Does It Show When Someone Has Hostile Intentions? When In Doubt, Replicate!

Bachelor Thesis

Lisa Ebben, s1814699

First supervisor: Dr.Ir. P.W. de Vries

Second supervisor: Dr. M.W.M. Kuttschreuter

University of Twente

22 July 2019

Abstract

Background. Contradicting opinions exist about whether humans are able to predict people's deceptive intentions just by observing them. It is suggested that increasing liars' cognitive load makes it easier for people to detect their deception. This research tries to find the same results of a previous study by Wijn and colleagues (2017), who discovered that exposing deceptive people to an environmental cue and high cognitive load improved judges' ability to detect their hostile intent. A new addition to the replication is the investigation of the possible role of gender and personality traits in judges' accuracy to detect deception.

Method. Judges (n = 23) were shown videos from the original study for which they had to rate if the participants in the videos had hostile intentions or not. Participants in the videos were subjected to a 2 (hostile vs. non-hostile intentions) x 2 (high vs. low cognitive load) x 2 (exposure to an environmental cue vs. no cue) design.

Results. The current study did not obtain the same results as the original study: the environmental cue and high cognitive load did not have a significant effect on participants' accuracy to detect hostile intent and innocence. It appeared that the participants in this study were no better than chance at detecting hostile intent. However, it was found that males were better at detecting whether females, compared to males, were actually innocent. Besides, they rated innocent males as guilty more often than innocent females. Females did not show a difference in accuracy. Also, it was easier for judges to detect hostile intent in males, and more difficult to detect hostile intent in females. Lastly, judges' scores on Neuroticism, Openness and Agreeableness did not influence their ability to detect hostile intent or innocence.

Discussion. In order for the results of the original study to become useful in real-life situations, other research needs to be conducted that can confirm that cognitive load and an environmental cue can increase judges' accuracy to detect deception. The findings related to gender can be valuable in training judges to detect deception more accurately.

Keywords: deception, detection, cognitive load, environmental cue, personality traits, gender

Does It Show When Someone Has Hostile Intentions? When In Doubt, Replicate!

Since 2002, the whole world, except North America, experienced an increase in terrorism (Institute for economics & peace, 2017). This increase led to a heightened necessity to detect deceptive intentions and prevent aggressive behaviour. In order to ensure protection, systems were developed which are able to screen behaviour, facial expressions and physiological elements, such as respiration. One of these systems is called Screening Passengers by Observation Technique (SPOT; Homeland Security, 2011), created to identify possibly threatening individuals on airports. Individuals are selected by so-called Behaviour Detection Officers (BDO's). The SPOT is based on Ekman's Facial Action Coding System (FACS; Ekman & Friesen, 1978), which he designed to analyse facial expressions and later used as a method to teach people how these expressions were connected to certain unseen emotions, including hostile ones. Especially 'microfacial' expressions are of importance, such as raising a brow, as they can indicate a person's effort to mask one's true feelings. Ekman declares that a 70% accuracy can be reached if trained observers detect deception using facial expressions alone and a 100% accuracy if gestures and body movement are also considered (Ekman & O'Sullivan, 1991).

Both the FACS and SPOT have received criticism. Honts and Raskin, as reported by Weinberger (2010), criticised Ekman on his coding system after not being able to replicate his findings (Weinberger, 2010). Regarding SPOT, in its first phase, BDO's identified more than 232,000 people to go through a second screening involving an inspection of bags. Of those people, 1,710 were actually arrested (Weinberger, 2010). Proponents of the programme point out that this number of arrests proves the programme's effectiveness. Critics, however, remark that fewer than 1% of referred people were arrested for criminal activities which did not involve terrorism, proving the programme's ineffectiveness and BDO's inability to detect intentions to deceive. The important question underlying the contradicting findings and criticism: Is it possible to predict people's deceptive intentions just by observing them?

Theoretical framework

To deceive is defined as "to intentionally cause another person to have or continue to have a false belief that is known or truly believed to be false by bringing about evidence on the basis of which the person has or continues to have the false belief" (Mahon, 2007, p. 189). This definition implies several things, namely that deceiving is intentional, that the deceived person obtains or resumes to have a false belief, and that the deceiver needs to know that the belief the deceived person has obtained or resumes to have is false. For instance, people can deceive by using signs that involve resemblance (icons), such as posting a happy face

emoticon about an item on social media that one is actually unhappy about (Mahon, 2015). Also, it is possible to deceive by using non-linguistic signs (symbols), such as wearing a wedding ring when one is not married. Besides, a person can deceive by speaking and thereby making statements, or by remaining quiet and thereby purposefully not making statements to withhold the truth.

Deception can thus come in a variety of types, ranging from direct fabrication to more subtle forms such as half-truths and concealments (Carlson, George, Burgoon, Adkins, & White, 2004). This entails that deception and lying have a different meaning, although the two terms are used interchangeably. A lie is defined as "an intentionally false statement" ("Lie", n.d.). Hence, deception refers to a process, whereas lying regards an explicit approach that can be used during this process. However, since liars have the same intention as people who try to deceive, namely to mislead, and lying and deception often have the same consequences, namely that people will receive false information (Bivins, 2008), lying and deception can be seen as concepts similar to each other. Results found by previous studies that focused on lying will thus be considered applicable to deception as well.

The ability of humans to accurately detect deception is the same as or somewhat better than flipping a coin, with a success rate of 54% (Aamodt & Custer, 2006; Bond & DePaulo, 2006; DePaulo et al., 2003; Hartwig & Bond, 2011). Generally, people categorised 47% of lies and 61% of truths correctly (Bond & DePaulo, 2006). Traditionally, police practices in deception detection were based on theories that assumed liars would show stress because of their fear of getting caught and their guilt about lying (Zimmerman, 2016). These theories made people pay attention to behavioural cues, without success. After finding out the ineffectiveness of these practices, psychologists have now developed other strategies to improve humans' ability to detect, rather than those of looking at visual cues that might signal deception, such as avoiding eye contact or fidgeting (Zimmerman, 2016). One of these strategies is increasing people's cognitive load.

Cognitive load. Deception can be considered an ability, which requires work and concentration to achieve its purpose (Spence, 2004). When one is fatigued, their deceiving ability can be inconsistent and one will be more prone to accidentally confess their lies. It is therefore likely that deception involves cognitive processes, such as the monitoring of reactions of both the liar and the victim when they interact. Besides, a liar must be able to do at least two actions at the same time: create new information (the lie) and withhold information (the truth). It can thus be suggested that deception requires extra cognitive effort than telling the truth does (Spence, 2004).

