Does it show when someone has hostile intentions? When in doubt, replicate!

Name: Annemiek Fokkens
Student number: s1762311
Date: 20-06-2018
Mentor: Dr.Ir. P.W. de Vries
Second mentor: Dr. J.M. Gutteling
University of Twente

Abstract

Security officers have the important task to keep public places safe by observing and detecting people with possible hostile intentions. Findings of previous search about this subject are diverse and ambiguous, some researchers have stated people are able to observe and detect deviant behavior, whereas others state that people are no better than chance level in detecting hostile intentions. A recent study from Wijn, Van der Kleij, Kallen, Stekkinger and De Vries (2017) about hostile intentions, however, stated that some conditions could enhance detection levels. The current study will replicate the study from Wijn et al. (2017) to see if cognitive load and environmental cues could improve detection performance. In addition the original study, police officers took part in the experiment, and a comparison with college students is made concerning hostile detection. No main or interaction effects from the different conditions on detection performance were found after performing a 2 x 2 repeated measures ANOVA. Furthermore, there was no difference in detection levels when comparing police officers with college students. However, work experience seemed to be a significant predictor of the detection performance. This current study had different findings from the original study from Wijn et al. (2017). Replication is crucial and in order to get more insight in this field of psychology, future research on this topic is suggested.

Introduction

On a daily basis, security officers have to judge people's intentions at security checkpoints, border security and other important checkpoints. With training and the help of security devices like body scanners they have to detect potential deceptive behavior. In the USA they have been using a program called Screening Passengers by Observation Technique (SPOT) (Weinberger, 2010). With this technique, they claim that officers can identify people who could be a possible threat for the safety of the airport. A great number of researchers are skeptical of these kind of methods and state, according to Weinberger (2010), that there is too little scientific evidence to support these kind of techniques. Even more, research has shown after analyzing the accuracy of deception judgements, that people achieve an average of 54% correct lie-truth judgements (Bond & DePaulo, 2006). However, other researchers have shown that there could be scientific evidence for judging deceptive intentions. In this study, we will look into detecting deceptive behavior and the fundamental and scientific base for this argument. The research question of this study is: are people able to detect hostile intentions by observing behavior, and are there any differences between experts and lay people?

According to DePaulo et al. (2003) deception is defined as a deliberate attempt to mislead others. When people are deceiving others, there are many different behavioral cues they could show (DePaulo et al., 2003; Ekman & Friesen, 1969; Zuckerman, DePaulo & Rosenthal, 1981), which could lead to deviant behavior. When trying to detect this deviant behavior, it is crucial to have a clear image of what deviant behavior is and what the theoretical base is. Deviant behavior is any behavior that is contrary to the dominant norms of society (Crossman, 2018). There are many different theories which support and explain why people enact deviant behavior. For example, a theory which explains deviant behavior is the Labelling Theory (Crossman, 2018; Wellford, 1975). According to Tannenbaum (1938) deviant behavior could be a form of self-fulfilling prophecy, because people expect a person to enact criminal behavior, the person will do so. Even though this theory was first mentioned years ago, it is still one of the more prominent explanations of deviant behavior (Crossman, 2018). Deviant behavior may or may not become behavior driven by hostile intentions.

Over the last few years, there is a growing interest in preventing criminal behavior by detecting deviant behavior which is caused by hostile intentions (Elias, 2009). Research has shown particular behavior that could be linked to hostile intentions at train stations. Behaviors that are seen linked to hostile intentions are, for example: deviant walking patterns, prohibited actions such as making photos, and waiting without taking the train (Burghouts et al., 2011).

Terrorists, for example, are known to enact hostile reconnaissance, which likely leads to experiencing more stress what could be detected, due to behavioral changes (Eachus, Stedmon & Baillie, 2013). Detecting hostile intentions by observing behavior, however, has been a controversial subject (Ormerod & Dando, 2015; Weinberger, 2010).

However, there have been people who developed a system to train people to detect people with hostile intentions. One example of such a program is Search, Detect & React (SDR), founded by Ran Cohen. The idea behind the company is to train people as a form of "human engineering" (SDR, 2018). Trained people should be more aware of their surrounding and people's behavior. According to SDR (2018), in order to detect deviant behavior, it is important to establish normal behavior in that environment. Besides establishing the normal situation, there are more crucial aspects in detecting deviant behavior. In the SDR training, people are taught that there are three characteristics that play a part in observing behavior: eyes, hands and muscle tension. For example, when a person is covering his hands, he could be carrying a weapon and thus be a possible threat. So, according to SDR there is a way to detect deviant behavior through trained observers. This is in line with previous mentioned research and the study by Wijn et al. (2017).

