads as follows "I see myself as someone who is reserved" (See Appendix C). The scores range from 1-5 with 1 = "Strongly disagree" and 5 = Strongly agree".

Solely the characteristic of Agreeableness is measured by one additional item. This optional element has been included by the creators of the BFI-10 (Rammstedt & John, 2007) and is implemented in this study due to special interest in the Agreeableness score. For each of the five components one item is reverse scored and the other true scored.

The BFI-10 possesses moderate psychometric qualities (Rammstedt & John, 2007). As an example, Rammstedt & John (2007) compared the BFI-10 with a more extensive 44-item Inventory (9-items per scale), known as the "BFI-44". Their results suggest a mean part-whole correlation of 0.83 ranging from 0.74 for Agreeableness up to 0.89 for Extraversion. In other words, the BFI-10 is capable of predicting nearly 70% of the variance of the BFI-44 scale (Rammstedt & John, 2007). For a more detailed review see Appendix D.

Video Evaluation

To bring the variable of deception detection into play, the research design includes specific video material. A total of 7 different videos (see Appendix B) which varied in duration from under a minute up to approximately four minutes were utilized in the study. They were randomized in order and the participants were not able to stop or skip them at any point in time.

The video excerpts portrayed real people that claimed to plead for help to receive support from the public regarding the safe return of a missing loved one. However, some of the recorded individuals were trying to deceive the viewer in the sense of acting out certain types of emotion. This has been the case for 4 of the seven video stimuli. Thus, we were

provided with additional information regarding the truthfulness of the video protagonists by tenBrinke and Porter (2012) to utilize the video material as a measure of deception detection (See Appendix B). Their information has also been verified by the official news sources who originally published the video material.

The actual measurement of our participants' lie detection capability took the form of three questions per video (See Appendix B). Question 1 reads as follows: "To what extent are the emotions as displayed in the video authentic?" using a scale from 0 = not authentic at all to 100 = very authentic. The second question focuses on the construct of empathy and asks: "To what extent did you feel empathy for the people/person in the video?", the response is recorded using a 7-point Likert scale with a score from 1-7. The lowest score of 1 = no empathy at all and 7 = high amounts of empathy. The last question is the most important one since it focuses on the deceptiveness of the recorded individuals. Therefore, it reads as follows: "Do you think the people in the video were trying to deceive the viewer?". This question is answered by either a "Yes" or "No"-response.

These questions were the same for each of the video stimuli and assessed three different constructs. Since all videos are presented in a row, accompanied by the same three questions for each stimulus, we carried out an ANOVA to check for variance between videos. The ANOVA contains the participants' score of correct deception identification (0-7) as the dependent variable and the repetition factor of the video stimuli as the independent variable. The ANOVA has been assessed for each of the three questions. Therefore, it should indicate whether there is an actual difference between videos or if every video has been judged equally. The first question focused on the authenticity of the people's emotions. According to Mauchly's $W(20) = 0.11$, $p = .010$ the assumption of sphericity is not met for question item 1. In this case correction is needed, therefore Huynh-Feldt = 0.86 $F(6, 114) = 3.53$, $p = .005$, which indicates that the videos were evaluated differently for the first question.

The second question item aimed at the amount of empathy that our subjects felt for the people in the video. According to Mauchly's $W(20) = 0.25$, $p = .290$, the assumption of sphericity has been met. In this case correction is not needed, therefore $F(6, 114) = 4.61$, $p \leq .001$, which indicates that the videos were also evaluated differently for the second question.

Lastly, the final video question concerns the actual deceptiveness of the recorded individuals. According to Mauchly's $W(20) = 0.10$, with $p = .008$, the assumption of sphericity is not met. Thus, the Huynh-Feldt $= 0.86$ correction is needed $F(6, 114) = 3.11$, $p = .011$. This suggests a significant effect among the different video stimuli for question item 3 as well.

Thus, all three of the video questions show a statistically significant variance between the videos which is essential for further analysis. In other words, the videos and their corresponding questions were not answered equally for each video.

Originally, the study consisted of three different parts which were put together in the process. As a starting point, two distinct Qualtrics surveys were designed. The first one contained the demographic information as well as the consent form (See Appendix A) and was therefore included at the beginning of the study. The second one embodied the corresponding video questions and the Big Five Inventory. This questionnaire was inserted at the end of the survey since the participants had to watch the video material to answer related questions.

Solely the video footage had to be included on a different platform using iMotions. This software was originally selected due to its capability of measuring and analyzing facial expressions by making use of the participants' web camera. Thus, the participants were recorded while being presented with our video footage. However, the iMotions data were not used in the end, since a facial expression analysis would have been too complex for a sample size of 20 participants.