The act of deceiving is connected to activity in certain brain areas. In previous studies, the frontal executive region of the brain has been constantly found to be connected to deception (for reviews, see Abe, 2009; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Gombos, 2006; Hughes et al., 2005; Spence, 2004; Zeki et al., 2004; Spence & Kaylor-Hughes, 2008). Cognitive areas in the brain, such as the dorsolateral prefrontal cortex, the ventrolateral prefrontal cortex, the anterior prefrontal cortex, and the anterior cingulate cortex, have been suggested to be active during deception. The dorsolateral prefrontal cortex, for example, is involved in monitoring and manipulating working memory (Owen, 1997), response selection (Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000), and cognitive control (MacDonald, Cohen, Stenger, & Carter, 2000). The ventrolateral prefrontal cortex is associated with task switching (Dove, Pollmann, Schubert, Wiggins, & Von Cramon, 2000), reversal learning (Cools, Clark, Owen, & Robbins, 2002), and response inhibition (Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007). The anterior prefrontal cortex is connected to the integration of multiple operations to obtain a higher behavioural goal (Ramnani & Owen, 2004). The anterior cingulate cortex has been found to play a role in processes such as conflict detection (Kerns et al., 2004) and emotional processing (Phan, Wager, Taylor & Liberzon, 2002). Results from lie detection studies, developmental psychology, and neurocognitive imaging persistently indicate executive functions, such as inhibition, attention, metacognition, and management of working memory, to be the central mechanisms in deception (Gombos, 2006). These executive functions are identified as core features by theories that try to explain the process of deception, which predict that lying will become increasingly difficult as access to these functions is restricted. Given these findings that deception is connected to various cognitive processes, the brain regions discussed above presumably cooperate to form a complex cognitive performance such as deception.

As it appears that deceiving is a cognitively demanding process, increasing liars' cognitive load by means of executing another cognitively demanding task at the same time could make it more difficult for them to lie without getting caught. One underlying mechanism causing this could be that the cognitive demands of lying and another mentally challenging task are gathered from a shared, restricted pool of cognitive resources (Vrij et al., 2008). Lying performance will become worse when the necessary cognitive resources for the two demanding tasks surpass the limit of attention, called divided-attention (Johnston, Greenberg, Fisher, & Martin, 1970). Another possible process is that one must quickly switch their attention between the two demanding tasks when performing them simultaneously (Broadbent, 1957). The more cognitively demanding a task is, the more attention it requires,

leaving less attention to be dedicated to the second task (lying). Consequently, the less attention that is dedicated to a task, the worse performance will be (Kahneman, 1973).

Since Zuckerman, DePaulo, and Rosenthal's study (1981), who proved that lying is more cognitively demanding than telling the truth, research demonstrated that increasing the cognitive load of liars indeed made it easier for people to distinguish them from truth tellers (Vrij et al., 2008; Vrij, Mann, Leal, & Fisher, 2010). There are several reasons why lying demands more cognitive load than truth telling. First, making up a lie is cognitively demanding in itself (Vrij, Granhag, & Porter, 2010). When lying, somebody needs to make up a story, which comes across as believable to observers. Then, liars have to remember what they made up in order to keep consistent and believable. Besides, liars need to prevent slips of the tongue and avoid giving new leads (Vrij, 2008). Second, compared to truth tellers, liars are more prone to monitor and control their behaviour in order to come across as honest (DePaulo & Kirkendol, 1989), which is something that requires cognitive load (Baumeister, 1998). A guilty suspect, for example, may experience strong emotions, such as fear, anger or excitement, that must be repressed or faked (Porter & ten Brinke, 2010). Third, during lying, people are aware that observers pay attention to their behaviour (Buller & Burgoon, 1996). Therefore, liars might monitor their interviewer's reactions to see if their lies seem like the truth, which is an action that is also cognitively demanding (Buller & Burgoon, 1996; Schweitzer, Brodt, & Croson, 2002). Fourth, liars might be engaged with acting and role playing (DePaulo et al., 2003). Fifth, liars have to withhold the truth, which is something that requires cognitive effort (Spence et al., 2001). Last, compared to activating a truth, which is an automatic process, activating a lie is a process that requires more thought (Gilbert, 1991; Walczyk, Roper, Seemann, & Humphrey, 2003; Walczyk et al., 2005).

One can distinguish liars from truth tellers by examining their levels of cognitive load. Liars' level of cognitive load is higher than that of truth tellers, as lying takes up more cognitive resources than telling the truth. This creates vulnerability, and if even more cognitive load is imposed, by for example creating additional demands, liars might not cope as well as truth tellers with these demands (Vrij, Granhag, & Porter, 2010). As mentioned above, these implications also apply to deception, because lying is a form of deception.

This information implies that putting cognitive burden on liars will increase the chance his or her lying behaviour is more visible. As research has revealed that liars are often not visibly nervous and that cues to deception are too weak to notice (Vrij, Granhag, & Porter, 2010), increasing cognitive load can be a better way of detecting deception.

Gender and personality traits. Gender and personality traits have been suggested to play a role in deception detection by some studies, yet in other studies no significant results were found about their possible influence. For example, several studies have shown that males and females have different ways of determining whether someone is honest or not, paying attention to different factors (Levitan et al., 2015; Lewicki & Robinson, 1998; Tilley, George, & Marett, 2005). During communicative situations, males tend to focus on status, independence, and competition cues while females tend to focus on intimacy and social connection cues (Tannen, 1990). All three studies concluded that females are better at detecting deception than males. A possible reason for this is that females are better than males at processing emotional expressions (Collignon et al., 2010). However, another study found that males are slightly better than females at detecting deception (Heijboer et al., 2015). Yet other outcomes are described by a meta-analysis, namely that sex does not affect the ability to detect deception (Aamodt & Custer, 2006). They speculate that the reasons for this finding could be that people are in general bad at detecting deception regardless of their sex, or that the experimental situations used in the studies that were investigated, were inappropriate to test for deception detection. As can be seen, findings are contradicting.

DePaulo, Stone and Lassiter (1985) discovered that the accuracy of deception detection improved when people judged the opposite sex. They suggest that in opposite-sex situations, people become more motivated to lie successfully. As people become more motivated, it will become easier to distinguish and detect their lies from their truths in channels including nonverbal cues, than in the channel that includes only verbal cues. Being more motivated to lie successfully, people will try harder to monitor and control all of their behaviours. However, nonverbal behaviour might be less easily controlled than verbal behaviour, which can lead to a decrease of control in the nonverbal channels. Another study found similar results, which showed that in opposite-sex situations the detection accuracy was significantly higher (Porter, Campbell, Stapleton, & Birt, 2002). Forrest and Feldman (2000) also found a gender bias in detection accuracy. In their study, female judges considered female targets as more honest than male targets, and male judges estimated male targets as more honest than female ones. The judges were more prone to approve the statements made by same-sex targets than those of opposite-sex targets.