Training programs as mentioned above may beg the question if people could be more able to detect hostile intentions when being trained. For example, this would mean that police officers would be better in detecting hostile intent than non-experts. According to Bond and DePaulo (2006), generally speaking, people are poor at the detection of possible deceit. This is in line with conclusions of Stromwall, Granhag and Hartwig (2004) who stated that people's ideas about deception clues are not matching with the reality of deceit. Research has shown that police officers are not significantly better in lie detection when being compared to lay people (Aamodt & Custer, 2006; DePaulo & Pfeifer, 1986; Ekman & O'Sullvian, 1991). Furthermore, accuracy in lie detection was not related to work experience or experience with interviewing suspects (Vrij & Mann, 2008). In line with these studies, there can be concluded that detecting lies and deception is not something that can be trained.

However, other research suggests quite the opposite. For instance, Ekman, O'Sullivan and Frank (1999) showed that professionals were highly accurate when judging videos of people who were or were not lying. Also, the previous mentioned research of Ekman and O'Sullivan (1991), showed that Secret Service agents were more accurate than chance level. This is in line with research that suggested that training does raise accuracy levels when detecting lies (Frank & Feeley, 2011). The reason why findings are so diverse could be the

difference between high- vs low-stake situations, where police officers have been proven to be better lie detectors in high-stake situations (Wright & Wheatcroft, 2017). Another reason what could explain the differences, is the sort of cues people tend to focus on. For example, people, police officers as well as lay people, often try to detect liar by looking at non-verbal cues like gaze aversion, whereas verbal cues have been proven to be more accurate (Bogaard, Meijer, Vrij & Merckelbach, 2016; Vrij & Mann, 2000). This is in line with research of Mann, Vrij and Bull (2004) where accuracy was negatively correlated with behavioral, non-verbal cues, such as gaze aversion and fidgeting. Parallel with this finding, highly trained police officers with a high accuracy level of lie detection mainly focused on verbal cues (Mann et al., 2004). The latter findings suggests that experts are better in detecting deception than lay people, which suggests that deception could be detectable by observing behavior.

However, research mentioned above is primarily about deceit in the sense of lie detection, but not much is known about other kinds of deceit. Recent research about deviant behavior have tried to link deviant behavior to visible cues. In the study of Wijn, Van der Kleij, Kallen, Stekkinger and De Vries (2017) participants were informed to either bring a book or a dummy explosive device through a security checkpoint. Whilst in line for the checkpoint they were filmed, and other participants, called judges, had to decide whether or not the person was guilty of smuggling an explosive device. Other variables were cognitive load (high vs. low) and cues (present vs. absent). In one condition (cue present) participants passed a device that produced a noise resembling a police walkie-talkie, while they walked through the corridor towards the checkpoint. Especially with this manipulation, they showed in their research that environmental cues that signal risk of exposure can improve detection of hostile intentions.

Current study

In the current study, there will be a replication of the study of Wijn et al. (2017). According to Wijn et al. (2017) increasing environmental cues can improve detection of hostile intent. However, Meijer, Verschuere and Merckelbach (2018) criticized the Wijn et al. (2017) research, stating that the judges were unable to detect hostile intent. In order to get more clearance about the subject, replication is essential. Therefore, the first two hypotheses of the current study will be the same as in the original study:

Hypothesis 1: Environmental cues signaling 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 behaviors, and activation control of desirable behaviors is hampered.

In addition to the original study, the current one will also look at the possibility that there could be a difference between experts and lay people when it comes to judging deviant behavior in combination with hostile intentions. Previous research outcomes have been ambiguous concerning this question. Besides this difference, the effect of possible training programs will be included in this research as well. Subjects who have had specific training will be compared to subjects that did not follow any kind of training to see if there could be a difference between these groups. This leads to the following exploratory research questions:

Research Question 1: Is there a difference between police officers and college students in recognizing hostile intent? And if so, could work experience influence their detection performance?

Research Question 2: Are subjects who have had extra training in detecting hostile intentions able to perform better when judging possible wrongdoers than subjects who have not had such training?

