

**The Effect of Guilt Presumptions and Evidence Strength within Police Interviews:**  
Presuming a Suspect's Guilt and the Strength of Evidence and its Influence on Interviewer's  
Ultimate Judgments

Luca Marie Hülscher

s2305119

July 1<sup>st</sup>, 2022

**UNIVERSITY OF TWENTE.**

Psychology of Conflict, Risk, and Safety

Bachelor Thesis

First supervisor: Dr. Steven Watson

Second supervisor: Dr. Iris Sintermaartensdijk

### **Abstract**

This study investigated the impact of guilt presumptions and evidence strength within investigative interviews. Research showed interviewers presuming a suspect's guilt exhibited a confirmation bias (Hill et al., 2008) resulting in them misinterpreting the suspects' stories, leading to incorrect guilt judgments. Therefore, it was hypothesized guilt presumptions prior to the interview result in higher guilt judgments after the interview even when the suspect is innocent. This effect was expected to be stronger when strong evidence is presented compared to weak evidence. Moreover, this study examined the effect of guilt presumptions and evidence strength on trust and rapport. The study involved a 2x2 between-subjects factorial design using Evidence Strength (Weak and Strong) and Guilt Presumption (Not Presuming Guilt versus Presuming Guilt) as independent variables. The dependent variables were Guilt Judgment, Trust and Rapport. Participants played the role of a police officer. After receiving information about the case, participants interviewed the suspect by following an interview script. Around one-third of participants changed their mind from judging the suspects to be guilty to judging them to be innocent after listening to their story in an interview. Participants were less confident in judging the suspect to be guilty after the interview compared to before. Strong evidence resulted in higher guilt judgments before and after the interview. Interviewer's Trust and rapport ratings were generally lower than the interviewee's implying interviewers perceive the basis of the interaction differently than suspects. One practical key implication of this study is to shed light on the high rate of guilt assumptions and how hard individuals change their mind about the suspect's guilt especially when evidence initially seems to be strong.

**Keywords:** investigative interviews, guilt presumptions, evidence strength, guilt judgments, confidence, trust, rapport.

## **Introduction**

### **Investigative Interviews**

Investigative interviews aim to elicit an accurate, complete and detailed account from the suspect (Milne & Powell, 2010). Investigative interviewing is important as it can heavily impact the juridical decisions made and is an essential tool for information collection (Schollum, 2005). Thus, it is crucial interviews are conducted well and fairly so that investigations correctly identify the perpetrators of crime. For this, a new framework for interviewing suspects termed PEACE was introduced in England and Wales in the 1990s (Williamson, 1993). It was designed to decrease the number of false confessions evoked by aggressive interviewing tactics (for a fuller discussion of the PEACE model, see Shepherd and Griffiths, 2013). PEACE shifted the goal of police interviewing from eliciting confessions to collecting evidence (Newburn et al., 2012; Williamson, 1993).

One key principle of PEACE is police officers must act fairly when questioning suspects. This means the investigator should take an objective stance and avoid prejudice when evaluating evidence (Shepherd and Griffiths, 2013). However, interviewers often exhibit biased beliefs and judgments toward the suspect inclining a subjective stance (Hill et al., 2008; Kassin et al., 2003). This biases the interview itself as well as the interviewer's decisions made after it (Gudjonsson, 2003). For example, it results in inaccurate information presented in the interview provoking the suspect to be wrongfully convicted (Gudjonsson, 2003). Even if the suspect is not ultimately convicted, it could still significantly affect the suspect's life by temporarily imposing the status of being a suspect or even a "guilty criminal" which is accompanied by several negative consequences and significant disruption to life. Thus, it is important to avoid gathering inaccurate information as well as misinterpreting the evidence by being aware of the interviewers' biases within investigative interviewing.

The current study investigates the effect of guilt presumptions and evidence strength on the interviewers' judgments about the suspects' guilt. The study examines the impact of guilt presumptions and evidence strength on the interviewers' and suspects' perceived trust and rapport. Furthermore, the study investigates whether there is a stronger effect of guilt presumptions on guilt judgments, trust and rapport if the interviewer is presented with strong compared to weak evidence.

### **Guilt Presumptions**

When presuming the guilt of interviewees prior to the investigative interview, the police officers' aim is more likely to move away from gathering information to obtaining a confession (Gudjonsson, 2007; Kassin et al., 2003). Kassin et al. (2003) conducted a study manipulating interviewers into thinking the suspect is either guilty or innocent before the interview. Presuming guilt caused a behavioural change of the interviewers. They used more guilt-presumptive questions, utilized more coercive interrogation techniques, and exerted more pressure to get a confession (Kassin et al., 2003). Presumptions of guilt elicit a more accusatory interview style and cause the interviewer to put more effort into obtaining a confession (Adams-Quackenbush et al., 2019; Kassin et al., 2003). Both the interviewer and a neutral observer judged the suspect to be more guilty showing the investigative interview does not only influence the interviewers' judgments but also results in higher guilt judgments from other individuals (Kassin et al., 2003). Even if the interviewee is innocent, preconceived judgments of a suspect's guilt could result in the investigator misinterpreting the suspect's story and declaring the suspect guilty (Kassin et al., 2003). While the police not ultimately decide whether the suspect is sentenced or not, they do present a biased collection of evidence to the prosecutors which in turn results in a trial. If police declare the suspect guilty they restrict the suspect's life while continuing the investigation and heighten the risk of a false confession. Hence, what happens in the interview room has an impact in court and is likely

to bias other actors in the legal process, and thus, impacts the suspect's sentence even if the police do not ultimately decide about the suspect's guilt (Wanner, 2020).

