# Bachelor Thesis

## Truth or lie?

The difference between three types of judges in deception detection

Janet Kraft

Bachelor Thesis, June 17, 2013

Faculty Behavioral Science, University of Twente

Bachelor Psychology of Conflict, Risk & Safety


Graduate Commission:

First supervisor: Prof. dr. Ellen Giebels

Second supervisor: Dr. Elze Ufkes

**Summary**

The main goal of this study was to look at differences in deception detection during interviews between three different types of judges: the interviewing parties, untrained observers and trained observers. The recent meta-analysis of Hartwig and Bond (2011) was used as a starting point. Our expectation was that the interviewers would be worse in deception detection than the observers. Also was studied, if the trained observers are better at deception detection than the untrained observers and interviewers. Further, and as Hartwig and Bond (2011) analysis suggests that interventions could enhance the relatively weak signals, we examined the effect of a relational intervention: mimicry. We expected that when mimicry is repressed, the accuracy in deception detection increases.

We tested our expectations on a previously collected dataset with 63 videotapes of interviews in which students either lied or told the truth. Our design was a  3 (mimicry versus no mimicry versus control) X 3 (interviewer versus untrained observer versus trained observer ) design.

The results show that the trained observers did worst in deception detection than the interviewees and the untrained observers. Also, the interviewers who mimicked the interviewee, were better in deception detection than the interviewers who repressed mimicry.

**Introduction**

Almost every day we are confronted with news about people claiming to be innocent. Just think of Bill Clinton in 1998, where he claimed that he did not have an affair with Monica Lewinsky. Another example is Joran van der Sloot in 2005 who first denied killing Natalee Holloway. As you can see, both claimed to be innocent and most people believed it in the first place. But afterwards, it turned out they both lied. Thus, deception detection might not be as simple as many people think.

Hartwig and Bond (2011) did a meta-analysis looking into why people fail to detect lies. This article will be used as a basis for the current study. Hartwig and Bond (2011) suggest there are two reasons why people are not good in recognizing lies. For one, people are looking at invalid cues when they want to judge deception. The second possibility they explore us that the signals of deception simply are too weak. Their analyses clearly provide support for the latter: people tend to look at the right indicators but the signals are simply too weak.

The current study elaborates on this notion in two ways. First, we examine whether it makes a difference who is judging deception. Because interviews are one of the most common methods to discover the truth, we used this method to analyze in which way there is a difference in deception detection (recognition of lies), between interviewers and untrained and trained observers. Hartwig and Bond (2011) further suggest interventions to increase the strength of cues of deception. Until now, most interventions focused on cognitive interventions (such as telling a story in reversed order, Vrij & Mann, 2004). In this study we will focus on a relational intervention, namely mimicry, to see if there is an improvement in deception detection.

In the following section we will first explain what deception detection exactly is and then further explore the effects of mimicry.

*Deception*

Masip, Garrido and Herrero (2004, p.148) describe deception as follows: deception is *the deliberate attempt, whether successful or not, to conceal, fabricate, and/or manipulate in any other way factual and/or emotional information, by verbal and/or nonverbal means, in order to create or maintain in another or in others a belief that the communicator himself or herself considers false.* In the literature, deception is also called lying or misleading. Most people lie to protect themselves, to avoid conflict or tension, or not  to hurt someone else (Vrij,

Granhag, & Porter, 2010). Individuals usually have some expectations about the behavior of liars: they expect them to behave nervous and show discomfort, to avoid eye contact, change their posture, speak with a higher pitched voice and will make speech errors (Hartwig & Bond, 2011). Furthermore, liars are expected to speak faster and louder and make more pauses too (DePaulo, Malone, Lindsay, Muhlenbruck, Charlton, & Cooper, 2003). In reality, almost none of people's expectations are true indicators of deception. However, Hartwig and Bond (2011) analysis revealed that despite these expectations, in reality, people look at the right indicators but these are just very weak.

That people are not good at detection liars is unfortunate particularly when high stakes are involved, such as in a police interview. People usually think that police officers (as professionals) should be better in deception detection than other people (laypersons), who are not so much involved in deception detection. But the truth is, like laypersons, the chance that a police officer catches a liar is 50 percent (Vrij et al., 2010). One possibility, why the police are poor at detecting lies is that they are involved during the interview. First, listening and interpreting the incoming information is mentally very demanding (Vrij, 2008). Secondly, a relationship may develop between suspect and interviewer, which may make deception detection a more difficult task (cf. Stel et al., 2009). The question is therefore whether an outside observer (e.g. behind a one-sided mirror) might produce higher detection rates because these cognitive and affective processes do not intervene with their task.