Method

Design

A between-subjects experimental design was employed. Group 1 consisted of college students, group 2 consisted of police officers. The results were analysed using a mixed design approach. Participants from both groups were asked to judge videos form earlier research. They were shown video material of participants from an earlier 2 x 2 x 2 design with experimentally induced hostile (vs. non-hostile) intentions, with high (vs. low) cognitive load, and who were exposed to an environmental cue (vs. no cue) while they walked a pre-defined route. The independent variables were environmental cues (cue vs. no cue) and mental load (load vs. no load). The dependent variable was recognition of hostile intention (intent vs. no intent) by the judges. The judges were unaware of these conditions.

Participants

There were 25 participants in total, consisting out of 11 police officers and 14 college students. The mean age of all participants was 30.64 (*SD*=11.32). The mean age of the police officers was 40.00 (*SD*=11.27), and the mean age of the students was 23.29 (*SD*=2.55). There were 14 men and 11 women participating. All participants were recruited through convenience sampling. Police officers from a police station in Amsterdam were directly approached by their superior and asked if they would be willing to take part in the experiment. Two of the police officers had followed an extra training to detect deviant behaviour, the so-called "Spottersopleiding". Students participated in exchange for Bachelor course credits. Participants had to be at least 18 years old and speak the Dutch language fluently.

Materials

For the study, videos from previous research of Wijn et al. (2017) were used. These videos contained clips of participants walking through the corridor towards a checkpoint. Approximately half of them were carrying books in their bag, whereas the other half was carrying a dummy bomb device in a bag. Participants that carried a dummy explosive were motivated to not stand out at the checkpoint because of a potential loss (i.e., risk of time-consuming questionnaires) and a potential gain (i.e., chance of winning 100 euros). In the current study, the judges had to watch these videos and indicate after every video if they thought this person was innocent or not. These videos were shown to the judges through the use of Matlab. In addition to the study of Wijn et al. (2017), participants had to fill out a questionnaire. The questionnaire was shown in Qualtrics.

Procedure

First of all, participants were welcomed and were seated behind a laptop. A brief explanation of the current study was given to them, to give them an impression of what was expected of them. After this explanation, they were instructed that they had to answer a few questions at first, then they would watch the videos and after that they would had to answer a few more questions. In Qualtrics, first of all, an informed consent was shown. When participants agreed to these conditions, a brief explanation of the study was given in Qualtrics as well. Then they had to fill out a few question about their demographic data, work or study related questions, and an open question about what kind of behaviour they would describe as suspicious if they would work as a security guard at an airport. When they arrived at the page where was stated they were going to watch videos they had to give the researcher a heads up

(see Appendix A). Then, Matlab was started and the original instruction of Wijn et al. (2017) was shown. In this instruction, it was mentioned that the ratio of participants who were carrying books or dummy bombs were 50/50. After this instruction, participants were able to starts with five practice videos. After every video participants had to choose: was this person carrying a bomb or was this person carrying books? Because of a lack of time and concentration, police officers watched less videos than the college students did. However, the videos that they watched were selected in a way that they still saw an even distribution of all the different conditions. After the participants watched all the videos, Matlab closed automatically and they had to give another heads up to the researcher. Qualtrics was shown again and they had to fill out a few more questions. They were asked to describe the specific behaviour which made them decide that people were carrying a bomb. This question, as well as the previous open question, was part of a different, related study. These questions will not be discussed and analysed any further in the current study. Finally, they were asked about their participation and concentration whilst watching the videos.

Analysis

The data was analysed using mixed methods. To answer the two hypotheses, the raw data of each participant was exported from Matlab to Excel. Two files could not be exported due to a faulty results-file and one due to a faulty test run, it concerned one police officer and one college student. For the remaining participants, the answers of 15 videos had to be deleted because the conditions of these videos were unknown. All files were merged and demographic data that was measured by the separate questionnaire in Qualtrics was added manually. To be able to give feedback to the participants about how they had performed, their scores were calculated with an Excel code.

Then, all data was imported to SPSS. Descriptive statistics and frequencies of age, gender, group, function, and work experience were examined. A new variable was added to differentiate the conditions, numbered 1-8 (hostile intent vs non-hostile intent, cue present vs absent and high vs low cognitive load). Next, the hit rates (H) and false alarm rates (F) for each participant and for each condition were calculated. The hit rate was the number of people rated guilty divided by the number of people with hostile intent and the false alarm rate was the number of people rated guilty divided by the number of people with non-hostile intent. These rates were then entered into an Excel file by Zhang and Mueller (2005, see https://sites.google.com/a/mtu.edu/whynotaprime/), which was provided to calculate A. A, being the improved non parametric index of sensitivity, was calculated for each condition of

the 2 (cognitive load: high vs. low) x 2 (cue: present vs. absent) design.