Hill et al. (2008) indicate a presence of a confirmation bias within investigative interviews. Confirmation bias can be described as “the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectation, or a hypothesis at hand” (Nickerson, 1998, p.175). Confirmation bias operates based on two mechanisms, namely selective information search and biased interpretation of available information (Ask & Granhag, 2005). Thus, if police officers expect suspects to be guilty (although they are innocent) officers tend to seek information confirming their expectations and avoid information disconfirming their beliefs. Consequently, evidence interpretable in multiple ways might be interpreted in a way making the suspect look guilty, although the suspect is innocent.

Next to the guilt judgment itself, it might be important how confident the interviewers are in their guilt judgments as this may influence the suspect's legal process. Simulated jurors being confident in their guilt judgments resulted in more severe verdicts (Sue et al., 1973). Higher guilt presumptions might lead to higher confidence in the judgments about the innocent suspect's guilt. Next to this, interviewers presuming guilt might confirm their expectations by misinterpreting the evidence against the innocent suspect as incriminating.

In most studies about this topic, the interviewers choose their own questions (Hill et al., 2008; Kassin et al., 2003). However, when using the methodology of interviewers generating their own questions, it is not possible to isolate the effects of guilt presumptions on the guilt judgments from the differences in questioning (Hill et al., 2008). More specifically, it cannot be determined whether differences in interviewers' judgments are due to the aspect of the confirmation bias regarding the interviewer's selective information search, and thus, the difference in questions, or are due to the biased interpretation of available information. The current study solves this

methodological obstacle by utilising a script keeping the questions and responses the same allowing to isolate the effect of how identical information is processed. Therefore, the current study can determine whether guilt presumptions bias the processing of otherwise identical information.

### **Evidence strength**

Evidence strength, like presumptions of guilt, affects the interviewer's behaviour (Moston et al., 1992; Stephenson & Moston, 1993). The study of Sue et al. (1973) showed this by outlining stronger evidence resulted in more guilty court judgments, as well as higher confidence about the suspect's verdict. The strength of evidence influenced the interviewer's decision to make an allegation during questioning (Moston et al., 1992). Police officers utilised a confession-based approach when evidence was strong, and an information-gathering approach when evidence was weak (Stephenson & Moston, 1993). A possible explanation for this could be police officers presented with weak evidence might exhibit a weaker confirmation bias than officers presented with strong evidence. Thus, they follow the goal of eliciting an accurate, complete and detailed account from the suspect by gathering information instead of eliciting a confession. Therefore, evidence strength, like presumptions of guilt, can affect the interviewer's behaviour. However, similar to guilt presumptions, it is unknown whether there is an effect of evidence strength if interviewers' questioning is controlled for. However, it is expected confirmation bias again plays a key role. Interviewers only focus on information supporting their hypotheses about the suspect's guilt (O'Brien & Ellsworth, 2006). Thus, even if the suspect is innocent, the interviewer could misinterpret the suspect's explanation of the evidence (O'Brien & Ellsworth, 2006). Evidence strength is likely to reinforce the strength of initial presumptions which could make it especially difficult for people to change their minds with strong evidence compared to weak evidence. Thus, if there exists strong evidence, the baseline guilt presumptions may be higher and guilt scores after the interview may be more similar to the baseline guilt judgment with strong compared to weak

evidence. Thus, guilt presumptions might be more likely to provoke incorrect guilt judgment when evidence is strong.

### **Relationship between Interviewer and Interviewee**

The relationship between the interviewer and the interviewee has been declared to be one of the most important factors for a successful interview (Williamson, 2013). If the quality of the relationship between the investigator and the suspect is neglected, even the best interview techniques might fail and the aim of the investigative interview cannot be reached.

### ***Rapport***

One element of the relationship between interviewer and interviewee is rapport which fosters cooperation between the interviewer and the interviewee (Gabbert et al., 2021). There is consensus on rapport being crucial for a successful interview (Collins & Carthy, 2018; Gabbert et al., 2021). Rapport can be defined as “smooth, positive interpersonal interaction” (Abbe & Brandon, 2014, p. 208) and increases the quantity of information displayed by the interviewee (Collins et al., 2002). Moreover, investigators gathered more forensically relevant and accurate information when they built rapport (Collins & Carthy, 2018).

To establish rapport, the interviewer must take a non-judgmental approach in interviewing (Alison et al., 2013). However, if suspects (whether innocent or guilty) resist, the interviewers discard their initial efforts to build and maintain rapport, and increasingly offer confrontation (Izotovas et al., 2021). Therefore, if interviewers presume guilt, but the suspects do not acknowledge this guilt (e.g., because they are innocent), the interviewer might perceive the suspect as offering resistance and might neglect rapport. Thus, guilt assumptions affect rapport-building behaviours. However, it is to be determined whether guilt presumptions themselves are enough to affect rapport even when the interviewer’s behaviour is fixed by interview scripts, or whether additional cues, like the strength of evidence, must be taken into account. When there is strong

evidence, the interviewers might be biased to misinterpret social cues indicating rapport. A possible explanation for this is interviewers form a negative first impression based on the strong evidence. Due to confirmation and belief perseverance biases, the interviewers maintain these negative impressions, leading them to misinterpret social cues, and thus, resulting in the low rapport ratings although the suspect had a reasonable explanation for the evidence against them (Grahe & Bernieri, 1999).

### *Trust*

Next to rapport, the concept of trust is important as several studies stressed the vital role of trust in the quality and quantity of information gathering in police interviews (Roberts, 2010; Wilson & Walsh, 2019). Trust can be defined as “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party” (Mayer et al., 1995, p.712). This willingness to trust depends on the perceived trustworthiness of the other persons. This trustworthiness of the others is determined by their traits and actions (Mayer et al., 1995). Moreover, to undertake a trusting action individuals must take risks (Mayer et al., 1996). If individuals engage in this risky behaviour or not depends on the perceived risk of the situation. Thus, the higher individuals perceive the risk within an investigative interview, the lower their trust might be. If police interviewers presume guilt and face strong evidence, they might perceive the highest risk resulting in the lowest trust ratings.