Levitan and colleagues (2015) investigated if people's personality traits matter in their accuracy of judgment. They found that for females but not for males, the trait Neuroticism appeared to be negatively correlated with the ability of deception detection. The authors suggest that people with this trait possess an inflated need to presume people are generally

truthtellers and trustworthy, as a neurotic person suffers more than others when confronted with worrisome thoughts or negative perceptions. In addition, another study found that the traits Openness and Agreeableness enhance the capacity of recognising deceiving intentions, next to again a negative correlation between Neuroticism and the ability to detect deception (Enos et al., 2006). Enos and colleagues give a possible explanation for these results. An extremely high score in Agreeableness is connected to dependent personality disorder, of which intense attention to the opinions of others and sensitivity to affect are some of the symptoms. They argue that this sensitivity improves the ability to perceive deception. Contrary to these positive results, other studies found that personality traits do not play a role in the ability to detect deception (Ask & Granhag, 2003; Spencer, 2017).

For the reason that results from previous studies differ and mostly concern the overt style of lying (verbal deception) instead of nonverbal deception, the current study wants to examine if gender and personality traits can influence people's detection accuracy and whether the results found in previous studies also apply to deception detection.

Present study

This study partially aims to replicate the experiment that was executed in a previous research by Wijn and colleagues (2017), who examined whether environmental cues and imposing cognitive load can improve deception (i.e. hostile intentions) detection accuracy. A new addition to the replication is the examining of the possible role of gender and personality traits in judges' accuracy to detect deception. The original research showed successful results, however much criticism remains on whether it is possible for people to detect hostile intentions (Meijer, Verschuere, & Merckelbach, 2018). Therefore, it is important to know whether these results can be reproduced and if the validity and reliability of the results will increase. As this research intends to find the same outcomes as the original study found, its purpose is to also investigate "the role of cognitive load in combination with environmental cues on the sensitivity of judges (i.e., the ability to successfully distinguish participants with a hostile intent from participants with non-hostile intent)" (Wijn et al., 2017). Additionally, it is questioned whether gender and personality traits improve judges' accuracy to detect hostile intent. Expecting to find the same results as the original study, the same hypotheses will be tested:

> Hypothesis 1: Environmental cues signalling risk of exposure lead to perceptible reactions of individuals with hostile (vs. non-hostile) intentions and therefore to a better ability to distinguish individuals with hostile intentions from individuals with non-hostile intentions.

Hypothesis 2: Chances of detection of individuals with hostile intentions will only be higher when these individual's cognitive capacity is constrained, so that and inhibitory control of impulsive or reflexive behaviours, and activation control of desirable behaviours is hampered.

Based on the positive results found by earlier studies, the following hypotheses were formulated to examine whether this study is able to find the same results about gender and personality traits in the context of deception:

Hypothesis 3: Females are better at detecting deception than males.

Hypothesis 4: Accuracy of deception detection is higher when opposite-sex judgments are made than when same-sex judgments are made.

Hypothesis 5: Deception detection accuracy is higher in people who score high on Agreeableness and Openness and low on Neuroticism.

Method

Design

This study follows the design of the previous research by Wijn and colleagues (2017). They implemented a 2 (hostile vs. non-hostile intentions) x 2 (high vs. low cognitive load) x 2 (exposure to an environmental cue vs. no cue) repeated measures design, which their participants were subjected to. In their experiment, judges had to watch videos of these participants who walked a pre-defined route. For this study, the factor gender is added, as well as the personality traits as covariates. The independent variables are cognitive load, environmental cue, gender of participants (in the videos) and gender of judges. The dependent variable is accuracy of (non-)hostile intentions detection. Judges were unaware of these conditions.

Participants

After performing a G-power analysis, it was concluded that 60 participants were required in this study. Due to time issues, however, this number was not achieved. Instead, 23 participants took part, who were all students (15 females, 8 males, M_{age} = 22.35, SD = 2.31, age range: 19-29 years), gathered through convenience sampling. 60,9% was German, 30,4% was Dutch, and 8,7% had another nationality. 16 of the participants were psychology students, one studied communication sciences, and six were doing another study. After participating, participants would receive cookies and credits they needed in order to graduate their studies. Participants had to be able to understand and speak English.

Materials

Videos from the original research by Wijn and colleagues (2017) were used (see Image 1 and 2), by employing the programme Matlab. In these videos, 143 participants (64 males and 79 females), who were randomly assigned to the conditions (hostile intentions x environmental cue x cognitive load), are walking with a laptop bag through a 20 meter long hallway towards a checkpoint. The device producing the environmental cue (i.e. police walkie-talkie's static noise) was hidden halfway. The laptop bags contained either a dummy explosive device or books, of which the participants were aware. The participants carrying a dummy explosive device were told not to be stopped at the checkpoint and get their bag searched, as they would have to fill in time-consuming questionnaires, when otherwise having a chance of winning 100 euros. Judges had to watch these videos and assess whether the person in the video is having hostile intentions (i.e. carrying the dummy explosive device) or not (i.e. carrying the books). Also, the participants carrying the bags had to fill in a questionnaire after the experiment, in which they were asked to report their feelings during the task. However, this part of the original study will not be further examined.

To test the hypotheses on personality traits, the NEO Personality Inventory – 3 (NEO-PI-3; McCrae, Costa Jr, & Martin, 2005) was administered to the judges. The NEO-PI-3 is the revised version of the NEO Personality Inventory – Revised (Costa & McCrae, 1992). The test contains 240 statements, in order to measure the Big Five personality dimensions extraversion (outgoing vs. reserved), neuroticism (emotional stability vs. emotional instability), openness to experience (curious vs. conventional), agreeableness (altruistic vs. egocentric) and conscientiousness (ambitious/goal-oriented vs. easy-going/careless) (De Fruyt, Hoekstra, & Ormel, 1996). Statements are responded to by means of a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Judges were asked to respond to the statements of the dimensions Neuroticism, Openness and Agreeableness, the only dimensions of interest in this study, based on what they thought was similar to their most typical behaviour. The Cronbach's alphas of the dimensions Neuroticism, Openness and Agreeableness are .87, .89 and .83, respectively, indicating high internal consistency.

In addition to the NEO-PI-R statements, some demographic questions were asked at the beginning of the questionnaire, including age, gender, nationality and study. The end of the questionnaire contained a debriefing, in which judges were informed about the conditions the participants in the videos were subjected to. The complete questionnaire can be found in Appendix A, except for the NEO-PI-R statements, which are excluded because spreading them is forbidden.