Data Analysis

The data has been collected by two different softwares. Firstly, the Qualtrics software which enabled the creation of the given surveys and secondly the iMotions platform which functioned as a tool to implement the video material. In addition, the iMotions data has been checked in terms of quality by the program itself which relied on the positioning of the subject's head during the entire process of presenting the video stimuli. Subsequently, the data were stored on a separate iMotions server.

Variables

To investigate the given research question and corresponding hypotheses the following quantitative variables were tested. On the one hand, the subjects' capability of detecting false emotions, which is measured in the form of a score of 1-7 *(1 point for each correct deception identification for a total of 7 videos)*. This value may be defined as the dependent variable in this study. On the other hand, the participant's scores on the BFI-10 Inventory which is the independent variable.

Statistical Software and Testing

To analyze the given information, a dataset has been created by using the recorded responses obtained by Qualtrics, excluding the data recorded by iMotions. Every given step has been executed utilizing the IBM SPSS Statistics software 28th edition.

As a starting point, we computed mean scores of each Big Five subscale as well as those of the correct deception identification score to get a general overview of the data.

Furthermore, the data has been analyzed using a ANOVA to investigate whether there is an actual measurable effect between the questions concerning each video stimulus. This analysis has been carried out for each of the three corresponding video questions which is already included in the Instruments section of this paper.

To look at the possible relationship between a subject's Big Five score and the score of correct deception identification, the data had to be checked for normal distribution since a decision between Spearman's Rho and Pearson's *r* had to be made. Therefore, a Shapiro-Wilk test and a Kolmogorov-Smirnov test have been computed as well as the Skewness and Kurtosis values of our variables. Thus, a bivariate analysis has been carried out including Spearman's correlation coefficient.

Lastly, to look at the predictive value for each of the five dimensions of the independent variable a regression analysis has been executed.

Results

As a starting point for further analysis, the outcomes of the BFI-10 and the scores of the dependent variable of correct deception identification were calculated and summarized in the tables below (see Table 1&2). In addition to that, the data has been checked for normal distribution using Skewness and Kurtosis as well as Shapiro-Wilk and Kolmogorov-Smirnov. Out of these parameters all of them indicate that the dimension of Agreeableness shows a non-normal distribution. However, not all the given statistical tests need to be significant to indicate a deviation from normality. Looking at the dimension of Conscientiousness, it is evident that even though the Skewness value of 0.97 seems to be relatively normal, the other three parameters suggest something different (See Table 1). Furthermore, the BFI-subscale of Agreeableness possesses a Skewness value of -1.88 and a Kurtosis value of 4.98. In other

words, a Skewness greater than its cut-off score of 1 suggests a substantially skewed distribution of the variable of interest (Hair, Hult, Ringle & Sarstedt, 2022). Whereas a Kurtosis value greater than the corresponding cut-off estimate of 1 suggests that the distribution of Agreeableness scores is heavily peaked in this case (Hair et al., 2022). In addition, the significance levels of the Shapiro-Wilk with $p \leq .001$ and Kolmogorov-Smirnov with $p=.018$, seem to be statistically significant as well, which further supports the assumption of a non-normal distribution. Interestingly, the Kolmogorov-Smirnov test seems to be significant for each variable in our study design.

Furthermore, the dependent variable of correct identification (of deception) also suggests a statistically significant deviation from normality (see Table 1).

Table 1

Normal parameters and Assumption checks of BFI and Correct Identification

	Mean	*SD*	*Skewness (SE)*	*Kurtosis (SE)*	*Kolmogorov-Smirnov*	*Shapiro-Wilk*
Extraversion	3.20	0.15	0.19 (0.51)	-0.86 (0.99)	**.007**	.098
Agreeableness	3.40	0.14	**-1.88 (0.51)**	**4.98 (0.99)**	**.018**	**<.001**
Conscientiousness	3.30	0.14	0.97 (0.51)	**1.33 (0.99)**	**.015**	**.024**
Neuroticism	3.20	0.16	-0.23 (0.51)	0.31 (0.99)	**.024**	.124
Openness	3.83	0.21	-0.38 (0.51)	-0.84 (0.99)	**.011**	.144
Correct Identification	3.50	1.19	0.52 (0.51)	**1.05 (0.99)**	**<.001**	**.012**

Note. *Bold-printed parameters show deviation from normality

Table 2

Frequencies of correct deception identification

	Correct Identification	
Score	Frequency	Percentage (%)
1	1	5
2	1	5
3	10	50
4	5	25
5	1	5
6	2	10
7	0	0
Total	20	100

Note. *The total scores range from 1-7, one point for each correct answer with a total of 7 videos

Therefore, the assumption of normal distribution is not met. Given the non-normal distribution, Spearman's Rho is more appropriate than Pearson's *r* since it has a higher level of accuracy in this case. Consequently, a bivariate correlation analysis has been executed using Spearman's rank order coefficient.