According to Mayer et al., (1995), trust has three constituent parts, namely (1) ability, (2) integrity and (3) benevolence. Ability is defined by the characteristics, skills and competencies ensuring an individual can be trusted in a specific area, such as conducting an investigative interview. Integrity includes the perception of whether the trustee utilises a specific set of principles and rules which are considered appropriate by the trustor. The trustor expects the trustee to be



honest and to have no hidden agenda in the interview. However, if the police officer presumes guilt and if there is strong evidence, the police officer might perceive a lower degree of integrity, and thus, trust. Not only the police officer, but also the suspect might perceive a lower degree of integrity, and thus, trust, because the police officers might change their behaviour toward the suspects when they perceive low trust.

Benevolence is the extent to which a trustee is believed to want to do good to the trustor, aside from an egoistical profit motive. If interviewers presume guilt, they might interpret the suspects' explanation of evidence as uncooperative. The interviewer might perceive the suspect as less benevolent, thus, less trustworthy. The lower trust might lead to the interviewer exhibiting a more accusatory interviewing approach. This could result in the interviewee perceiving the suspect as less benevolent, thus, less trustworthy.

To conclude, there might be a difference in interviewers' and interviewees' overall trust ratings as the three parts of trust might be affected differently for the two parties. Therefore, the study measures the concept of trust individually for the interviewee and interviewer.

### **Purpose of this study**

Investigative interviews are crucial for the police to gather information. Conducting the interviews in an objective manner and obtaining reliable and accurate information from the suspect is of utmost importance. However, interviewers can be misguided by biases and preformed beliefs towards the suspect. Two factors possibly triggering the biases are guilt presumptions and evidence strength. Therefore, this study focuses on the interviewer's presumptions of guilt and the evidence strength and their effect on the interviewer's guilt judgment about the suspect. Moreover, the study examines the effect of guilt presumptions and evidence strength on trust and rapport in the interviewer's and suspect's interaction. The mock suspects in this study will in fact be innocent and will provide an account giving a plausible alternative explanation for the evidence against

them. To control for different interview techniques, this study keeps the interviewers' questions and the interviewees' responses the same by providing a pre-formed script for the interaction. This enables investigating how interviewers' biases affect the interpretation of the suspects' narrative.

From this, the research question follows. What is the effect of guilt presumptions and evidence strength on the interviewer's perception of the suspect in the context of an investigative interview?

### **Hypotheses**

**H1:** Interviewers presuming guilt prior to the interview will be more likely to assume innocent suspects are guilty after the interview.

**H2:** Evidence against innocent suspects appearing strong compared to weak will make interviewers more likely to assume guilt after the interview.

**H3:** When interviewers presume guilt, the perceived rapport and the perceived trust will be reduced in both the suspects and the interviewers after the interview.

**H4:** When there is strong evidence and additionally interviewers presume guilt, the guilt judgments after the interview will be the highest of all conditions.

**H5:** When there is strong evidence and additionally interviewers presume guilt, the perceived rapport and the perceived trust after the interview will be the lowest of all conditions in both interviewers and suspects.

### **Methods**

#### **Design and Manipulation**

This study consisted of a 2x2 between-subjects factorial design using Evidence Strength (Weak and Strong) and Guilt Presumption (Not Presuming Guilt versus Presuming Guilt) as independent variables. The dependent variables were Guilt Judgment, Trust and Rapport.

## **Participants**

The participants were recruited via convenience sampling through an institutional test-subject system (SONA). Through participation, they earned course credit. Other participants were personally recruited through the researchers' social media channels (i.e., Facebook, Instagram). Some of the participants were acquaintances of one of the researchers. Participants who signed up outside of the participant pool were given no incentive. The prerequisite for participation was individuals were at least 18 years old and could speak and read English well.

In total, 65 participants took part in this study. The participants' mean age was 22.81 years (SD= 5.82, Range = 19-55). The majority of participants were German (N=56). Most of the participants were either students from the University of Twente (N=24) or from another university (N=13), compared to not being a student (N=23). The participants were assigned in order of entry to one of the four experimental conditions: (A) Weak Evidence - No Guilt Presumptions (N=17); (B) Weak Evidence - Guilt Presumptions (N=15); (C) Strong Evidence - No Guilt Presumptions (N=20); (D) Strong Evidence - Guilt Presumptions (N=13). Some last-minute dropouts of participants resulted in an unequal number of participants in each of the conditions. The participants were guaranteed anonymity. The BMS Ethics Committee authorized the present research (Approval code: 220428).

## **Materials and Procedure**

### ***Information sheet***

After the participants signed up for the study via SONA or via the researchers, the researchers sent them an information sheet with instructions via email (Appendix A). This was done at least 24 hours before the scheduled interview to ensure participants had enough time to prepare.

The information sheet instructed participants to prepare for the interview by reading a vignette and the questions they asked the suspect which were also sent to the participants via email. Moreover, they were informed they must (1) fill out some questions before the interview, (2) conduct the interview, and (3) fill out the rest of the questionnaire.

### *Vignette*

Next to the information sheet, the researchers sent the participants background information about the case and evidence against the suspect as well as the questions the participants will ask the suspect during the interview (Appendix B).

The vignette instructed participants to put themselves in a police officer's position and informed participants their supervisor has asked them to interview the suspect involved in a new case. The participants were informed there was a woman arrested for dealing drugs, including the exact place, date, and time of this crime. The suspect was alleged to be her accomplice and was therefore also suspected of dealing drugs. Moreover, the vignette included the evidence against the suspect. It also repeated the participant's task to read the questions to the suspect. Moreover, the participants were informed the introduction part of the interview was already over, and they can directly start by asking the questions provided to ensure all interviews were identical.