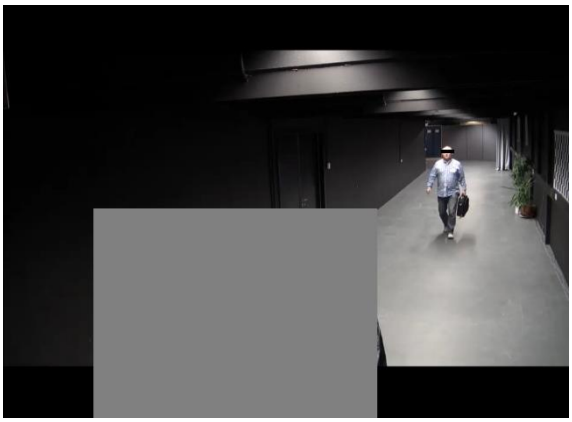

Image 1. Person walking down the hallway (example video).

Image 2. Decision screen at the end of each video.

Procedure

In the first place, the judges were welcomed and seated behind a laptop. The current study was briefly explained, after which judges received the instruction of watching the videos and having to fill in some demographic questions, as well as questions about which indicators they focused on when deciding whether a participant had hostile intent or not (used in another study) and the NEO-PI-R afterwards. Before participating in the experiment, judges were asked to sign the informed consent in Qualtrics. After, Matlab was started, which contained the videos of the original study (Wijn et al., 2017). In Matlab, first the instructions for watching the videos were explained. Judges were notified that approximately half of the participants in the videos were carrying a dummy explosive device, and the other half was carrying books. In reality, 73 participants were guilty and 70 participants were innocent. Then they would start with five practice videos, having to choose after each one if the person was carrying a bomb in their bag (yes/no). After the practice videos, the instructions were again briefly shown, and the experimental videos were started. Having watched all 143 videos, Matlab was closed automatically and judges had to call the researcher to start Qualtrics. In Qualtrics, they were asked to fill in the questionnaire (Appendix A).

Analysis

Due to technological issues, the data could not be analysed by calculating a specific accuracy measure, which was done in the original study. Instead, detection accuracy was measured by analysing the Hits (true positives), Misses (false negatives), False Alarms (false positives) and Correct Rejections (true negatives) the judges made (Figure 1). First of all, the data were transmitted from Matlab to SPSS. Descriptive statistics of age, gender, nationality and study were explored. Judges' answers to the videos (yes/no) were recoded into new

variables called Hit/Miss/False Alarm/Correct Rejection, distinguished by the condition of the video and the gender of the person in the video (examples are cue x load x male and no cue x no load x female). From these variables, the means were computed per category (e.g. the variable MissesCueLoadMale contains the mean of Misses a judge made in all videos of male participants in the condition cue x load) and with these new variables, a repeated measures ANOVA was carried out.

Figure 1

Description of a Hit, Miss, False Alarm and Correct Rejection

		Hostile Intent	
		Yes	No
Rated	Yes	Hit	False Alarm
	No	Miss	Correct Rejection

Results

Judges all claimed they took their participation in this study seriously, as was indicated by their score (M = 4.61, SD = 0.50) on the five-point Likert scale (1 = strongly disagree, 5 = strongly agree). On average, judges were able to concentrate (M = 3.57, SD = 0.95), and were not distracted (M = 4, SD = 0.85).

First, judges' Hits, Misses, False Alarms, and Correct Rejections were more closely looked at (Table 1). Overall, judges rated fewer than half of the people guilty (M = 55.22, SD = 18.18), despite being told that half of the participants in the videos was guilty and the other half innocent.

Table 1

Judges' Rates of Misses, Hits, False Alarms and Correct Rejections

	Minimum	*Maximum*	*M*	*SD*
Misses	.42	.93	.63	.14
Hits	.07	.58	.37	.14
False Alarms	.06	.57	.40	.13
Correct Rejections	.43	.94	.60	.13

What can be concluded from Table 1, is that judges scored a higher number of Misses (M = .63) and Correct Rejections (M = .60) than Hits (M = .37) and False Alarms (M = .40). They scored Hits the least and Misses the most.

Next, a repeated measures ANOVA was conducted with cue, cognitive load and gender of participants in the videos as independent variables, Hits & Misses and Correct Rejections & False Alarms as dependent variables, with gender of judges as between-participants factor, and their scores on the personality traits Neuroticism, Openness and Agreeableness as covariates. The results are displayed in Table 2.

Table 2

Hits & Misses and False Alarms & Correct Rejections as a function of cue, load, gender and personality traits

	Hits & Misses			False Alarms & Correct Rejections		
Factor	*df*	*F*	*p*	*df*	*F*	*p*
Cue	(1,18)	0.72	.41	(1,18)	3.72	.07
Load	(1,18)	0.90	.36	(1,18)	1.65	.22
GenderVideos	(1,18)	8.09	.01	(1,18)	2.53	.13
Cue x Load	(1,18)	1.02	.38	(1,18)	0.18	.68
Cue x GenderJudges	(1,18)	0.71	.41	(1,18)	1.09	.31
Load x GenderJudges	(1,18)	1.20	.29	(1,18)	0.04	.84
Cue x Load x GenderJudges	(1,18)	0.29	.60	(1,18)	0	1
GenderVideos x GenderJudges	(1,18)	1.38	.26	(1,18)	6.69	.02
GenderVideos x ScoreOpenness	(1,18)	5.49	.03	(1,18)	0.04	.84
Load x ScoreNeuroticism	(1,18)	2.59	.13	(1,18)	5.65	.03
Cue x ScoreAgreeableness	(1,18)	0.13	.72	(1,18)	4.62	.05
Load x GenderVideos x ScoreNeuroticism	(1,18)	0	.96	(1,18)	4.56	.05

First of all, to clarify: Hits and Misses are each other's inverses, which also applies to False Alarms and Correct Rejections. For example, when a judge got a Hit for a participant, he or she automatically did not get a Miss. Being each other's inverses, it means that the Hits and Misses (and False Alarms & Correct Rejections) got the same outcomes after executing the repeated measures ANOVA.

As can be seen, most factors did not significantly interact with each other. Both cue and cognitive load had no significant effect on neither the Hits & Misses nor the False Alarms & Correct Rejections judges made. This means that hypothesis 1 and 2 are rejected. However, gender from the participants in the videos seemed to have an effect on the Hits & Misses of judges, F (1,18) = 8.09; $p < .05$. It appeared that it was slightly easier for judges to score Hits for males (M = .38, SD = 0.18) than for females (M = .35, SD = 0.19). Also, it was slightly easier to score Misses for females (M = .65, SD = 0.19) than for males (M = .62, SD = 0.18).

For the False Alarms & Correct Rejections, an interaction was found between gender of participants in the videos and gender of judges, F (1, 18) = 6.69; $p < .05$. Pairwise comparisons showed that there was a small difference in accuracy of Correct Rejections of males when they judged females compared to males, and of False Alarms when males judged males compared to females ($M_{difference}$ = .12, SE = .04; F (1, 21) = 7.36, $p < .05$). Females showed no difference in accuracy of neither False Alarms nor Correct Rejections based on the gender of the participants in the videos ($M_{difference}$ = .01, SE = .03; F (1, 21) = 0.15, $p > .05$). While hypothesis 3 is rejected, as females did not have a higher accuracy in detecting hostile intent than males did, hypothesis 4 can be partially accepted. This because for Correct Rejections, opposite-sex situations seemed to make a difference in accuracy score for males. However, it appeared that same-sex situations decreased males' accuracy to detect innocence.