The analysis indicates statistically significant correlations among certain BFI-subscales (see Table 3). As a result, the dimension of Neuroticism seems to positively correlate with that of Extraversion, $r(18) = .53$, $p = .017$. In addition, the variables of Agreeableness and Neuroticism seem to be even more strongly associated, $r(18) = .59$, $p = .006$.

In contrast, the analysis suggests a statistically significant negative correlation between certain BFI dimensions and the given dependent variable which is the score of correct identification of deception (See Table 3). For example, the Agreeableness score seems

to be strongly associated with the dependent variable, $r(18) = -.68$, $p \leq .001$. In other words, if a participant's score on Agreeableness increases, the corresponding correct identification score tends to decrease. This assumption also applies to the subscales of Conscientiousness and Neuroticism in relation to the dependent variable (see Table 3).

Table 3

Correlation coefficients using Spearman's Rho

	1	2	3	4	5
1. Extraversion					
2. Agreeableness	0.14				
3. Conscientiousness	0.32	0.21			
4. Neuroticism	**0.53**	**0.59**	**0.46**		
5. Openness	0.04	-0.06	0.30	0.13	
6. Correct Identification	-0.28	**-0.68**	**-0.46**	**-0.63**	-0.35

Note. *Bold-printed parameters hint at a statistically significant correlation

As a last step in the data analysis, a regression has been executed. This analysis should function as a tool to examine whether the Big Five dimensions are capable of predicting a shift in the dependent variable. As a result, the ANOVA indicates a statistically significant connection, $F(5, 14) = 3.16$, $p = .041$. This indicates that the BFI subscales may indeed possess some predictive value regarding the correct identification score of our participants. However, the corresponding p-values of the unstandardised regression coefficients are not statistically significant for the given variables (see Table 4). Therefore, it is crucial to take the collinearity statistics into account.

Consequently, it is evident that all the BFI dimensions have a variance inflation factor (VIF) greater than 1 (See Table 4). This indicates a moderate correlation between variables which underlines the results of Spearman's Rho. Looking at the variable of Agreeableness with a VIF-value of 2.04, this would mean that the variance of a certain regression coefficient is 104% bigger than it would be if no collinearity is present. Consequently, it is questionable whether the given regression model is powerful enough to significantly predict the correct identification score based on all Big Five dimensions taken together. Thus, it is difficult to distinguish in this case, which of the independent variables accounts for the change in the dependent variable, since some of the BFI subscales are highly correlated. This could also explain the non-significant p-values of the given regression coefficients B.

Table 4

Regression coefficients and VIF for BFI dimension

				Collinearity Statistics
	Regression coefficient B	*SD*	*Sig.*	*VIF*
Extraversion	0.03	0.39	0.95	1.48
Agreeableness	-0.73	0.49	0.16	2.04
Conscientiousness	-0.25	0.43	0.57	1.56
Neuroticism	-0.40	0.53	0.46	3.17
Openness	-0.38	0.25	0.15	1.10

Discussion

The goal of this study was to shed light on a possible connection between deception detection and personality research. Therefore, several hypotheses were investigated to answer the research question which reads as follows; "To what extent do individual differences in specific personality traits based on the Big Five personality dimensions influence the probability of detecting deception?".

The corresponding assumptions focused on individual characteristics as measured by the BFI-10. Therefore, it has been hypothesized that a participant's score on the subscale of Openness has an influence on that subject's score of deception identification. The results of our data analysis indicate that there seems to be a negative relationship between the independent variable of Openness and the dependent variable of deception identification. However, this association does not seem to be statistically significant (See Table 3).

In addition to that, we hypothesized that there might be a negative association between the Neuroticism score and that of deception identification. This assumption is supported by our analysis which indeed indicates a statistically significant negative relationship between Neuroticism and deception identification (See Table 3).

Lastly, it has been presumed that people who score higher on Agreeableness are worse at identifying deception. This hypothesis is also supported by the bivariate correlation analysis (See Table 3).

Therefore, the statistical analysis supports the notion that once a participants score on Big Five dimension Agreeableness or Neuroticism increases, the corresponding score on deception identification decreases.