The calculated values were manually entered back into SPSS. Q-Q plots were created to examine whether the *A* scores from the different conditions were normally distributed, which was required for the repeated measures analysis. After plotting the *A* scores, a 2 x 2 repeated measures ANOVA, with the four conditions (cue & cognitive load, no cue & cognitive load, cue & no cognitive load, no cue & no cognitive load) as independent variables and the *A* scores of each condition as dependent variables was performed. This 2 x 2 repeated measures ANOVA was conducted to examine if there were significant differences in *A* scores between the different conditions.

In order to answer the explanatory research questions, first of all an independent samples t-test was performed. The independent variable in this t-test was function (student vs police officer) and the dependent variable the mean scores. Second of all, a one-way ANOVA was performed to measure a possible influence of years on the job on the mean scores of the police officers on hostile detection. At last, an independent samples t-test was executed once more to measure if training had a positive influence on hostile detection. Here, the independent variable was training (yes or no) and the independent variable the mean scores of the police officers.

Results

When looking at the results, one participants stands out because this person judged every person in the videos as non-guilty. This results in an average score of 51,6%, but one could question the meaning of this participant's data. Because of the fact that this participant was instructed that the distribution of guilty vs non-guilty was 50/50, this participant was deleted out of the dataset. The same goes for one student, who only filled out 8 participants to be guilty. So in total, two participants were removed from the dataset because of deviant answering patterns. Furthermore, in order to test the participation level, all participants were asked to fill out how seriously they took part in the study on a five point Likert scale. There were no outliers, and all participants scored above average on the five point Likert scale (*M*=4.39, *SD*=0.68). Police officers scored higher (*M*=4.80 (*SD*=0.28) than students did (*M*=4.02, *SD*=0.73).

Hits and misses

Because participants wanted to know how they scored, every score was calculated. The mean score of all participants (M=0.47) was slightly below chance level. To have a better indication of the dataset the table 1 was generated. Most common amongst all participants

were false negatives (*M*=.29, *SD*=.05). The least common were true positives (*M*=.21, *SD*=.05). Overall, participant judged the videos as negative more than they judged the videos as positive. This resulted in more true negative answers (*M*=.26, *SD*=.06) than in true positives (*M*=.21, *SD*=.05). When looking more closely at the answer patterns of the police

Table 1
Descriptive Statistics: hits and misses amongst all participants

	N	Minimum	Maximum	Mean	Std. Deviation
True Positive	23	,10	,31	,2161	,05425
False Positive	23	,10	,36	,2370	,05996
True Negative	23	,14	,39	,2576	,05870
False Negative	23	,19	,42	,2887	,05627

officers, the same can be stated (table 2). The differences between the four answer categories in the police participants, however, are smaller than amongst all the participants. Furthermore,

Table 2
Descriptive Statistics: hits and misses amongst police officers

	N	Minimum	Maximum	Mean	Std. Deviation
True Positive	10	,17	,31	,2210	,04654
False Positive	10	,19	,36	,2580	,05633
True Negative	10	,14	,31	,2420	,05633
False Negative	10	,19	,33	,2790	,04654

police officers gave more false positive answers than true negative answers, whereas that was the other way around for college students (table 3). So where college students had the

Table 3
Descriptive Statistics: hits and misses amongst college students

|	N	Minimum	Maximum	Mean	Std. Deviation
True Positive	13	,10	,30	,2123	,06112
False Positive	13	,10	,31	,2208	,05965
True Negative	13	,18	,39	,2696	,05981
False Negative	13	,21	,42	,2962	,06358

tendency to judge the participants more as negative, this was not the case for the police officers.