Lastly, no significant effects were found from the personality scores on Hits, Misses, Correct Rejections and False Alarms. Interactions that were found are, for the Hits, one between gender of people in the videos and score on Openness, F (1, 18) = 5.49; $p < .05$. For the False Alarms, load and score on Neuroticism seemed to interact, F (1,18) = 5.65; $p < .05$, as well as cue and score on Agreeableness, F (1, 18) = 4.62; $p < .05$. Also, an interaction was found between load, gender of the people in the videos, and score on Neuroticism, F (1, 18) = 4.56; $p < .05$. Further analyses on these interactions fall outside the scope of this research. Hypothesis 5 is rejected, because personality traits did not affect judges' accuracy score.

Discussion

Replication

This research has failed to replicate the findings of the research by Wijn and colleagues (2017). They found that judges' accuracy of detecting hostile intent in people improved when those people had high cognitive load and were exposed to an environmental cue. Results of the current study show that an environmental cue and high cognitive load did

not have a significant effect on judges' accuracy to detect hostile intent and innocence. Hypothesis 1 and 2 are therefore rejected.

Besides, it appeared that the judges in this study were unable to detect hostile intent in many participants and on average, they even scored below the chance level which is claimed to be the accuracy level of humans for detecting deception (Aamodt & Custer, 2006; Bond & DePaulo, 2006; DePaulo et al., 2003; Hartwig & Bond, 2011). They did score above chance level on Correct Rejections, however also on Misses, indicating that these scores are probably caused because judges scored a lot more people innocent than guilty when this ratio should have been close to 50/50, which can be considered a form of response bias.

The fact that this study was not able to find the same positive results as the original study does not indicate that those results were found by chance. It is proven almost impossible to exactly replicate a study (Raloff, 2015). Besides, there were fewer judges in this study, and they were all students. The original study did not have only students as their judges. It could be that students in particular are bad at detecting hostile intent or just the ones that participated in this study, which would explain the results. The second assumption is more likely, as studies have found that students are not less accurate at detecting deception than other citizens (Aamodt & Custer, 2006). In addition, the small sample size could have decreased the power of this study so that possible smaller effect sizes were not detected. Also, this study did not follow the exact same analyses of the original study. A different measure was used to check for accuracy of judges' deception detection, namely the Hits, Misses, False Alarms and Correct Rejections. However, with the new accuracy measure, the possible influence of cognitive load and the environmental cue could still be examined, so eventually the results of the original study and this one could be compared properly. In order for the results of the original study to become useful in real-life situations, other research needs to be conducted that can confirm that cognitive load and an environmental cue can increase judges' accuracy to detect deception. When confirmed, increasing cognitive load and exposing deceivers to an environmental cue can become valuable techniques to increase the accuracy of judges to detect deceivers.

Gender and personality traits

Further analyses found some differences in gender concerning deception detection accuracy. Firstly, males and females both lacked the accuracy to judge whether a participant was guilty. This means that hypothesis 3 is rejected, as females did not differ in their ability from males to judge hostile intent. This outcome corresponds with the meta-analysis which also found that sex did not influence the ability to detect hostile intent (Aamodt & Custer,

2006). Hypothesis 4 is partly supported as males were better at judging innocence in opposite-sex situations. However, male judges rated innocent males as guilty more often than innocent females. Females did not show an accuracy improvement in opposite-sex situations. Another study which studied gender differences in deception detection found that females' honesty, compared to males' honesty, is easier to detect (Kuzio, 2018), which confirms the finding that males were good at detecting females' innocence and thought innocent males (compared to innocent females) were guilty more often. Additional findings were that it was easier to detect hostile intent in males, and more difficult to detect hostile intent in females. This implies that not only the gender of the judges, but also the gender of the deceivers influenced judges' accuracy of deception detection.

Thus, the findings about gender tell us that males are better at detecting an innocent female than an innocent male, that males rate innocent males as guilty more often than they rate innocent females guilty, that it is more difficult to detect a guilty female than a male, but more easy to detect a guilty male than a female, and that females and males are inaccurate in detecting hostile intent. Based on these findings, a possible recommendation for real-life practice is that people need to be trained in order to be able to detect deception. Studies have found that trained judges are more accurate than untrained judges (DeTurck, 1991; DeTurck, Harszlak, Bodhorn, & Texter, 1990; DeTurck & Miller, 1990). Furthermore, as it is more difficult to detect deception in a female, training should focus on this difficulty. Also, males need more training in detecting innocence in males.

Considering personality traits, hypothesis 5 is rejected because judges' score on Neuroticism, Openness and Agreeableness did not influence their accuracy to detect hostile intent or innocence. This coincides with results from other studies which also indicated that personality traits did not play a role in the ability to detect deception (Ask & Granhag, 2003; Spencer, 2017). Nonetheless, some interesting interactions were found between personality scores and other factors. This indicates that in combination with other factors (such as gender) personality traits could influence judges' accuracy, but that no direct effect was found. These interactions were unexpected and were not further analysed in this study, however it can be interesting to explore these further in future studies.

Strengths

This research covered the topic deception detection, which has not been studied as much as the topic of lying. Especially the aspect of underlying intent, and not verbal communication, has not been researched often. Therefore, this research is considered as supplementary relevant knowledge in a field which needs more exploring. In addition, it

examines the role of gender and personality traits in the ability to detect, which is also valuable additional knowledge as these possible influences have not been studied much before this research.

Besides, as mentioned before, this research is a replication of another study. The fact that the results of the current research differ from the original research sheds light on the importance of replications. Replications are necessary because, if they come up with the same results, they create confidence that previous results are true. They increase validity and reliability of the previous results of a research. It appears that many psychology studies are unable to be replicated, which questions the scientific process underlaying those studies and their trustworthiness (Diener & Biswas-Diener, 2019).

Future research

In this study, judges' ratings of whether a participant was guilty or innocent were skewed, that is, overall they scored too many participants as innocent. To minimise this response bias, future replications could include two bars inside Matlab which track the responses of judges, so they can see at any time during the experiment how many participants they have rated guilty and innocent.

Due to having to sit for 90 minutes with having a break of only maximum five minutes in the same room, it is possible that after watching the videos, judges were not concentrated anymore on filling out the questionnaires at the end. Also, although judges stated they were concentrated when watching the videos, many of them commented on the amount of videos. Fatigue could have influenced their answers to the questionnaires and/or videos and consequently the results as well. Future replications could decrease the amount of videos judges have to watch, as this is a time-consuming and exhausting exercise, and because 26 videos yield the same results (Wijn et al., 2017).