Main findings

The given research is based on a correlational design which has been the main focus. Therefore, the bivariate analysis indicates that some of the BFI dimensions indeed correlate with a participant's score on deception identification. Looking at the subscale of Openness, we supposed that there might be a positive relationship, since Openness has been found to correlate with intelligence as well (Costa & McCrae, 1995). This is however not the case in this study specifically. Therefore, we reject this hypothesis.

Furthermore, we investigated the relationship between Neuroticism and deception detection. This relationship turned out to be statistically supported and of negative nature. Thus, we accept our hypothesis regarding the negative influence of Neuroticism on our subjects' deception identification. It is presumed that neurotic individuals perform worse on deception detection tasks due to their heightened negative perception (Hirschberg et al., 2006). However, further research is needed to replicate this finding and uncover the reasoning behind it.

Lastly, the association between a subject's score on the subscale of Agreeableness and that of deception identification has been statistically supported. This correlation is of negative nature as well. Thus, we accept the corresponding hypothesis. This assumption may have been the most promising in this paper. Looking at frequently made mistakes when trying to identify deceivers, it is evident that many people are biased in the sense of taking the statements of other people at face value (Bond & DePaulo, 2006; Vrij, 2008). Therefore, people tend to assume that their opposite is telling the truth, rather than questioning his or her honesty (Bond & DePaulo, 2006; Vrij, 2008). In addition to that, people that are highly agreeable usually tend to avoid conflict which would logically further support the given assumption (Costa & McCrae, 1995). This finding is also in line with those of Hirschberg et al. (2006).

Furthermore, it is evident that the statistical analysis not only indicates correlations between single Big Five dimensions and the deception identification score but also among the BFI subscales themselves. As an example, the characteristics of Agreeableness and Extraversion seem to be highly correlated with that of Neuroticism (See Table 3). While this is not an uncommon phenomenon regarding the Big Five taxonomy or the BFI-10 specifically (Rammstedt & John, 2007), it does have an influence on the power of the given regression analysis. Consequently, the regression model shows that the BFI possesses some predictive value regarding the dependent variable. However, the analysis is not able to identify which of the given independent variables has the strongest predictive power. This may be the case, due to the multicollinearity among the BFI subscales (See Table 3) accompanied by the small sample size. In addition to that, the variance inflation factors (VIF) of the BFI might appear particularly normal in the analysis, but they may still be too high given the small number of participants (See Table 4).

Study Limitations

The given study was originally planned to focus on a more extensive sample than that of the current research. However, the study's online procedure brought about great advantages e.g. in terms of time and effort, but also lead to a greater error. This error has been detected in the middle of data collection since many subjects reported that they were not able to take part in the study using a specific internet browser. Therefore, participants using the "Safari"- browser were not able to take part in the study, which resulted in particularly high drop-out rates. Thus, the original notion of the study being centred around a more extensive mimicry analysis had to be changed to fit a smaller sample size which is why several conditions were excluded from the analysis. As a result, it is evident that the relatively small number of participants influences the power of our statistical analysis, especially regarding the

regression model. Thus, we were not able to statistically identify which of our predictors is indeed the most powerful one considering the change of the dependent variable.

Furthermore, due to the original focus on mimicry and facial expression analysis, a shorter Big Five Inventory has been selected for implementation. Therefore, it is crucial to be aware of the shortcomings of smaller Big Five scales. As an example, the minimal number of items may not be able to adequately represent the measured construct (Clark & Wilson, 1993; Yarkoni, 2010). As a consequence, the psychometric qualities of the Inventory may suffer in terms of reliability and predictive validity (Carvalho, Nunes, Primi & Nunes, 2012; Credé, Harms, Niehorster & Gaye-Valentine, 2012). Thus, it is of great importance to take the complexity of measuring human personality into account.

Nevertheless, the creators of the BFI-10 have proven the usefulness of their Inventory, especially in research settings like ours, where time and resources are limited (Rammstedt, B., & John, 2007). In addition, Rammstedt and John (2007) compared the BFI-10 scale to a more extensive Big Five Inventory with 44 items (BFI-44). As a result, they found out that the BFI-10 is indeed able to portray 70% of the BFI-44 variation and maintain 85% of its retest reliability (Rammstedt & John, 2007) (See Appendix D). Therefore, it remains a matter of resources when considering the advantages and shortcomings of abbreviated measurements.

Directions for future research

Lastly, despite the occurrence of unforeseen circumstances which increased the difficulty of executing this research, the study was still able to bring about a direction for future investigation. For example, it could be of great value to stick to the original study design following tenBrinke and Porter (2012) and engage in facial expression analysis. In addition, personality research is a complex undergoing that possibly benefits from a greater

research scope and more extensive time resources. Thus, there may still be a possible link between deception and personality research yet to be revealed in the future.