Based on this, the first hypothesis in our study will be:

H1: The interviewers will be worse than the observers in deception detection.

Like we discussed above, people are poor at detecting lies. Training can be a possibility to improve the skills of deception detection (Jensen, Meservy, Burgoon, & Nunamaker, 2010), particularly when one looks at global impressions rather than at single cues (Schmid Mast, Bangerter, Bulliard, & Aerni, 2011; see also Hartwig & Bond, 2011). In our study, we had the untrained observers, who looked at specific details from the interviewees, and the trained observers, who looked at global impressions from the interviewees. Based on that, the next hypothesis is:

H2: The trained observers will perform better at deception detection than the untrained observers.

As mentioned previously, Hartwig and Bond (2011) point at the importance of administering behavioral interventions to increase the strength of deception cues. Until now, most researches about deception detection are basically cognitive interventions (Hartwig, Granhag, & Vrij, 2005). We will take a closer look at relational interventions by using mimicry.

*Mimicry*

Mimicry is the *automatic tendency to imitate the behaviors of other people* (Stel, van Dijk, & Olivier, 2009, p.693). Behavior includes the facial expressions, speech and posture of people, with whom we interact (Bourgeois & Hess, 2007). For social interaction, mimicry is beneficial because if we imitate someone else, we feel closer to the other one and it enhances the liking for each other (Stel, Rispens, Leliveld, & Lokhorst, 2011). Mimicry also includes emotions, like guilt or fear. Stel et al. (2009) are one of the first researchers who examined mimicry in combination with deception detection. They examined if deception detection increases if mimicry is applied. Their results suggest, that deception detection is higher if no mimicry is used. One possible reason for that is, that listening and interpreting incoming information is mentally demanding (Vrij, 2008) and the objective judgement is obstructed by the emotional attachment (Stel et al., 2009). Based on that, the hypothesis will be:

H3: Accuracy in deception detection increases when the interviewers represses mimicry rather than uses mimicry.

**Method**

The dataset with 63 interviews previously collected by Van Leusen (2012) was used to examine the hypotheses. These interviews included people who told the truth and lied. In this study, also the interviewers and untrained observers gave judgements about the veracity of the statement of the interviewee. Additionally, and for the current study, three second year students of the University of Twente watched each video. Their answers are also included in this study.

*Participants and design*

A total of 63 interviews with 126 students took place. This group consisted of 80 men and 46 women. This included 27 male teams (a man was interviewed by a man), 10 female teams (a woman was interviewed by a woman) and 26 mixed teams (8 men were interviewed by a woman and 18 women were interviewed by a man). The mean age was 21.5 years and 32 branches of studies were followed.

All participated students were from the University of Twente and were contacted by email, flyers and personally if they wanted to participate in the experiment. Some of these students participated because they are obliged to participate some hours in study experiments. For that, they get so called "credits". Furthermore, for their participation, students got some financial compensation of 6 euro or the credits. Moreover they were told, if they perform very well, they will get more chance of getting a VVV cheque at the value of 40 euro.

The study design included a repeated measure. In this study, we have a within subject design of 3 (judge: interviewers vs. trained observers versus untrained observers) and a between dyad measure (intervention: mimicry versus no mimicry versus control).

*Procedure*

At the beginning of the experiment, the participating students were picked up by one of the two experimenters (the untrained observers). Then the students were assigned at the role of the interviewee or the interviewer randomly. The experimenter brought first the student who got the role of the interviewee in an experiment room. The interviewees were told that the experimenter will leave the room and get the instructions. The experimenter left for 2 minutes. In the meantime, the student with the role of the interviewee got the chance to donate money to Giro 555 in a donation pot. On the donation pot was written: "Honger in de

Hoorn van Africa, Giro 555". The interviewee got 5 minutes to read the instructions of telling the truth about donating money or lie about it (see Materials), after the experimenter was back. After that they handed in the instructions, the experimenter took the donation pot and counted the money. In the meantime, the student with the role of the interviewer was brought to another experiment room, where he/she also got 5 minutes to read the instructions. There were three different conditions for the interviewer: the interviewer got the instruction to mimic, the interviewer got the instruction to repress mimic or the interviewer got no instructions about mimicry (see Materials).

Before the interview started, every participant had to sign a consent form in which was said, that the participation was voluntary and that the interview would be recorded audio visually for further research. To be sure that they understood the instructions, the experimenter asked a control question.

The interview lasted for 3 minutes. The interviewee and the interviewer were brought to another room and sat in front of each other. Between them were two tables. At the edge of each table was a camera to film both of them. The camera for the interviewee was on the edge of the table of the interviewer and the camera for the interviewer was on the edge of the table of the interviewee.