Another limitation is the fact that the people in the videos are not actual terrorists. They were asked to imagine they were carrying a bomb, which is a complete different situation than a terrorist knowing he is carrying an actual bomb. This could mean that the people were unable to picture themselves as carrying a 'real' bomb and that their behaviour is not displaying they are hiding something, as they do not feel like they are hiding something which could have real consequences. Consequently, judges could and should not use the people's behaviour as an indicator of hostile intent. Future studies should aim to simulate such real terrorist situations as closely as possible, or use videos of terrorists (or other criminals) made at an airport to see whether judges are able to select them from the crowd.

While this research did not succeed in the desired replication of results from the original study (Wijn et al., 2017), it did manage to find some interesting additional outcomes. It appeared that gender and personality traits, when interacting with other factors, can play a role in deception detection. Future research should focus on replicating and extending what was found in this research. As many studies have shown that humans are not accurate at detecting deception, but they still need the ability to do this (e.g. the BDO's in SPOT), it is important to keep investigating whether certain methods can improve their accuracy.

References

Aamodt, M.G. & Custer, H. (2006). Who can best catch a liar?. *Forensic Examiner, 15*(1), 6-11.

Abe, N. (2009). The neurobiology of deception: Evidence from neuroimaging and loss-of-function studies. *Current Opinion Neurology, 22*, 594–600.

Ask, K. & Granhag, P.A. (2003). Individual determinants of deception detection performance: Need for closure, attributional complexity and absorption. *Göteborg Psychological Reports*, 33, No. 1.

Baumeister, R.F. (1998). The self. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., 680–740). Boston: McGraw-Hill.

Bivins, T.H. (2008). The ethics of strategic communication: Winter 2009. Retrieved July 1, 2019, from https://pages.uoregon.edu/tbivins/stratcomweb/notes/lying-notes.pdf

Bond Jr, C.F. & DePaulo, B.M. (2006). Accuracy of deception judgments. *Personality and social psychology Review*, *10*(3), 214-234.

Broadbent, D.E. (1957). A mechanical model for human attention and immediate memory. *Psychological review*, *64*(3), 205-215.

Buller, D.B. & Burgoon, J.K. (1996). Interpersonal deception theory. *Communication Theory*, *6*, 203–242.

Carlson, J.R., George, J.F., Burgoon, J.K., Adkins, M., & White, C.H. (2004). Deception in computer-mediated communication. *Group decision and negotiation*, *13*(1), 5-28.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of cognitive neuroscience*, *19*(1), 69-80.

Christ, S.E., Van Essen, D.C., Watson, J.M., Brubaker, L.E., & McDermott, K.B. (2009).

The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cerebral cortex, 19*(7), 1557-1566.

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia, 48*(1), 220-225.

Cools, R., Clark, L., Owen, A.M., & Robbins, T.W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22*(11), 4563-4567.

Costa, P.T. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

DePaulo, B.M. & Kirkendol, S.E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *NATO Advanced Science Institutes series. Series D: Behavioural and social sciences, Vol. 47. Credibility assessment* (51-70). New York, NY, US: Kluwer Academic/Plenum Publishers.

DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin, 129*(1), 74.

DePaulo, B.M., Stone, J.I., & Lassiter, G.D. (1985). Telling ingratiating lies: Effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology, 48*(5), 1191-1203.

DeTurck, M.A. (1991). Training observers to detect spontaneous deception: Effects of gender. *Communication Reports, 4*(2), 81-89.

DeTurck, M.A., Harszlak, J.J., Bodhorn, D.J., & Texter, L.A. (1990). The effects of training social perceivers to detect deception from behavioral cues. *Communication Quarterly, 38*(2), 189-199.

DeTurck, M.A., & Miller, G.R. (1990). Training observers to detect deception: Effects of self-monitoring and rehearsal. *Human Communication Research, 16*(4), 603-620.

De Fruyt, F., Hoekstra, H.A., & Ormel, J. (1996). *Handleiding NEO persoonlijkheidsvragenlijsten* [Manual NEO personality questionnaires]. Lisse, the Netherlands: Swets Test Services.

Diener, E. & Biswas-Diener, R. (2019). The replication crisis in psychology. In R. Biswas-Diener & E. Diener (Eds), *Noba textbook series: Psychology*. Champaign, IL: DEF publishers.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C.J., & Von Cramon, D.Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Cognitive brain research, 9*(1), 103-109.

Ekman, P. & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Consulting Psychologists Press, Palo Alto.

Ekman, P. & O'Sullivan, M. (1991). Who can catch a liar? *American psychologist, 46*(9), 913-920.

Enos, F., Benus, S., Cautin, R.L., Graciarena, M., Hirschberg, J., & Shriberg, E. (2006). Personality factors in human deception detection: Comparing human to machine performance. In *Ninth International Conference on Spoken Language Processing*.

Forrest, J.A. & Feldman, R.S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin, 26*(1), 118-125.

Gilbert, D.T. (1991). How mental systems believe. *American Psychologist, 46*, 107–119.

Gombos, V.A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, social, and general psychology monographs, 132*(3), 197-214.

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), 643-659.

Heijboer, J., Kitselaar, M., van Orsouw, W., van Rosmalen, M., Sedy, M. & Verhoeven, S. (2015). *Gender differences and confidence levels in lie detection, based on facial expressions and upper body language*. Article in preparation.

Homeland Security (2011, August 5). Screening of Passengers by Observation Techniques (SPOT) Program. Retrieved February 28, 2019, from https://www.dhs.gov/publication/screening-passengers-observation-techniques-spot-program

Hughes, C.J., Farrow, T.F., Hopwood, M.C., Pratt, A., Hunter, M.D., & Spence, S.A. (2005). Recent developments in deception research. *Current Psychiatry Reviews*, *1*(3), 273-279.

Institute for economics & peace (2017). *Global terrorism index*. Retrieved February 28, 2019, from http://visionofhumanity.org/app/uploads/2017/11/Global-Terrorism-Index-2017.pdf

Johnston, W.A., Greenberg, S.N., Fisher, R.P., & Martin, D.W. (1970). Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology*, *83*(1p1), 164.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., & Carter, C.S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*(5660), 1023-1026.

Kuzio, A. (2018). The role of sex differences in detecting deception in computer-mediated communication in English. *Journal of Computer-Assisted Linguistic Research*, *2*(1), 39-53.

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

Lewicki, R.J. & Robinson, R.J. (1998). Ethical and unethical bargaining tactics: An empirical study. *Journal of Business Ethics, 17*(6), 665-682.