**Weak versus Strong Evidence Conditions.** Depending on the condition, the suspect received a vignette including Weak or Strong Evidence against the suspect. There were five pieces of evidence in both conditions and the underlying key piece of evidence was the same, the pieces just differed in strength. The difference in Strength of Evidence was based on the Evidence Framing Matrix (Granhag et al., 2013). The pieces of evidence in the Strong and Weak Evidence Conditions were almost identical to each other, only differing in the dimension of the strength of source and specificity of evidence (Granhag et al., 2013) (Appendix B). Strength of source related to how strong or weak individuals perceive the evidence to be. Specificity concerned the perceived degree

of precision of evidence varying from low to high. Evidence with a high degree of precision is exact while evidence with a low degree of precision is rather vague.

In the Weak Evidence Condition, the evidence was stemming from an unreliable and weak source. An example of Weak Evidence is a witness statement from an old woman suffering from schizophrenia, as evidence stemming from individuals with a mental condition is often perceived as less reliable and credible (Gatenby, 2020; Reavey et al., 2016). Moreover, Weak Evidence has little specificity, such as a phone call on the suspect's phone from the woman convicted for dealing drugs on the day of the drug deal.

In the Strong Evidence Condition, the pieces of evidence were from a reliable and strong source such as a surveillance camera showing a clear image. Evidence shows individuals often overestimate the credibility and reliability of surveillance camera footage resulting in them perceiving camera footage as Strong Evidence (Granot et al., 2018). Furthermore, in this condition, the evidence is quite specific, thus, having a high degree of precision, such as a specific message including the time and place of the drug deal on the suspect's phone coming from the woman convicted for dealing drugs.

At the end of the vignette, the participants were provided with the questions to ask the suspect during the interview (Appendix C). The questions in the different conditions were identical other than the specific evidence presented to the suspect (Weak or Strong).

**Table 1***Examples of Pieces of Evidence*

| Weak Evidence Condition   | Strong Evidence Condition  |
|---|--|
| An old woman suffering from schizophrenia saw someone that looked like the suspect together with Mrs Brown in the park, five minutes before and after Mrs Brown dealt the drugs | There is surveillance camera footage that the suspect was together with Mrs Brown in the park, five minutes before and after Mrs Brown dealt drugs |
| There were traces of marijuana found in the car of the suspect  | There were traces of the same drugs (Heroin, morphine, LSD, Marijuana) in the car from exactly the same batch as Mrs Brown dealt                   |

*Note.* Based on the Evidence Framing Matrix (Granhag et al., 2013).

**Guilt versus No Guilt Presumption conditions.** The second manipulation in the study was whether the participants were primed to believe the suspect is guilty or not. This was achieved by instructing the participants their “supervisor” is either (1) very sure (Guilt Presumption) or (2) not sure (No Guilt Presumption) if the suspect is guilty of drug dealing.

***Pre-interview***

There were two researchers, one taking the role of the experimenter and one of the suspect. To prevent biased results, all participants were interviewed by a researcher who was unfamiliar with them. When the participants joined the scheduled online meeting via the platform ZOOM, the experimenter awaited them with a turned-on camera and microphone. The experimenter ran again verbally through the information sheet with the participant and asked if the participant had any questions. The suspect has not joined the meeting yet. The participant was provided with the link

of the online survey via the chat function and it was clarified the experimenter waited in the online meeting until the participant finished the first part of the survey in case the participant had any questions. The participants were asked to turn off their camera and microphone and fill out the questionnaire until the questionnaire instructed the participants to go back to the meeting.

### ***Pre-questionnaire***

Qualtrics was used to collect the measures. First, the participant gave informed consent. Within the questionnaire, the participants were asked to read the vignette again and to leave this document open, as they could use it to read out the questions in the interview. The participants were instructed to exactly adhere to the questions provided to ensure the content of all interviews was identical.

To check whether the Guilt manipulation worked, the participants were asked the following question “Based on the information I have about the current case, I believe the suspect is....” and could choose from the answers “guilty” or “not guilty” which represented the variable Binary Guilt Pre-Measure. Moreover, they were asked “How sure are you about the suspect's guilt?” and could answer on a 5-point Likert scale (1= very sure the suspect is innocent; 5= very sure the suspect is guilty) which represented the variable Guilt Pre-Measure.

To check whether the Strength of Evidence manipulation worked, the participants answered “I perceive the evidence against the suspect as...” on a 5-point Likert scale (1 = very weak, 5 = very strong).

Lastly, Qualtrics notified the participant the questions before the interview were finished, but it is important the participant leaves the questionnaire open. Qualtrics instructed the participant to go back to the online meeting and stressed the importance of waiting until after the interview to continue with the second part of the questionnaire.

### ***Interview***

After finishing the questionnaire and joining the online meeting again, the experimenter asked if the participant was ready to start the interview or if the participant had any questions. The experimenter turned off the camera and microphone, and the suspect joined the meeting. It was ensured the participant and suspect did not know each other to prevent biased results. Moreover, the participant thought the suspect was just another participant in the study and not a researcher to make the interaction more realistic.

As soon as the suspect joined the meeting, the participants asked the questions (Appendix C). The questions aimed at getting more information about the suspect's account and were formulated neutrally. In total, the interviewer asked nine questions. The suspect's answers were identical in all interviews, independent of condition.

This made the interviews as objective as possible for the interviewer and created the same conditions ensuring the effects on the dependent variables were isolated. The suspect always provided an account explaining away all evidence against them and was in fact innocent. The interview was about five minutes long. When they finished the interview, the suspect left the online meeting and the experimenter turned on the camera and microphone again. The experimenter asked the participant to complete the post-interview questionnaire. The experimenter stayed in the online meeting as long as the participant filled out the questionnaire in case the participant had any questions.