During the interview, one of the experimenters was behind a one-sided mirror and acted as an observer. The observer could see the interviewer and the interviewee but they could not see the observer. None of the participants knew about that. Like the interviewer, the observer had to decide if the student told the truth or lied. The trained observers watched the videos and also had to make the same decision as the untrained observer and the interviewer. None of the observers knew which instructions the student and the interviewer got. The untrained observer, thus one of the experimenters, monitored the time. After the interview, the interviewee, the interviewer and the observers had to fill in a questionnaire (see Materials). The untrained observers, the interviewers and the interviewees got questions about their estimate of deception, their estimate about emotion of the interviewee and their estimate about the contact between interviewer and interviewee. The trained observers were also asked about their estimate of deception detection and about their global impression of the student. The participants got the money afterwards. Furthermore, they were informed by email who won the VVV cheque.

*Materials*

<u>Instruction deception</u>

The content for the instructions for the interviewees was that they will be questioned by a fellow student. They got one of the following tasks: to tell the truth or to lie about donating money. If the student donated money and got the instruction to lie about it, than the student has to tell the interviewer that he/she did not donate money. In both instructions it was important to convince the interviewer. It is important that the interviewer must believe the story of the interviewee, in the lie condition. According to these results, the

<u>Instruction interviewers</u>

The content of the instructions of the interviewers was that they will ask the fellow student about donating money. Their task was to find out if the interviewee was telling the truth or telling a lie. There were three conditions: mimicry, no mimicry and control group. In the "mimicry" condition, the interviewer was told to mimic the interviewee. They got hints to look at the face of the interviewee and pay special attention to eyes, eyebrows and mouth. Also, they should look at the body of the interviewee and pay attention to arm, leg and shoulder movements. It was important to mimic the facial and the body movements during the whole interview. In the "no mimicry" condition, the students were told that they are not allowed to imitate the movements and facial expressions of the interviewee. They got as tip for repressing mimicry, that they could press their jaws on each other, to keep the face almost static and to sit straight in the chair with the hands on the knees. The control group was given no instruction at all except for the assignment to find out if the student was telling the truth or a lie. All were promised 1 extra Euro, if they got the correct answer..

*Observers*

<u>Untrained observers</u>

The untrained observers followed a training of five hours to minimize the judgement differences between them. In the training, the observers reviewed the questionnaire, chose the position behind the one sided mirror and discussed the observation during a pilot. The goal of the training was to pay attention to the same aspects of every single interviewee and to judge these aspects in a similar way. The untrained observers were tested on their assessment of deception, where no significant difference was found. During their observation, they paid

attention to specific details, like which kind of emotions (think of fear, nervousness, anger and more) were felt by the interviewees, the interviewers and the untrained observers. Also they looked at the opinion of each party about the conversation, like how nice the other one was or feeling of trust (for more information, look at the master thesis of Van Leusen (2012).

Trained observers

The trained observers followed a training of about 30 hours. In this training, they watched many different videos and judged them separately using a scoring form. After each video they compared their answers. If there were any disagreements between the observers, they discussed it to get a collective assessment. Further, the scoring form was corrected during the training. The goal of this training was to get a good inter-rater reliability. Inter-rater reliability between the three trained observers was good; 0.717. We created an average judgement by choosing that answer that 2 out of three chose. In contrast to the untrained observers, the trained observers looked at global impressions, which will be explained below.

*Variables*

The independent variable was the mimicry condition, in which the interviewers got the instruction to mimic, to repress mimic or no instructions at all. As an independent variable we also included who made the judgement (interviewer, untrained observer, trained observer). The dependent variable was the assessment of deception (yes/no).

## Results

We used a Pearson Chi-Square to test the hypotheses, including a significance level of 0.05.

*Deception detection*

32 interviewees got the task to tell the truth about donating money and 31 interviewees got the task to lie about donating money. With use of a Pearson Chi-square, we found a significant difference between the three parties, $\chi2 = 20.378$, $p < .001$. It turned out that he untrained observers assessed 66.7 percent correctly and in 33.3 percent of the 63 interviews wrong. They had a false positive (thinking that the interviewee was lying but in reality the student was telling the truth) of 37.5 percent (12 of the 32 interviews in the condition of truth telling). Also they had a false negative (thinking that the interviewee was telling the truth but actually the student was telling a lie) of 26 percent (9 of the 31 interviews in the condition of telling a lie). The trained observers got 27 percent correct in total. These observers got a false positive of 62.5 percent (20 of 32 interviews in the truth condition) and a false negative of 83.9 percent (26 of 31 interviews in the lie condition). The interviewers did a 52.5 percent correct assessment. Their false positive was 38.7 percent (12 of 31 interviews. One opinion was missing) and the false negative was 56.7 percent (17 of 30 interviews). The difference of the total correct and wrong answer between the untrained observers and the trained observers was significant, with a Pearson Chi-Square of $\chi2 = 19.922$, $p < .001$. Between the untrained observers and the interviewers, the difference of the total answers was not significant and had a Pearson Chi-Square of $\chi2 = 2.600$, $p = .107$. A significant difference between the trained observers and the interviewers was found with a Pearson Chi-Square of $\chi2 = 8.415$, $p = .002$. Thus, the first hypothesis and second hypothesis cannot be confirmed.