Lie (n.d.). In the Oxford English Dictionary. Retrieved June 2, 2019, from https://en.oxforddictionaries.com/definition/lie

MacDonald, A.W., Cohen, J.D., Stenger, V.A., & Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*(5472), 1835-1838.

Mahon, J.E. (2007). A definition of deceiving. *International Journal of Applied Philosophy, 21*(2), 181-194.

Mahon, J.E. (2015, December 25). The definition of lying and deception. Retrieved June 2, 2019, from https://plato.stanford.edu/entries/lying-definition/

McCrae, R.R., Costa, Jr, P.T., & Lord, W. (2015). *Neo Personality Inventory - 3 (UK Edition)*(Rev. ed.). Oxford: Hogrefe Ltd.

McCrae, R.R., Costa, Jr, P.T., & Martin, T.A. (2005). The NEO–PI–3: A more readable revised NEO personality inventory. *Journal of personality assessment, 84*(3), 261-270.

Meijer, E.H., Verschuere, B., & Merckelbach, H. (2018). Failing to tell friend from foe: A comment on Wijn et al. (2017). *Legal and Criminological Psychology, 23*(1), 2-4.

Owen, A.M. (1997). The functional organization of working memory processes within human lateral frontal cortex: The contribution of functional neuroimaging. *European Journal of Neuroscience*, *9*(7), 1329-1339.

Phan, K.L., Wager, T., Taylor, S.F., & Liberzon, I. (2002). Functional neuroanatomy of

emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16*(2), 331-348.

Porter, S. & ten Brinke, L. (2010). Truth about lies: What works in detecting high-stakes deception? [Special issue]. *Legal and Criminological Psychology, 15*, 57–76.

Porter, S., Campbell, M.A., Stapleton, J., & Birt, A.R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *34*(3), 172.

Raloff, J. (2015, September 11). When a study can't be replicated. Retrieved July 20, 2019, from https://www.sciencenewsforstudents.org/article/when-study-can%E2%80%99t-be-replicated

Ramnani, N. & Owen, A.M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature reviews neuroscience*, *5*(3), 184.

Rowe, J.B., Toni, I., Josephs, O., Frackowiak, R.S., & Passingham, R.E. (2000). The prefrontal cortex: response selection or maintenance within working memory?. *Science*, *288*(5471), 1656-1660.

Schweitzer, M.E., Brodt, S.E., & Croson, R.T.A. (2002). Seeing and believing: Visual access and the strategic use of deception. *The International Journal of Conflict Management, 13*, 258–275.

Spence, S.A. (2004). The deceptive brain. *Journal of the Royal Society of Medicine*, *97*(1), 6-9.

Spence, S.A., Farrow, T.F.D., Herford, A.E., Wilkinson, I.D., Zheng, Y., & Woodruff, P.W.R. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport: For Rapid Communication of Neuroscience Research, 12*, 2849–2853.

Spence, S.A. & Kaylor-Hughes, C.J. (2008). Looking for truth and finding lies: the

prospects for a nascent neuroimaging of deception. *Neurocase*, *14*(1), 68-81.

Spencer, S.D. (2017). Examining personality factors in deception detection ability. *Psi Chi Journal of Psychological Research*, *22*(2).

Tannen, D. (1990). *You just don't understand*. New York: Ballantine.

Tilley, P., George, J.F., & Marett, K. (2005, January). Gender differences in deception and its detection under varying electronic media conditions. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (pp. 24b-24b). IEEE.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* (2nd ed.). Chichester, England: Wiley.

Vrij, A., Granhag, P.A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological science in the public interest, 11*(3), 89-121.

Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*, 253–265.

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*, 327–348.

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

Walczyk, J.J., Schwartz, J.P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person-to-person about live events: A cognitive framework for lie detection. *Personnel Psychology, 58*, 141–170.

Weinberger, S. (2010). Airport security: Intent to deceive? *Nature, 465*(7297), 412-415.

Wijn, R., van der Kleij, R., Kallen, V., Stekkinger, M., & de Vries, P.W. (2017). Telling

friend from foe: Environmental cues improve detection accuracy of individuals with hostile intentions. *Legal and Criminological Psychology*, *22*(2), 378-399.

Zeki, S., Goodenough, O.R., Spence, S.A., Hunter, M.D., Farrow, T.F., Green, R.D., ... & Ganesan, V. (2004). A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1451), 1755-1762.

Zimmerman, L. (2016, March). Deception detection. Retrieved March 3, 2019, from https://www.apa.org/monitor/2016/03/deception

Zuckerman, M., DePaulo, B.M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In *Advances in experimental social psychology* (Vol. 14, pp. 1-59). Academic Press.

Appendix A

INFORMED CONSENT

This research is conducted by Julie Müller and Lisa Ebben. We invite you to participate in this research project, which is part of our bachelor's thesis. We are investigating whether people are able to detect hostile intentions. Research already showed that people are indeed quite good at spotting malicious intent, but insight is not yet gained in the indicators that people use to decide whether someone has hostile intentions. In our study, we therefore want to gain more information about what people focus on when determining if someone has hostile intentions. Furthermore, we are examining the role of personality traits and gender in the ability to detect.

Procedures

You will participate in a study lasting approximately 90 minutes. You will watch videos of people walking down a hallway with a laptop case towards a security checkpoint where the person's bag will be checked by a security man. Either their bags contain something illegal (a bomb) or legal (books) and it is up to you to decide whether people are carrying the bomb (i.e. have hostile intentions), or the books. Afterwards you will be asked to fill in a questionnaire. Please be aware that you need to be at least 18 years old to participate in our study.

Potential Risks or Discomforts

There are no obvious physical, legal or economic risks associated with participating in this study. You do not have to answer any questions you do not wish to answer. Your participation is completely voluntary and you are free to discontinue your participation at any time.

Confidentiality

Your privacy will be protected to the maximum extent allowable by law. No personally identifiable information will be reported in any research product. Moreover, only trained research staff will have access to your responses. Within these restrictions, results of this study will be made available to you upon request. At the start of the research your name will be replaced by a pseudonym; your name will be coded. At the end of the study you can put your email on a list, if you wish to receive your results of the study.

STATEMENT OF CONSENT

Your signature indicates that you are at least 18 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction and you voluntarily agree that you will participate in this research study. I agree to participate in a research project led by Julie Müller & Lisa Ebben. The purpose of this document is to specify the terms of my participation in the project through being interviewed.