Hostile intent

To measure how the different conditions influenced the detection of hostile intent, a sensitivity measure (*A*) for all of these conditions was calculated. First of all, a Q-Q plot was made to examine whether *A* were normally distributed (figure 1). After it was established that

Figure 1

Q-Q plot of A
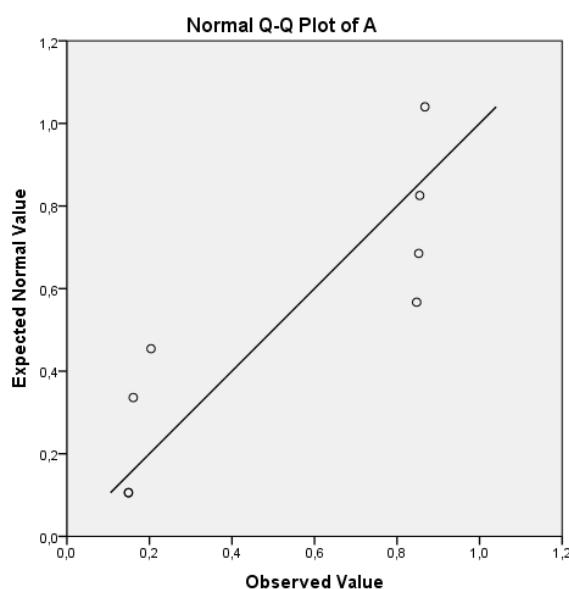
A was normally distributed, a 2 x 2 repeated measure ANOVA was conducted. Independent variables were the four conditions (low vs high cognitive load and cue present vs absent) and the dependent variable were the *A*-scores of every condition. As in the original study, it was expected that high cognitive load and a present environmental cue would increase hostile intent detection. First of all, when analyzing the results there was no significant interaction effect of *A* with the variables cognitive load (high vs low) and cues (present vs absent), $F(1,1)=0.099$, $p=.806$. Furthermore, no significant simple main effect was found for the cognitive load conditions, $F(1,1)=0.339$, $p=.664$. At last, the simple main effect for the environmental cues was also not significant, $F(1,1)=2.020$, $p=.390$. So, there were no significant differences amongst conditions, which results in rejecting hypotheses 1 and 2.

Differences between groups

Furthermore, an independent samples t-test was performed for the explanatory research question, with independent variable function (student vs police officer) and dependent variable detection score. This t-test was conducted to examine whether there could

be a difference between police officers and college students concerning their detection score. When comparing the group of students with the group of police officers, no significant difference was found, $t(23)=-,404$, $p=,690$. To see if there could be a correlation of the work experience to the detection performance of the police officers, a linear regression was performed with the years on the job as an independent variable and the mean score of only the police officers as a dependent variable. This regression analysis showed that number of years on the job was a significant predictor for their detection score, $F(1,9)=5.595$, $p=.042$.

Table 4

Regression analysis working experience

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,017	1	,017	5,595	,042[b]
	Residual	,028	9	,003		
	Total	,046	10			

a. Dependent Variable: Score2
b. Predictors: (Constant), Dienstjaren

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	,550	,038		14,536	,000
	Dienstjaren	-,006	,002	-,619	-2,365	,042

a. Dependent Variable: Score2

In order to compare police officers who had training in detecting deviant behavior with officers who did not have such training, an independent samples t-test was performed. The independent variable being training (yes vs no) and the dependent variable being the mean detection score. Unfortunately, only two officers had followed training concerning detection of hostile intent. When comparing these two police officers to the rest of the police officers, no significant differences were found, $t(9)=-,340$, $p=,742$.

Discussion

When analyzing hostile intent detection, the different conditions were not significant. This means that judges did not score better when participants were administered a cue or had a high cognitive load during the experiment, as was originally found in the study of Wijn et al. (2017) and as was hypothesized in the current study. These results led to rejection of both the

hypotheses. Furthermore, when comparing the two different groups of police and college students, no significant differences were found in their detection performances. This means that police officers were not better at detecting deceit than the college students in this study. Initially it was suggested that they could score better than the college students. However, when analyzing the number of years on the job as a predictor for the detecting performance of the police officers, it showed to be a significant predictor. This could mean that work experience could increase the ability to deceit detection. Finally, police officers who have had training in detecting deceit were compared to police officers who did not follow these kind of training. The difference between the groups were not significant, however the number of police officers who have had these training was only three out of eleven. The analysis of this research question was insufficient, because it would have been better if the number of police officers who had followed these training programs were higher.