In conclusion, the current research attempted to uncover a potential connection between deception detection capability and specific personality characteristics. The statistical analysis suggests that there is a statistical significance regarding this assumption. Thus, the given research question is answered as follows. Individual differences in personality traits based on the Big Five personality dimensions do influence the probability of detecting deception. However, it is crucial to consider that these findings are not generalizable among the BFI dimensions but solely apply to the subscales of Agreeableness and Neuroticism for this study design specifically.

References:

Abrams, E. M., & Greenhawt, M. (2020). Risk communication during COVID-19. *The Journal of Allergy and Clinical Immunology: In Practice*, *8*(6), 1791-1794. https://doi.org/10.1016/j.jaip.2020.04.012

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs*, *47*(1), i. https://doi-org.ezproxy2.utwente.nl/10.1037/h0093360

Berk, L. E. (2013). *Child Development* (ninth edition). Pearson Educacion.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological bulletin*, *100*(1), 3. https://doi-org.ezproxy2.utwente.nl/10.1037/0033-2909.100.1.3

Boiten, F. (1996). Autonomic response patterns during voluntary facial action. *Psychophysiology*, *33*(2), 123-131. https://doi-org.ezproxy2.utwente.nl/10.1111/j.1469-8986.1996.tb02116.x

Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and social psychology Review*, *10*(3), 214-234. https://doi-org.ezproxy2.utwente.nl/10.1207%2Fs15327957pspr1003_2

Block, J. (1995). Going beyond the five factors given: Rejoinder to Costa and McCrae (1995) and Goldberg and Saucier (1995).

Carvalho, L. D. F., Nunes, M. F. O., Primi, R., & Nunes, C. H. S. D. S. (2012). Unfavorable evidence for personality assessment with a 10-item instrument. *Paidéia (Ribeirão Preto)*, *22*, 63-71. https://dx.doi.org/10.1590/S0103-863X2012000100008

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological assessment*, *7*(3), 309. https://doi.org/10.1037/14805-012

Costa Jr, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment*, *64*(1), 21-50. https://doi-org.ezproxy2.utwente.nl/10.1207/s15327752jpa6401_2

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of personality and social psychology*, *102*(4), 874. http://dx.doi.org/10.1037/a0027403

Darwin, C. (1965). 1872. The Expression of Emotions in Man and Animals. *Chicago. University of Chicago Press*.

Duchenne (de Boulogne) GBA. Mécanisme de la physionomie humaine, ou analyse éléctro-physiologique de l'expression des passions, applicable à la pratique des arts Plastiques. One volume and a photographic atlas. Paris: Renouard, 1862. [2nd edition, Paris: Baillière, 1876; English translation by Cuthbertson RA. The mechanism of human facial expression, New York: Cambridge University Press, 1990]; 7.

Ekman, P. (1985). *Telling lies. Clues to deceit in the marketplace, marriage and politics.* New York: Norton (Paperback ed., Berkeley books, New York, 1986)

Ekman, P. (1993). Facial expression and emotion. *American psychologist*, *48*(4), 384. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0003-066X.48.4.384

Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, *98*(45-60), 16.

Ekman, P. (2003). Darwin, deception, and facial expression. *Annals of the New York Academy of sciences*, *1000*(1), 205-221. https://doi-org.ezproxy2.utwente.nl/10.1196/annals.1280.010

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *semiotica*, *1*(1), 49-98. https://doi-org.ezproxy2.utwente.nl/10.1515/semi.1969.1.1.49

Ekman, P., Friesen, W. V., O'sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., ... & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0022-3514.53.4.712

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar?. *American psychologist*, *46*(9), 913. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0003-066X.46.9.913

Freud, S. (1905). On psychotherapy. *Standard edition*, *7*(267), 64-145.

Global Deception Research Team. (2006). A world of lies. *Journal of cross-cultural psychology*, *37*(1), 60-74. https://doi-org.ezproxy2.utwente.nl/10.1177%2F0022022105282295

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), Review of personality and social psychology (Vol. 2, pp. 141 -165). Beverly Hills, CA: Sage

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0022-3514.59.6.1216

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). 3rd Ed.* Thousand Oaks, CA: Sage

Hassin, R., & Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of personality and social psychology*, *78*(5), 837. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0022-3514.78.5.837

Hartwig, M., & Bond Jr, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological bulletin*, *137*(4), 643. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/a0023589

Hess, U., & Fischer, A. (2013). Emotional mimicry as social regulation. *Personality and social psychology review*, *17*(2), 142-157.

https://doi-org.ezproxy2.utwente.nl/10.1177%2F1088868312472607

Hess, U., & Fischer, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social and personality psychology compass*, *8*(2), 45-57. https://doi-org.ezproxy2.utwente.nl/10.1111/spc3.12083

Hirschberg, J. B., Enos, F., Benus, S., Cautin, R. L., Graciarena, M., & Shriberg, E. (2006). Personality factors in human deception detection: Comparing human to machine performance. https://doi.org/10.7916/D85146PW

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues.