1. I have been given sufficient information about this research project. The purpose of my participation in this project has been explained to me and is clear.
2. My participation in this project is voluntary. There is no explicit or implicit coercion whatsoever to participate.
3.I have the right not to answer any of the questions. If I feel uncomfortable in any way during the research, I have the right to withdraw from the study.
4. I have been given the explicit guarantees that, if I wish so, the researcher will not identify me by name or function in any reports using information obtained from this study, and that my confidentiality as a participant in this study will remain secure.
5.I have been given the guarantee that this research project has been reviewed and approved by Julie Müller & Lisa Ebben and by the BMS Ethics Committee. For research problems or any other question regarding the research project, the Secretary of the Ethics Commission of the faculty Behavioural, Management and Social Sciences at University Twente may be contacted through ethicscommittee-bms@utwente.nl.
6.I have read and understood the points and statements of this form. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this study.

If you have any complaints about this research, please direct them to the secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences at the University of Twente, Drs. L. Kamphuis-Blikman P.O. Box 217, 7500 AE Enschede (NL), telephone: +31 (0) 53 489 3399; Email: l.j.m.blikman@utwente.nl). For general questions about the participation, confidentiality or data processing, direct them to Julie Müller, j.c.muller@student.utwente.nl

INSTRUCTION

Every answer you give in this research will be dealt with confidentially and anonymously.Please do not discuss with other people during this research.

This research is about recognising incidents. The goal is that you will try to recognise deviant (odd) behaviour. This is the behaviour that precedes an incident.

In a few minutes, you will watch a lot of short videos. In these videos, you will see people who are participating in a research to test the security of the Marechaussee (an institute that is concerned with the security of the Netherlands).

Each person in the videos carries a laptop bag. *Half of the people carry a bag with an explosive*, which they try to smuggle past the Marechaussee.

We would like to know if you are able to detect who these people are, by looking at who is showing deviant behaviour.

You will get a point for every explosive-smuggler you recognise correctly, but one point will be subtracted each time you think an innocent person is a smuggler.

Please look at the behaviour of the people very carefully. Sometimes it may feel as if you are taking a wild guess. Still, try to do it seriously and trust on your intuition.

You will first see 5 videos to practice with.

A tip: pay special attention to the last part of the hallway, where the black and white chequered board is hanging on the right side. Here you will sometimes see more deviant behaviour.

When you are ready and have understood the instructions, please give the researcher a sign.

DEMOGRAPHICS

How old are you?

What is your gender?

- Male
- Female
- Other

What is your nationality?

- Dutch
- German
- Other

What do you study?

- Psychology
- Communication Sciences
- Other

PARTICIPATION

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I took my participation in this study seriously.	•	•	•	•	•
During the study, I was able to concentrate.	•	•	•	•	•
During the study, I was not distracted.	•	•	•	•	•

BEHAVIOURAL INDICATORS

What behavioural indicators did you use for your decision?

Please shortly describe which features of the person you used to decide that he or she has hostile intentions.

BEHAVIOURAL INDICATORS

In the following, we want to find out how specifically people can detect hostile intentions. We are interested in the bodily indicators that show the observer if a person carries a bomb or only books.

Therefore, please try to remember these clips again and think about what you used to make your decision and then fill out the following questions.

INDICATORS OF HOSTILE INTENTIONS

Some short multiple-choice questions are presented about which features you used to determine if a person carries an explosive with them. Please indicate whether you used each of these in at least one of the videos. For each question you can check multiple features.

Did you use the body language of the person as an indication for hostile intentions?
If yes, which of the below were applicable?

- o The person was trying to look relaxed.
- o The person was trying too much to behave normally.
- o The person tried to hide him- or herself.
- o The person seemed uncomfortable.
- o The person seemed concerned.
- o The person displayed closed body language.
- o The person's behaviour was conspicuous.
- o I used none of the above.

Did you use the looking behaviour of the person as an indication for hostile intentions?
If yes, which of the below were applicable?

- o The person was looking backwards.
- o The person was looking towards the sides.
- o The person was looking around.
- o I used none of the above.

Did you use the person’s use of hands as an indicator?
If yes, which of the below were applicable?

- o The person was fiddling at their clothes.
- o The person had their hands in their pockets.
- o The person put their bag on the ground when reaching the checkpoint.
- o I used none of the above.

Did you use the person’s walking style as an indicator?
If yes, which of the below were applicable?

- o The person was walking slowly.
- o The person was walking too fast.

- The person had a normal way of walking.
- The person walked down the hallway in a straight line.
- The person's walking style was determined.

Did you use the facial expressions of the person as an indicator for hostile intentions?
If yes, which of the below were applicable?

- The person was smiling.
- The person was avoiding eye contact.
- The person was looking at the ground.
- The person was trying to look cool.
- The person had relaxed facial expressions.
- I used none of the above.

INDICATORS OF INNOCENCE

We would now like to know which factors indicated for you that someone was innocent and did not carry a bomb with them.
Did you use the body language of the person as an indicator for innocence?
If yes, which of the below were applicable?

- The person was relaxed.
- The person behaved normally.
- The person did not try to hide him- or herself.
- The person seemed comfortable.
- The person did not seem concerned.
- The person displayed open body language.
- The person was unobtrusive.
- The person behaved confidently.
- I used none of the above.

Did you use the looking behaviour of the person as an indicator for innocence?
If yes, which of the below were applicable?

- The person was looking backwards.
- The person was looking towards the sides.
- The person was looking around.
- I used none of the above.

Did you use the person's body gestures as an indicator for innocence?

If yes, which of the below were applicable?

- The person was fiddling at their clothes.
- The person had their hands in their pockets.
- The person put their bag on the ground when reaching the checkpoint.
- I used none of the above.

Did you use the person's walking style as an indicator for innocence?

If yes, which of the below were applicable?

- The person was walking slowly.
- The person was walking too fast.
- The person had a normal way of walking.
- The person walked down the hallway in a straight line.
- The person's walking style was determined.
- I used none of the above.

Did you use the facial expressions of the person as an indicator for innocence?

If yes, which of the below were applicable?

- The person was smiling.
- The person was avoiding eye contact.
- The person was looking at the ground.
- The person was trying to look cool.
- The person had relaxed facial expressions.
- I used none of the above.

OTHER BEHAVIOURAL INDICATORS

Are there any other behavioural indicators that are not mentioned above but were important during your decision while watching the videos?

The NEO-PI-R statements are left out because spreading them is forbidden.

DEBRIEFING

You came to the end of our study, thanks for your time and effort. Before ending your participation we want to inform you about some aspects of this study we did not tell you in the beginning in order to not influence your results.

The participants that you saw in the videos were manipulated by a cue: a static noise (similar to a walkie-talkie sound) was used when the participants walked by to influence their ability to hide their hostile intentions. Besides, we tried to increase participants' cognitive load by having them count their steps. With this cue and cognitive load manipulation, we expected that it would become harder for these participants to hide their hostile intent and, thus, that it would be easier for you to detect their hidden intentions.

If you have further questions please feel free to contact Julie Müller (j.c.muller@student.utwente.nl) or Lisa Ebben (l.ebben@student.utwente.nl).