**Table 1: Correct and incorrect answers**

| Condition | Untrained | | Trained | | Interviewer | |
|---|---|---|---|---|---|---|
| | Correct (%) | Wrong (%) | Correct (%) | Wrong (%) | Correct (%) | Wrong (%) |
| **Truth** | 62.5 | 37.5 | 37.5 | 62.5 | 61.3 | 38.7 |
| **Lie** | 71.0 | 29.0 | 16.1 | 83.9 | 43.3 | 56.7 |
| **Total** | 66.7 | 33.3 | 27.0 | 73.0 | 52.5 | 47.5 |

*Mimicry*

In the condition of mimicry, 22 participants got the instruction to mimic, 20 participants got the instruction not to mimic and 21 participants got no instructions. In the mimicry condition,

12 interviewees were supposed to tell the truth and 10 interviewees were supposed to tell a lie. In the no mimicry condtion, 10 of 19 interviewees had the instruction to tell the truth and 9 got the instruction to lie (one answer of the interviewer was missing in this condition). Further, in the control conditon 10 should tell the truth and 10 should tell a lie (also one missing answer of the interviewer). With use of a Pearson Chi-square, in the mimicry condition, 59.1 percent assessed correctly if the interviewee was lying or telling the truth. Also we found a false positive of 16.7 percent (2 of 12 interviews) and a false negative of 50 percent (5 of 10 interviews). 55 percent of the no mimicry condition was right about their estimation. Besides, they had a false positive of 40 percent (4 of 10 interviews) and a false negative of 66.7 percent (6 of 9 interviews). The control conidion did a correct assessment of 42.1 percent. A 40 percent of false positive (4 of 10 interviews) and a 80 percent false negative (8 of 10 interviews) were found. The outcome is that we can conclude, that the mimicry condition is better at deception detection than the others. There was a non significant difference, with a Pearson Chi-Square for the interviewers of $\chi2 = 1.256$, $p = 0.267$. The results are shown in table 2. Based on these results, the last hypothesis cannot be confirmed.

**Table 2: Answers mimicry**

| Condition | Interviewer | | | | | |
|---|---|---|---|---|---|---|
| | Mimicry | | No Mimicry | | Control | |
| | Correct (%) | Wrong (%) | Correct (%) | Wrong (%) | Correct (%) | Wrong (%) |
| **Truth** | 66.7 | 33.3 | 70.0 | 30.0 | 44.4 | 55.6 |
| **Lie** | 50.0 | 50.0 | 40.0 | 60.0 | 40.0 | 60.0 |
| **Total** | 59.1 | 40.9 | 55.0 | 45.0 | 42.1 | 57.9 |

**Discussion**

We had two main questions, which guided this study. The first one was to find out, if there was a difference in deception detection between three different types of judges: the interviewing parties, untrained observers and trained observers. Therefore, we had two hypotheses to test. Based on our results, the first hypothesis cannot be confirmed because the trained observers got less correct answers than the interviewers. Although the untrained observers got more correct answers than the interviewers, the difference was not significant. Also the following hypothesis cannot be confirmed because the trained observers did worst at deception detection than the other judges. It should be noted that the interviewers only gave about 50 percent correct answers, which is no more than guessing (Vrij, Edward, Roberts, & Bull, 2000). Besides, we took a look at false positives and false negatives. It is notable that there were more false negatives than false positives. This supports the statement that people are poor at detecting lies (Hartwig & Bond, 2011). One possibility, why the trained observers did worse than the other two parties is, that looking at global impressions may not improve deception detection as wanted. Hartwig and Bond (2011) suggest, that their use of cues of deception overlaps with actual cues of deception. This probably means, that people rely more on their intuition than on their explicit knowledge about deception. It could also mean, that the training was not good enough. Hartwig and Bond (2011) say that training does not show positive effects or just slightly improvements.