Kassin, S., Fein, S., Markus, H. R., McBain, K. A., & Williams, L. (2019). *Social Psychology Australian & New Zealand Edition*. Cengage AU.

Küntzler, T., Höfling, T. T. A., & Alpers, G. W. (2021). Automatic facial expression recognition in standardized and non-standardized emotional expressions. *Frontiers in psychology*, *12*, 1086. https://doi.org/10.3389/fpsyg.2021.627561

Krumhuber, E. G., Likowski, K. U., & Weyers, P. (2014). Facial mimicry of spontaneous and deliberate Duchenne and non-Duchenne smiles. *Journal of Nonverbal Behavior*, *38*(1), 1-11.

Leonard, M., Graham, S., & Bonacum, D. (2004). The human factor: the critical importance of effective teamwork and communication in providing safe care. *BMJ Quality & Safety*, *13*(suppl 1), i85-i90. http://dx.doi.org.ezproxy2.utwente.nl/10.1136/qshc.2004.010033

Levenson, R. W., Ekman, P., Heider, K., & Friesen, W. V. (1992). Emotion and autonomic nervous system activity in the Minangkabau of west Sumatra. *Journal of personality and social psychology*, *62*(6), 972.

Levitan, S. I., Levine, M., Hirschberg, J., Cestero, N., An, G., & Rosenberg, A. (2015). Individual differences in deception and deception detection. *Proceedings of COGNITIVE*, 331-348.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of personality*, *60*(2), 329-361. https://doi-org.ezproxy2.utwente.nl/10.1111/j.1467-6494.1992.tb00976.x

McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of personality and social psychology*, *56*(4), 586. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0022-3514.56.4.586

McCrae, R. R., & Costa Jr, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.4135/9781849200462.n13

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, *60*(2), 175-215. https://doi-org.ezproxy2.utwente.nl/10.1111/j.1467-6494.1992.tb00970.x

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and human behavior*, *26*(5), 469-480.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75-78. https://doi.org/10.1126/science.198.4312.75

O'Sullivan, M., & Ekman, P. (2004). 12 The wizards of deception detection. *The detection of deception in forensic contexts*, 269.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality*, *41*(1), 203-212. https://doi-org.ezproxy2.utwente.nl/10.1016/j.jrp.2006.02.001

Sato, W., Yoshikawa, S., Kochiyama, T., & Matsumura, M. (2004). The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *Neuroimage*, *22*(2), 1006-1013. https://doi-org.ezproxy2.utwente.nl/10.1016/j.neuroimage.2004.02.030

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model.

Ten Brinke, L., & Porter, S. (2012). Cry me a river: identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior*, *36*(6), 469. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/h0093929

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological review*, *93*(3), 239. https://psycnet-apa-org.ezproxy2.utwente.nl/doi/10.1037/0033-295X.93.3.239

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

Westerhof, G. J., Bohlmeijer, E. T., & van den Berg, S. M. (2014). The Individual.

Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of research in personality*, *44*(2), 180-198. https://doi-org.ezproxy2.utwente.nl/10.1016/j.jrp.2010.01.002

Appendix A

Online informed consent

Please read the following instructions carefully and do not proceed to the next question unless you are fully aware of the subject and its consequences!

Dear participant,
thank you for taking your time to participate in this study, enclosed you will find important information on the study procedure and content.
Do not hesitate to reach out to me if you have any questions regarding the research at hand.

Contact details of researcher: Valentina Leona Zurhold e-mail address: **v.zurhold@student.utwente.nl**

address: Klosterstraße 4, 48599 Gronau, Germany mobile phone: +49 157 75313442

Study procedure: In the following you will be presented with our consent form before actually starting to participate in our study. It is possible to withdraw your consent or even stop completing the survey at any moment in time without specifying any reason for it. If you do so, we will not use the data you provided and delete your information as soon as possible.

If you consent to taking part in our study, you will be asked to specify some demographic information in the beginning e.g. your age, gender and profession. Shortly after we will proceed by presenting you with specific video footage. It is necessary to pay close attention to the behaviour of the people in the videos, so that you are capable of answering questions after each video.