Previous research has stated diverse opinions about the possibility of people being able to detect hostile intentions. Furthermore, some researchers stated that this could be trained, whereas others stated this is not possible. In the current study, the findings were diverse as well. First of all, the different conditions did not influence the detection performance of the judges. Second of all, no difference was found between detection levels of college students and police officers. This finding is in line with research that claims hostile detection through observation is nor trainable nor possible. However, during the research, police officers stated they did not see a match between their daily tasks on the job and the task of the experiment. Some of the police officers suggested to let security guards participate in the experiment as well. Perhaps, in future research this would be a useful addition. Third of all, years on the job seemed to influence detection skills amongst police officers. This could possibly mean that people can learn this skill and it is trainable. On the other hand, having had training in this area did not increase detection levels. These latter two findings are contradictory, which could be looked into in future research as well. However, out of a population of ten police officers, only two had followed this kind of training. Perhaps these two police officers were randomly better than the other officers, which could explain this finding. To make a reasonable statement about this explanatory question, more research is needed.

In comparison with the original study, the participants differed. Where Wijn et al. (2017) only had college students as participants, the current study also had police officers participating. Literature might have suggested the possibility that police officers would score better than lay people, however this could not be concluded in this study. A possible

explanation of why police officers scored slightly worse than college students, could be due to concentration issues. Even though they scored above average on the question about their concentration levels, almost every single one of the participants commented on the length of the study. They mentioned that they were having trouble concentrating on one task for such a long time. The average response time of the police officers was approximately 45 minutes. For the students, this was at least over 90 minutes. Perhaps the design of this study was more designed for the attention span of students than for police officers, who are used to doing more practical tasks. If this were to be researched in the future, it could be suggested that the task of the experiment is more adjusted to the attention span of police officers.

Furthermore, a lot of the participants, especially police officers, complained about the videos they had to watch. They claimed that they were not observing real behaviour, instead they were watching people act out a role. For example, participants in the videos started to laugh when walking up to the checkpoint. Perhaps it could be possible to make use of actual footage of people with hostile intentions in future research. This could give a more clear picture of actual deviant behaviour, rather than people acting out possible deviant behavior. As reviewed in the introduction, people often have a false idea of suspicious behavior what does not match the reality (Bogaard, Meijer, Vrij & Merckelbach, 2016; Vrij & Mann, 2000). This tendency could maintain these false beliefs, namely people in the original study thought they knew how to act as a criminal, but in reality this behavior does not match actual criminal behavior. So perhaps it would be more meaningful to use video materials of actual criminal behavior in future research. In that case, the idea if people are able to observe and detect hostile intentions could be tested more realistically. Judges would then be observing actual behavior and possibly detect actual criminals.

Replication crisis

The current study was a replication of the study of Wijn et al. (2017). Reason for this replication was amongst others a critical note of Meijer et al. (2018), who commented that the original findings were controversial. The original researchers then wanted to replicate the study, to research of their findings could be expanded and be proven again. Recently, a lot of psychological studies have had critical notes because findings were not replicated for example. Although there have been many influential studies in the history of psychology, psychology is now facing a replication crisis. First of all, many studies have never been replicated, which does not enhance the quality of the findings of the research (Keps & McDaniel, 2013; Shrout & Rogers, 2018). Second of all, when studies are indeed replicated,

findings often differ from the original study (Maxwell, Lau & Howard, 2015). So, in order to make sure that findings actually prove something, replication is crucial. Especially when it comes to more controversial subjects like visible deviant behavior on people with hostile intentions. That is why in this study, there was a replication conducted of the previous mentioned study from Wijn et al. (2017) to research if the found effect would be the same again. In the current study, the findings differed from the original study. That is also why it is important to conduct further research in this area of the field, and even more important to keep replicating these studies.

Literature

Aamodt, M., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15*(1), 6–11.

Bogaard, G., Meijer, E. H., Vrij, A. & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PLoS ONE, 11*(6), e0156615. doi: 10.1371/journal.pone.0156615

Bond, Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234. doi:10.1207/s15327957pspr1003_2

Burghouts, G. J., den Hollander, R., Schutte, K., Marck, J., Landsmeer, S. & den Breejen, E. (2011). Increasing the security at vital infrastructures: Automated detection of deviant behaviors. Proc. SPIE 8019, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X, 80190C doi: 10.1117/12.884579

Crossman, A. (2018). Sociological explanations of deviant behavior: A look at four different theories. Retrieved from: https://www.thoughtco.com/sociological-explanations-of-deviant-behavior-3026269

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

DePaulo, B. M. & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology, 16*(3), 249–267. doi: 10.1111/j.1559-1816.1986.tb01138.x

Eachus, P, Stedmon, A. & Baillie, L. (2013). Hostile intent in public crowded spaces: A field study. *Applied Ergonomics, 44*(5), 703-709. doi: 10.1016/j.apergo.2012.05.009

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*(1), 88–106. doi: 10.1080/00332747.1969.11023575

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*(9), 913-920. doi: 10.1037/0003-066X.46.9.913

Ekman, P., O'Sullivan, M. & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10*(3), 263-266. doi: https://doi.org/10.1111/1467-9280.00147

Elias, B. (2009). *Airport passenger screening: Background and issues for Congress* (No. R40543), Congressional Research Service.