### ***Post-questionnaire for Participants***

**Guilt Judgment.** The first two questions after the interview were the same questions as the participant had already answered once before the interview. However, for clarity, the questions are outlined again. The first question after the interview was "Based on the information I have about the current case, I believe the suspect is...." with the answer options "guilty" or "not guilty" which



represented the variable Binary Guilt Post-Measure. The second question was “How sure are you about the suspect's guilt?” and participants answered on a 5-point Likert scale (1= very sure the suspect is innocent; 6= very sure the suspect is guilty) which represented the variable Guilt Post-Measure.

**Rapport.** The dependent variable of Rapport was measured based on the Rapport Scale for Investigative Interviews and Interrogations (Rs3I) Interviewee Version (Duke et al., 2018). This scale measured six dimensions of Rapport, namely expertise, cultural similarity, attentiveness, commitment behaviour and connected flow. Eighteen items measured Rapport which were answered on a five-point Likert scale (1= strongly disagree; 5= strongly agree). The higher participants scored on the scale, the more Rapport was present.

The original Rs3I was written from the perspective of the interviewee, therefore, for the interviewers' version, the statements were adjusted to make them appropriate to answer from the interviewer's perspective (Appendix D). “The interviewee was attentive to me” is an example of one of the items. The scale had good internal reliability ( $\alpha = .90$ ) in the sample of this study.

**Trust.** The dependent variable Trust was measured based on the scale from Mayer and Davis (1999) including (1) ability, (2) benevolence, and (3) integrity as subscales of Trust. The participants answered 17 items on a five-point Likert scale (1= strongly agree; 5= strongly disagree) to measure their Trust in the suspect. The items were adjusted so that it made sense to answer them from the perspective of the interviewer, see Appendix E for the list of items. “The interviewee seemed very capable of answering the questions” is one example of the trust items. The Cronbach's alpha for the overall trust scale in the current study was good ( $\alpha = .79$ ).

**Demographics.** The participants answered some questions about their age, nationality and occupation.

**Questions about participants' immersion.** Moreover, the participants were asked some questions reflecting their immersion and effort, for example, "The interview was somewhat realistic" (Appendix F). Furthermore, the question "Do you have anything left to say to the researcher?" was asked.

**Debriefing.** At the end of the questionnaire, a debrief was provided including information about the theoretical background of the study and the hypotheses of the study (Appendix G).

### *Post-questionnaire for Suspect*

Not only the participant filled out a questionnaire, but the researcher playing the suspect also completed a questionnaire on Qualtrics. The suspect was blind to the Guilt Condition, however, it was not possible to blind toward the Evidence Strength Condition as the evidence needed to be introduced during the interview. The questionnaire served the purpose of comparing the interviewer's ratings of Trust and Rapport with the suspect's. The suspect filled out items about Trust, based on the scale of Mayer and Davis (1999) and items about Rapport, based on the Rs3I of Duke et al. (2018).

### *Post-interview*

When the participants finished the whole questionnaire, the experimenter asked if they had any questions. If everything was clear, the experimenter thanked the participant again and ended the online meeting.

## **Data Analysis**

### *Confirmatory Analysis*

The data analysis program SPSS version 26 is used. All analyses are conducted considering an alpha level of 0.05.

To ensure the manipulation of the independent variables Guilt Presumption and Evidence Strength has worked, t-tests are used. It is checked whether participants perceived a difference in conditions (Strong versus Weak Evidence; Guilt versus No Guilt Presumptions).

To analyse the variables Guilt Judgment, and the variables Trust and Rapport, a two-way ANOVA is run. Through this, the differences in the dependent variables yielded by the different conditions of the independent variables are investigated. Moreover, it can be tested whether there is an interaction effect between Strength of Evidence and Guilt Presumption on the dependent variables. In case of significant results for the interaction effects, Tukey post-hoc tests are performed to further analyse the data.

To test whether there is a difference between Trust and Rapport ratings for the interviewer and the suspect, a Pearson's correlation and a t-test are conducted.

### *Exploratory analysis*

Moreover, exploratory analyses are conducted using a measure of immersion as a covariate. After excluding the participants who did not immerse themselves in the study, all analyses are performed again to see whether there are any differences.

## **Results**

### **Descriptive statistics**

There were two measures of Guilt, namely the binary Guilty Judgment and the Likert Guilt Judgment. First, only the binary measure of Guilt is considered to determine what proportion of participants assumed guilt before and after the interview. The majority of individuals found the suspect to be Guilty ( $N = 45, 69.2\%$ ) versus Not Guilty ( $N = 20, 30.8\%$ ) before the interview. When examining the binary variable Guilt Post-measure, the minority of participants selected Guilty ( $N = 28, 43\%$ ) versus Not Guilty ( $N = 37, 57\%$ ). There was a significant association between the binary Guilt Pre-Measure and the Post-Measure ( $\chi^2(1, N = 65) = 3.85, p = .05$ ). Around 35% ( $N =$

23) of participants changed their minds from judging the suspect to be Guilty before the interview to judging them to be Not Guilty after the interview. 7% (N = 5) of participants changed their judgments from judging the suspect to be Not Guilty to judging them to be Guilty after the interview.

After considering the binary Guilt-measure, the Likert Scale Guilt Measure is considered. Pre-Guilt scores ( $M = 3.54$ ,  $SD = 0.95$ ) were higher than Post-Guilt Scores ( $M = 2.94$ ,  $SD = 1.16$ ) meaning participants were more confident in the suspect's guilt before the interview compared to after the interview. Means of all other variables of interest are also represented in Table 2.