During the process of watching these videos you will be video-taped without audio for further analysis. However, these videos will solely be analysed by software and not by the researchers. The video recordings will be stored anonymously on a separate platform in Europe which adheres to the GDPR rules.

Once you are finished with the video-material and the corresponding questions, you will be asked to rate yourself with a short Inventory of 11-items in total. Afterwards, you will be asked whether you want to receive more information on the study you just took part in (optional) and we thank you for your participation.

Use of the information in the study: The information you provide will be used for the data collection and analysis procedure of a bachelor's thesis. Therefore, the results will be published and stored in the University of Twente library.
Any personal information you provide such as your age, gender or profession will not be passed to third parties beyond the research team.
For specific parts of our research report we may decide to quote some of your answers anonymously.

Trigger warning!: The video-material we are using contains intense displays of emotion which may be of triggering nature. The people in the videos may express great sadness,

hopelessness and anger. So, if you do not feel like taking part in the study based on this information and for any other reason we fully respect your decision and are thankful for your consideration.

Contact Information for Questions about Your Rights as a Research Participant: If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by **ethicscommittee-hss@utwente.nl**

Appendix B

Video Material

Video footage and its sources

Name of Missing Person	Name of Pleader	Relationship to Missing Person	Country of Origin	Innocence /Guilt	Source
Sharon Malone	Harry Clinch	father	UK	Innocent	BBC
Louise Brown	Paul Brown	father	UK	guilty	ITN Source
Linda Flemming	Derek Flemming	father	UK	guilty	ITN Source
Sarah Payne	Sara Payne	mother	UK	innocent	BBC
Karrissa Boudreau	Penny Boudreau	mother	Canada	guilty	Voxant - Global News
Shannon Matthews	Karen Matthews	mother	UK	guilty	ITN Source
Holly Wess & Jessica Chapman	Leslie Chapman & Sharon Chapman, Kevin Wells	father & mother, father	UK	innocent	BBC

Video questions

Question item 1:

To what extent are the emotions as displayed in the video authentic? (0 = not authentic at all;

100 = very authentic)

Note. *measured by a scale ranging from 0-100

Question item 2:

To what extent did you feel empathy for the people/person in the video?

Not at all	Very little	Little	Neutral	More than average	Moderately high	High

Question item 3:

Do you think the people in the video were trying to deceive the viewer?

Yes	No

Appendix C

BFI-10 Inventory and Scoring information

Instruction: How well do the following statements describe your personality?

I see myself as someone who …	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
… is reserved	(1)	(2)	(3)	(4)	(5)
… is generally trusting	(1)	(2)	(3)	(4)	(5)
… tends to be lazy	(1)	(2)	(3)	(4)	(5)
… is relaxed, handles stress well	(1)	(2)	(3)	(4)	(5)
… has few artistic interests	(1)	(2)	(3)	(4)	(5)
… is outgoing, sociable	(1)	(2)	(3)	(4)	(5)
… tends to find fault with others	(1)	(2)	(3)	(4)	(5)
… does a thorough job	(1)	(2)	(3)	(4)	(5)
… gets nervous easily	(1)	(2)	(3)	(4)	(5)

I see myself as someone who …	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
… has an active imagination	(1)	(2)	(3)	(4)	(5)

Optional additional Agreeableness item (true-scored)

…is considerate and kind to almost everyone.	(1)	(2)	(3)	(4)	(5)

Scoring information for BFI-10

Scoring the BFI-10 scales:

Extraversion: 1R, 6; Agreeableness: 2, 7R; Conscientiousness: 3R, 8; Neuroticism: 4R, 9; Openness: 5R; 10 (R = item is reversed-scored).

Appendix D

Table 5

Psychometric qualities of BFI-10 inventory according to Rammstedt & John (2007)

	Part-whole correlations				Test-retest stability		
	US-1	US-2	G-1	Mean	US-1	G-1	Mean
Extraversion	.87	.90	.90	.89	.79	.87	.83
Agreeableness	.74	.78	.70	.74	.69	.66	.68
Conscientiousness	.84	.77	.83	.82	.70	.83	.77
Neuroticism	.88	.85	.86	.86	.76	.71	.74
Openness	.79	.78	.80	.79	.65	.78	.72
Mean BFI-10	.83	.82	.83	.83	.72	.78	.75
Mean BFI-44	—	—	—	—	.83	.85	.84

Note. *The abbreviations "US-1", "US-2" and "G-1" refer to different samples that were measured in the USA and Germany respectively. The mean correlations were computed by Rammstedt & John using the Fisher *r-to-Z* transformation. The test-retest stability has been measured after a 6 week interval for the German sample and an 8 week interval for the US sample. All values are derived from the following paper by Rammstedt and John (2007);