Frank, M. G. & Feeley, T. H. (2011). To catch a liar: Challenges for research in lie detection

training. *Journal of Applied Communication Research, 31*(1), 58-75. doi: 10.1080/00909880305377

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in industrial and organizational psychology? *Industrial and Organizational Psychology, 6*(3), 252–268. doi: 10.1111/iops.12045

Mann, S., Vrij, A. & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89*(1), 137-149. doi: 10.1037/0021-9010.89.1.137

Maxwell, S. E.,Lau, M. Y.,Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70*(6), 487-498.

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. doi: 10.1111/lcrp.12118

Ormerod, T. C., & Dando, C. J. (2015). Finding a needle in a haystack: Toward a psychologically informed method for aviation security screening. *Journal of Experimental Psychology: General, 44*(1), 76-84. doi:10.1037/xge0000030

SDR (2018). Retrieved from: https://sdr.eu.com/nl/over-sdr/

Shrout, P. E. & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual review of Psychology, 69*(1), 487-510.

Stromwall, L. A., Granhag, P. A., & Hartwig, M. (2004). *Practitioners' beliefs about deception*. Cambridge: Cambridge University Press.

Tannenbaum, F. (1938). *Crime and Community*. London and New York: Columbia University Press.

Vrij, A. & Mann, S. (2000). Telling and detecting lies in a high-stake situation: the case of a convicted murderer. *Applied Cognitive Psychology, 15*(2), 187–203. doi: 10.1002/1099-0720(200103/04)15:2<187::AID-ACP696>3.0.CO;2-A

Vrij, A. & Mann, S. (2008). Who killed my relative? police officers' ability to detect real-life high-stake lies. *Psychology, Crime & Law, 7*(2), 119-132. doi: 10.1080/10683160108401791

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

Wellford, C. (1975). Labelling Theory and Criminology: An assessment. *Social Problems,*

22(3), 332–345. doi: 10.2307/799814

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. doi: 10.1111/lcrp.12107

Wright, C. & Wheatcroft, J. M. (2017). Police officers' beliefs about, and use of, cues to deception. *Journal of investigative psychology and offender profiling, 14*(3), 307-319. doi: 10.1002/jip.1478

Zhang, J. & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70*, 203–212. doi:10.1007/s11336-003-1119-8

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. InL. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, 1–59). New York: Academic Press.

Appendix A: Questionnaire part 1

Does it show when someone has hostile intentions?

Hartelijk bedankt voor het deelnemen aan ons onderzoek!

In het kader van onze Bachelorthese bij de afdeling Psychologie van Conflict, Risico en Veiligheid aan de Universiteit Twente gaan wij onderzoeken in hoeverre mensen met een vijandige intentie herkend kunnen worden van derden. Eerder onderzoek heeft aangetoond dat daders in sommige situaties (bijvoorbeeld, als zij worden afgeleid) vaker door derden worden herkend.

 Echter is er een gebrek aan onderzoek over potentiële verschillen tussen experts, die een opleiding hebben gevolgd met nadruk op het herkennen van misdaders, en leken. Verder weten wij niet welke nonverbale gedragingen van de daders hun ontmaskeren. Daarom onderzoeken wij drie dingen: Of storend geluid en cognitieve inspanning tot vaker herkennen leidt, of politieagenten beter zijn dan studenten in het herkennen, en welke non-verbale gedragingen tot het herkennen leiden.

 Het onderzoek duurt in totaal ongeveer 90 minuten. Ten eerste vragen wij u om een aantal vragen te beantwoorden, daarna ziet u een aantal video's waarin personen een tas met een bepaalde inhoud over een gang moesten smokkelen. Sommige tassen hadden een illegale inhoud (zoals een bom), sommige hadden een legale inhoud (zoals boeken). U moet gewoon beslissen of de persoon die u per video ziet schuldig of niet schuldig is.