**Table 2**

*Mean Scores of Variables of Interest*

| Variable                             | Mean (SD)   | Minimum | Maximum |
|--------------------------------------|-------------|---------|---------|
| Guilt Pre-Measure                    | 3.54 (.95)  | 2       | 5       |
| Guilt Post-Measure                   | 2.94 (1.56) | 1       | 5       |
| Evidence Strength Pre-Measure        | 3.25 (1.19) | 1       | 5       |
| Evidence Strength Post-Measure       | 2.78 (1.26) | 1       | 5       |
| Participants' Trust in Interviewee   | 3.22 (.44)  | 2.24    | 4.59    |
| Interviewees' Trust in Participant   | 3.46 (.42)  | 2.41    | 4.24    |
| Participants' Rapport in Interviewee | 3.56 (.63)  | 2.47    | 5.00    |
| Interviewees' Rapport in Participant | 3.90 (.41)  | 2.76    | 4.65    |

*Note.* All variables can be interpreted in terms of a 5-point Likert Scale.

Next, Pearson's correlation coefficients are reported. As Table 3 shows, Guilt Pre-Measure was strongly correlated with the Guilt Post-Measure and the Evidence Strength Pre-Measure ( $.34 < r < .50$ ). This means participants being confident a suspect was guilty before the interview were more likely to perceive evidence as strong before the interview and were more likely to be confident the suspect was guilty after the interview.

Guilt Post-Measure was correlated with Evidence Strength Post-Measure, and with the participants' Trust and Rapport in the interviewee ( $-.25 < r < .54$ ). This means participants being confident a suspect was guilty after the interview were more likely to perceive evidence as strong after the interview. Moreover, participants with higher Guilt Post-scores rated the established Trust and Rapport with the interviewee as lower.

Evidence Strength Pre-Measure was strongly correlated with the Post-Measure and with the participants' Rapport in the interviewees ( $.25 < r < .50$ ). This means participants perceiving the evidence as strong before the interview were more likely to perceive it as strong after the interview and it influenced the participants' Rapport in the interviewee.

The participants' Trust ratings were strongly correlated with the participants' Rapport ratings and the same applies to the interviewees' Trust and Rapport ratings ( $.66 < r < .72$ ). This means participants and interviewees giving a low Trust rating were more likely to give a low Rapport rating (and vice versa). None of the participants' ratings were correlated with the interviewees' ratings for either Trust or Rapport.

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

**Table 3***Pearson's Correlation Between the Dependent Variables of Interest*

|                              | Guilt<br>Pre-M. | Guilt<br>Post-M. | Evidence Strength<br>Pre-M. | Evidence Strength<br>Post-M | Trust<br>participant | Trust<br>interviewee | Rapport<br>participant | Rapport<br>interviewee |
|------------------------------|-----------------|------------------|-----------------------------|-----------------------------|----------------------|----------------------|------------------------|------------------------|
| Guilt<br>Pre-M.              | -               | .34**            | .50**                       | .24                         | -.06                 | .15                  | -.05                   | .19                    |
| Guilt<br>Post-M.             | -               | -                | .23                         | .54**                       | -.25*                | -.18                 | -.25*                  | -.09                   |
| Evidence Strength<br>Pre-M.  | -               | -                | -                           | .50**                       | .03                  | .11                  | .01                    | .25*                   |
| Evidence Strength<br>Post-M. | -               | -                | -                           | -                           | -.22                 | -.05                 | -.21                   | .17                    |
| Trust<br>participant         | -               | -                | -                           | -                           | -                    | -.01                 | .66**                  | -.04                   |
| Trust<br>interviewee         | -               | -                | -                           | -                           | -                    | -                    | -.19                   | .72**                  |
| Rapport<br>participant       | -               | -                | -                           | -                           | -                    | -                    | -                      | -.18                   |
| Rapport<br>interviewee       | -               | -                | -                           | -                           | -                    | -                    | -                      | -                      |

Note. \* $p < .05$ ; \*\* $p < .01$

### **Effectiveness of Manipulations**

For the manipulation checks, it was examined whether there is a difference in the participant's Guilt Presumptions prior to the interview based on the guilt manipulation and a difference in participant's perceived Evidence Strength prior to the interview based on the evidence manipulation. The t-test revealed there was no statistical difference between individuals in the Guilt Presumption Condition ( $M = 3.57, SD = .93$ ) and in the No Guilt Presumption Condition ( $M = 3.50, SD = 1.00$ ),  $t(63) = -0.28, p = .78$ . The Evidence was perceived as stronger in the Strong Evidence Condition ( $M = 3.61, SD = 1.17$ ) compared to the Weak Evidence Condition ( $M = 2.88, SD = 1.10$ ) which was a statistically significant difference,  $t(63) = -2.59, p = .01$ .

### **Guilt Judgment and Evidence Strength**

Originally it was planned to examine the effect of Guilt Presumptions on Guilt Judgment and Evidence Strength, however, the guilt manipulation did not work and, therefore, some of the original hypotheses involving the guilt manipulation (Hyp. 1 & 4) could not be tested. Nevertheless, as this study measured participants' guilt judgments before and after the interview, there existed an alternative to examine the variable Guilt Judgment. Therefore, it was decided to run a new model still allowing to estimate the extent the interview changed participants' initial assumptions. The new model was a 2x2 within-subjects factorial design using Time (Pre- vs Post-interview) as within participant variable and Evidence Strength (Weak vs Strong) as between participant variable. By running this one-way repeated measures ANOVA, it was determined whether there was a difference in Guilt Judgment before and after the interview. However, for transparency, the original analysis was also conducted and is included in Appendix H.