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality*, *41*(1), 203-212. https://doi-org.ezproxy2.utwente.nl/10.1016/j.jrp.2006.02.001

Appendix E

Spss Syntax

```
* Encoding: UTF-8.
*Correcting Variable Codes.
DATASET ACTIVATE DataSet1.
RECODE Task1_1 Task1_7 Task1_3 Task1_4 Task1_5 (1=5) (2=4) (3=3) (4=2) (5=1).
EXECUTE.


*Correcting Qualtrics Error.
RECODE Q12 Q15 (7=1) (8=2).
RECODE Q18 (9=1) (10=2).
RECODE Q21 (11=1) (12=2).
EXECUTE.


*Correcting Counterbalancing.
RECODE Q3 Q12 Q15 Q21 (1=0) (2=1).
RECODE Q6 Q9 Q18 (1=1) (2=0).
EXECUTE.
```

*Checking for differences between videos.

GLM Q1_27 Q4_27 Q7_27 Q10_27 Q13_27 Q16_27 Q19_27

 /WSFACTOR=Faktor1 7 Polynomial

 /METHOD=SSTYPE(3)

 /CRITERIA=ALPHA(.05)

 /WSDESIGN=Faktor1.

GLM Q2_1 Q5_1 Q8_1 Q11_1 Q14_1 Q17_1 Q20_1

 /WSFACTOR=Faktor1 7 Polynomial

 /METHOD=SSTYPE(3)

 /PRINT=DESCRIPTIVE

 /CRITERIA=ALPHA(.05)

 /WSDESIGN=Faktor1.

GLM Q3 Q6 Q9 Q12 Q15 Q18 Q21

 /WSFACTOR=Faktor1 7 Polynomial

 /METHOD=SSTYPE(3)

 /PRINT=DESCRIPTIVE

 /CRITERIA=ALPHA(.05)

 /WSDESIGN=Faktor1.

*Scoring for correct identification of deception.

** To do: Identify which videos are actually deceptive (1=D) (0=T).

COMPUTE CorrectIdentification=0.

IF (Q3 = 1) CorrectIdentification=CorrectIdentification+1.

IF (Q6 = 1) CorrectIdentification=CorrectIdentification+1.

IF (Q9 = 0) CorrectIdentification=CorrectIdentification+1.

IF (Q12 = 0) CorrectIdentification=CorrectIdentification+1.

IF (Q15 = 1) CorrectIdentification=CorrectIdentification+1.

IF (Q18 = 1) CorrectIdentification=CorrectIdentification+1.

IF (Q21 = 0) CorrectIdentification=CorrectIdentification+1.

EXECUTE.

*Compute mean scores for BFI-10 scales.

COMPUTE Extraversion=MEAN (Task1_1,Task1_6).

COMPUTE Agreeableness=MEAN (Task1_2,Task1_7,Task1_11).

COMPUTE Conscientiousness=MEAN (Task1_3,Task1_8).

COMPUTE Neuroticism=MEAN (Task1_4,Task1_9).

COMPUTE Openness=MEAN (Task1_5,Task1_10).

EXECUTE.

EXAMINE VARIABLES=CorrectIdentification Extraversion Agreeableness Conscientiousness Neuroticism

  Openness

 /PLOT BOXPLOT HISTOGRAM NPPLOT

 /COMPARE GROUPS

 /STATISTICS DESCRIPTIVES EXTREME

 /CINTERVAL 95

 /MISSING LISTWISE

 /NOTOTAL.

*Correlational analysis including Bivariate Correlation and Spearman correlation coefficient.

*CORRELATIONS

 /VARIABLES=Extraversion Agreeableness Conscientiousness Neuroticism Openness CorrectIdentification

 /PRINT=TWOTAIL NOSIG FULL

 /MISSING=PAIRWISE.

NONPAR CORR

 /VARIABLES=Extraversion Agreeableness Conscientiousness Neuroticism Openness CorrectIdentification

 /PRINT=SPEARMAN TWOTAIL NOSIG FULL

 /MISSING=PAIRWISE.

*Regression analysis to look for variation across findings and normal distribution (Shapiro-Wilk).

REGRESSION

 /MISSING LISTWISE

```
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT CorrectIdentification

/METHOD=ENTER Extraversion Agreeableness Conscientiousness Neuroticism Openness

/RESIDUALS DURBIN.


FREQUENCIES VARIABLES=CorrectIdentification

/ORDER=ANALYSIS.
```