The second main question was about the effect of mimicry. According to the results, the interviewers with the instruction to mimic, did better in deception detection. But it has to be mentioned that the difference between the interviewers was not significant. Based on that, the last hypothesis cannot be confirmed

*Limitations and follow-up research*

As the results show, every expectation before turned out to be false. There are some limitations which may explain why.

All trained observers were German. It is possible that they did not understand everything correctly. Some Dutch expressions have a different meaning in German and the other way around. This can be a handicap in deception detection, because maybe the body movements and the contents of the interviews were differently interpreted. For further

research it should be kept in mind that everyone speaks the same language to prevent misunderstandings.

Furthermore, there was a difference in how the interviewees were judged. The interviewers and the untrained observers attended at the experiment. The trained observers watched the videos of the interviews afterwards. The interviewers and the untrained observers could have unconscious pressure of decision taking whereas the trained observers were able to watch the videos multiple times and take a lot of time to make their decision. By comparison of untrained observers with trained observers, both should be watching the interview at the same time to be more comparable with each other.

Another limitation of this study is that there were only two untrained observers and only four trained observers. Also only 63 interviews took place. To improve the validity of this experiment and make it more generalizable, more observers, participants and interviews are needed.

As our results suggest, the training of the observers did not have much effect. For further research more attention should be paid at the training to get better results. Apparently are 30 hours not enough to make a positive difference between the other judges. If the training is improved, it could be interesting to compare professionals with laypersons with training. As Vrij et al. (2010) stated, professionals are not better than laypersons at deception detection. Maybe with more intensive training there will be less difference.

Further should be kept in mind that we used students with no deep experience in deception. To get more valuable results, professional participants (like as an example) police officers) should be used. Also should the setting be different. This study included low stake circumstances, thus the chance of getting away with a lie was high. To get more reliable results, the stakes have to be high, like in police interviews. Also should the setting be changed, to make it more comparable research with reality. Because we used students as participants and observers, we do not know how they would react when the setting were different and more realistic.

The false positives and false negatives need more research. Our results suggest that people think more often that they see a truth when actually the other person is lying. It would be interesting for follow-up research to examine why there are more false negatives than false positives.

Conclusively we can say, that this research can be improved by different changes in setting and participants. Also should more attention be paid at the different kinds of judges and at false negatives and false positives. This study is a good basis for follow-up research.

# References

Bourgeois, P., & Hess, U. (2007). The impact of social context on mimicry. *Biological Psychology, 77,* 343-352. Doi: 10.1016/j.biopsycho.2007.11.008

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

Hartwig, M., Granhag, A., & Vrij, A. (2005). Police interrogation from a social psychology perspective. *Policing and Society, 15,* 379-399.

Hartwig, M., & Bond, C.F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgements. *Psychological Bulletin,* 137, 643-659. Doi: 10.1037/a0023589

Jensen, M., Meservy, T., Burgoon, J., & Nunamaker, J. (2010). Automatic, multimodal evaluation of human interaction. *Group Decis Negot,* 19, 367-389. Doi: 10.1007/s10726-009-9171-0

Masip, J., Garrido, E., & Herrero, C. (2004). Defining deception. A*nales de psicología, 20,* 147-171.

Schmid Mast, M., Bangerter, A., Bulliard, C., & Aerni, G. (2011). How accurate are recruiters' first impressions of applicants in employment interviews? *International Journal of Selection and Assessment,* 19, 198-208.

Stel, M., Van Dijk, E., & Olivier, E. (2009). You Want to Know the Truth? Then Don't Mimic. *Psychological science, 20,* 693-699.

Stel, M., Rispens, S, Lelived, M., & Lokhorst, A. (2011). The consequences of mimicry for prosocials and proselfs: Effects of social value orientation on the mimicry-liking link. *European Journal of Social Psychology,* 41, 269-274. Doi: 10.1002/ejsp.790

Van Leusen, E. (2012). Leugendetectie. Het effect van mimicry op deceptie detectie. *Universiteit Twente,* 1-28.

Vrij, A., Edward, K., Roberts, K., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior,* 24, 239-263.

Vrij, A., & Mann, S. (2004). Detecting deception: the benefit of looking at a combination of behavioral, auditory and speech content related cues in a systematic manner. *Group Decision and Negotiation,* 13, 61-79.

Vrij, A. (2008). Nonverbal dominance versus verbal accuracy in lie detection. *Criminal Justice and Behavior,* 35, 1323-1336. Doi: 10.1177/0093854808321530

Vrij, A., Granhag, A., & Porter, S. (2010). Pitfalls and opportunities in Nonverbal and verbal lie detection. *Psychological Science in the Public Interest,* 11, 89-121. Doi: 10.1177/1529100610390861