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

There was a statistically significant main effect for Time,  $F(1,63) = 15.40, p < .001, partial \eta^2 = .20$ . The Guilt Judgment mean score decreased from Pre-Measure ( $M = 3.54, SD = .95$ ) to Post-Measure ( $M = 2.94, SD = 1.16$ ).

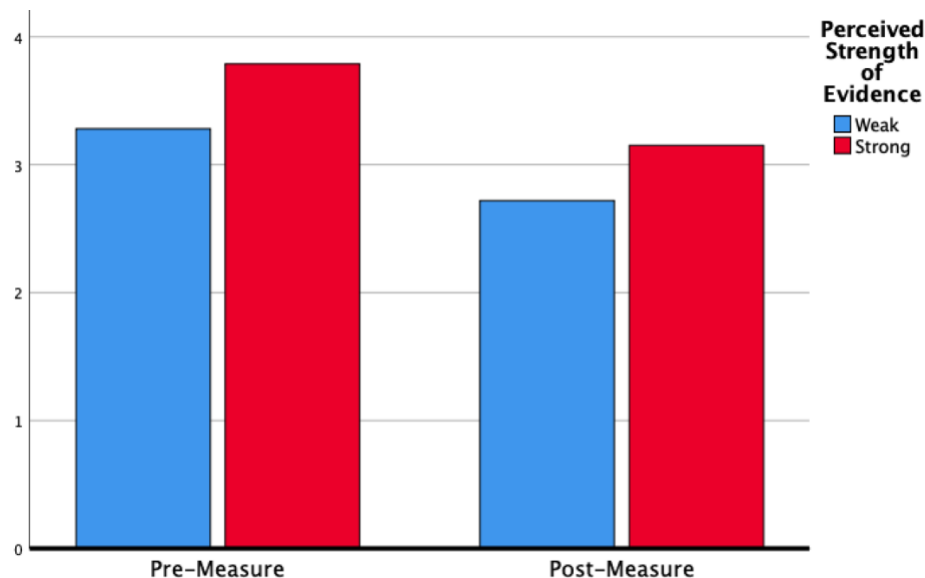
There was a significant main effect for Evidence Strength,  $F(1,63) = 5.08, p = .03, partial \eta^2 = .08$ . Guilt Judgments for individuals in the Strong Evidence condition were higher compared to the Weak Evidence condition ( $M = 3.47, SD = 0.15$  versus  $M = 3.00, SD = 0.15$ ). Thus, the results give support for hypothesis two, namely Strong Evidence makes interviewers more likely to assume Guilt after the interview compared to Weak Evidence.

There was no statistically significant interaction effect between the Guilt Measures and Evidence Strength,  $F(1,63) = 0.06, p = .81, partial \eta^2 < .001$ . Participants were most confident in judging the suspect's Guilt before the interview and when presented with Strong Evidence ( $M = 3.79, SD = 0.86$ ) compared to Weak Evidence ( $M = 3.28, SD = 0.99$ ). Moreover, compared to the Pre-Measures, Guilt Judgments decreased after the interview and when presented with Strong Evidence ( $M = 3.15, SD = 1.15$ ) and even more when presented with Weak Evidence ( $M = 2.72, SD = 1.14$ ). Figure 1 displays a bar chart visualizing these findings.



**Figure 1**

*Guilt Judgments Before and After the Interview Including the Effect of Evidence Strength*



### **Trust and Rapport Ratings**

To test for effects of Guilt Presumptions (Guilt Presumptions versus No Guilt Presumptions) and Evidence Strength (Weak versus Strong) on both the interviewers' and interviewees' Trust and Rapport, a factorial ANOVA was conducted. It should be noted the guilt manipulation did not work, however, it was decided to run the original analyses for these independent variables and test the original hypotheses (Hyp. 3 & 5). This way it was determined whether there were any effects of the manipulation post-interview even if they had not been noticeable pre-interview.

For both participants' and interviewees' Trust and Rapport ratings, there were no statistically significant main effects for Guilt Presumptions and for Evidence Strength (all  $F$ 's < 3.45; all  $p$ 's > .07) except the main effect of Evidence Strength on Participants' Rapport ratings. Here, there was marginally significant effect,  $F(1,60) = 3.45$ ,  $p = .07$ , partial  $\eta^2 = .05$ . There was

the possibility interviewers rated Rapport higher in the Weak Evidence condition compared to the Strong Evidence condition ( $M = 3.73$ ,  $SD = 0.61$  versus  $M = 3.41$ ,  $SD = 0.62$ ). Moreover, there were no significant interaction effects between Guilt Presumptions and Evidence Strength for both participants' and interviewees' Trust and Rapport ratings (all  $F$ 's  $< 0.66$ ; all  $p$ 's  $> .42$ ). The specific values of each effect and the means can be found in Appendix I.

Hypothesis three stated interviewers presuming Guilt result in reduced Rapport and Trust ratings from both participant and suspect after the interview. Since the effect of Guilt Presumptions was non-significant, hypothesis three was rejected. Hypothesis five stated when participants presume Guilt and additionally are presented with Strong Evidence, the Trust and Rapport ratings are lowest in both participant and interviewee after the interview. When comparing the means of Trust and Rapport between the conditions, Trust and Rapport ratings were indeed the lowest, however, the effect was not significant. Therefore, hypothesis five was rejected.

### ***Differences in interviewers' and interviewees' ratings***

Paired-samples t-tests were conducted to compare the Trust and Rapport ratings of the participants' and the interviewees'. The Trust ratings of the participants' were lower than the ones of the interviewees',  $t(63) = 3.02$ ,  $p = .004$ ,  $M = 3.22$ ,  $SD = 0.45$  versus  $M = 3.46$ ,  $SD = 0.43$ .

The Rapport ratings of the participants' were lower than the ones of the interviewees'  $t(63) = -3.22$ ,  $p = .002$ ,  $M = 3.57$ ,  $SD = 0.63$  versus  $M = 3.91$ ,  $SD = 0.42$ .