Toestemmingsverklaring

Ik verklaar op een voor mij duidelijke wijze te zijn ingelicht over de aard, methode, en doel van het onderzoek. Ik weet dat de gegevens en resultaten van het onderzoek alleen anoniem en vertrouwelijk aan derden bekend gemaakt zullen worden. Mijn vragen zijn naar tevredenheid beantwoord. Ik begrijp dat mijn antwoorden uitsluitend voor analyse en wetenschappelijke presentaties zal worden gebruikt. Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud me daarbij het recht voor om op elk moment zonder opgavevan redenen mijn deelname aan dit onderzoek te beëindigen.

◯ Akkoord (1)

Wat zijn uw initialen en de laatste twee nummers van u geboortedatum?
(Bijvoorbeeld, *Anna Schulz, geboren 01.01.1980*: **AS80**)

_____

Wat is uw leeftijd?

_____

Wat is uw geslacht?

◯ Man (1)

◯ Vrouw (2)

Wat is uw nationaliteit?

◯ Nederlands (1)

◯ Anders, namelijk (2) _____

Bent u student of werkt u bij de politie?

◯ Student (1)

◯ Politie (2)

Skip To: Q13 If Bent u student of werkt u bij de politie? = Student

Hoe lang werkt u al bij de politie (in jaren)?

_____

Wat is uw positie?

_____

Heeft u een van de volgende trainingen gevolgd?

☐ Spottersopleiding (1)

☐ Search, Detect, React (SDR) (2)

☐ Geen van beide (3)

Skip To: End of Block If Heeft u een van de volgende trainingen gevolgd? = Spottersopleiding

Skip To: End of Block If Heeft u een van de volgende trainingen gevolgd? = Search, Detect, React (SDR)

Skip To: End of Block If Heeft u een van de volgende trainingen gevolgd? = Geen van beide

Als je via SONA aan deze studie deelnemt, wat is jouw SONA id-code?

_____

Welke opleiding volg je op het moment?

○ Psychologie (1)

○ Communicatiewetenschappen (2)

○ Anders, namelijk: (3) _____

Neem a.u.b een moment om over het volgende scenario na te denken:
Stel u bent een agent bij de luchthaven-beveiliging. Uw taak is om bedreigingen voor de veiligheid vroegtijdig te herkennen en te voorkomen. U staat naast de veiligheid-scan en personen die opvallend gedragen worden door u opgepakt en naar een aparte kamer gebracht voor een ondervraging.

Op wat voor soort gedragingen zou u letten bij de gecontroleerde personen? Hoe zou volgens u bijvoorbeeld een terorrist die een aanslag wil gaan plegen zich gedragen? Noem a.u.b. ten minste drie concrete non-verbale gedragingen.

U gaat nu nog enkele video's bekijken.

Wanneer u iemand schuldig bevindt, let dan goed op **wat voor gedragingen ervoor hebben gezorgd dat u een persoon als schuldig heeft beoordeeld.** Door wat voor soort gedrag zonderde de schuldige mensen voor u van de onschuldige mensen af?

Kijk alstublieft alle video's en ga pas naar de volgende pagina als u ze allemaal gezien heeft.

Appendix B: Questionnaire Part 2

U heeft net een aantal video's gezien. Denk a.u.b even erover na **wat voor gedragingen ervoor hebben gezorgd dat u een persoon als schuldig heeft beoordeeld.** Door wat voor soort gedrag zonderde de schuldige mensen voor u van de onschuldige mensen af?

Noem a.u.b. ten minste drie concrete non-verbale gedragingen en leg deze a.u.b. zo specifiek mogelijk uit.

_____

De volgende vragen gaan over de mate van concentratie tijdens deelname aan dit onderzoek.

	Helemaal niet mee eens (1)	Niet mee eens (2)	Neutraal (3)	Mee eens (4)	Helemaal mee eens (5)
Ik heb serieus deelgenomen aan dit onderzoek. (1)	○	○	○	○	○
Tijdens het onderzoek kon ik mij goed concentreren. (2)	○	○	○	○	○
Tijdens het onderzoek werd ik niet afgeleid. (3)	○	○	○	○	○

Hartelijk bedankt voor deelname aan dit onderzoek. U bent aan het einde gekomen, eventuele vragen en/of opmerkingen kunt u hieronder noteren. Voor vragen of resultaten van deze studie kunt u mailen naar a.l.fokkens@student.utwente.nl

Ga naar de volgende pagina om het onderzoek af te sluiten.

_____

_____

_____

_____

_____