### **Exploratory analyses**

An independent-samples t-test was run to determine if there were differences in Immersion between the No Guilt Presumptions condition and Guilt Presumptions condition as well as between the Weak and Strong Evidence condition. This exploratory analysis was run because there existed

doubts about some participants' immersion in the study. The mean Immersion score of participants was around "Somewhat agree" on the Likert scale ( $M = 4.17$ ,  $SD = 0.69$ ) meaning most participants immersed themselves in the study.

The mean Immersion score for the No Guilt Presumption condition did not significantly differ from the Guilt Presumption condition ( $M = 4.24$ ,  $SD = 0.48$  versus  $M = 4.12$ ,  $SD = 0.82$ ,  $t(61) = 0.66$ ,  $p = .51$ ). The same goes for the Weak and Strong Evidence condition ( $M = 4.22$ ,  $SD = 0.64$  versus  $M = 4.12$ ,  $SD = 0.75$ ,  $t(61) = 0.62$ ,  $p = .54$ ).

After excluding the participants not immersed in the study (Immersion score  $< 3$ ,  $N = 3$ ) which were 4.8% of the sample, all analyses were conducted again. Participants scoring under three selected "Strongly agree" or "Somewhat agree" on at least some questions of the Immersion scale and, therefore, did not immerse themselves in the study.

All effects remained the same, thus, all effects were still non-significant except for the main effect of Time ( $F(1,58) = 12.11$ ,  $p = .001$ , *partial*  $\eta^2 = .17$ ), the main effect of Evidence Strength ( $F(1,58) = 5.26$ ,  $p = 0.03$ , , *partial*  $\eta^2 = .08$ ), and the difference between interviewers' and interviewees' Trust ratings,  $t(58) = 2.57$ ,  $p < .01$ , and Rapport ratings,  $t(58) = -2.82$ ,  $p < .01$ .

### **Summary of results**

In summary, one-third of participants changed their minds from judging the suspect to be Guilty before the interview to judging them to be Not Guilty after the interview. Participants were more confident in their Guilt Judgments before the interview than after the interview. Moreover, Guilt Judgments from participants presented with Strong Evidence were higher than those from participants presented with Weak Evidence which supported hypothesis two. Hypothesis three was rejected as interviewers presuming Guilt not resulted in reduced Rapport and Trust ratings. Hypothesis five was also rejected as interviewers presuming Guilt and presented with Strong

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

Evidence did not result in the overall lowest Trust and Rapport. When only considering Evidence-Strength, it becomes apparent Strong Evidence almost resulted in reduced participants' Rapport ratings. Furthermore, participants' and interviewees' Trust and Rapport ratings did not align as participants' generally rated the established Trust and Rapport to be lower. When excluding participants not immersed in the study, the mentioned effects stay the same.

### **Discussion**

#### **Reflection of the findings**

This study investigated the effect of guilt presumptions and strength of evidence within police interviews. More specifically, it was tested whether presuming a suspect's guilt versus not presuming it and strong versus weak evidence influenced guilt judgments, trust and rapport of the interviewer and suspect within the scope of a scripted online police interview.

#### ***Guilt Judgments***

When looking at the manipulation check for guilt, it turned out there was no difference between the two guilt conditions. Therefore, the hypotheses thematizing guilt presumptions could not be tested. However, the study entailed a binary and a Likert scale pre-and post-interview measure of guilt which could be compared. Around one-third of the participants changed their minds from judging the suspect to be guilty to not judging the suspect to be guilty after they conducted the interview. Some participants even changed their judgments from not guilty to guilty after listening to the suspect's account. Participants found the suspects to be less guilty or were less confident about their guilt judgment after the interview compared to before the interview. However, around 40% of participants still believed the suspect was guilty even after the interview where the suspects gave a clear argumentation of their innocence. These findings mostly coincide with the existing literature. Studies suggested some people show a confirmation bias, or a tunnel version

when testing their hypotheses (Hill et al., 2008; Nickerson, 1998; O'Brien & Ellsworth, 2006). When individuals expect a suspect to be guilty, they are generally more likely to seek information confirming their hypothesis, and thus, perceive the suspect as more likely to be guilty and are more confident about this (Ask & Granhag, 2005; Hill et al., 2008). Rollwage et al. (2020) showed confirmation bias is especially strong when individuals are confident in their judgments. Thus, if individuals are confident about the suspect's guilt before the interview, they are more likely to exhibit a confirmation bias during the interview. This amplifies the individuals' search for hypothesis-confirming information resulting in higher confidence judgements after the interview (Rollwage et al., 2020). Therefore, if individuals are confident in their guilt judgments before the interview, they may be less likely to change their minds about the suspect's guilt, even if the suspect is innocent.

Multiple studies indicated police officers interpret evidence in a way confirming their original thoughts about the suspects (Hill et al., 2008; Kassin et al., 2003). Police interviewers fail to neglect their original hypotheses about the suspect's guilt even when the evidence is objectively disconfirming (Ask & Granhag, 2005). Instead of interpreting ambiguous information about the suspect's innocence as exonerating, police officers interpret it in incriminating terms due to their confirmation biases (Ask & Granhag, 2005). Thus, multiple studies showed unfounded presumptions of guilt result in a confirmation bias leading to individuals being confident suspects are guilty even after the interview. However, this study showed at least some interviewers changed their minds after listening to the suspects' argumentation outlining their innocence. Directly influencing these guilt judgments is the strength of evidence presented against the suspect as shown by existing literature (Stephenson & Moston, 1993; Sue et al., 1973). However, it was not investigated so far whether strong evidence results in higher guilt judgments compared to weaker

### The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

evidence when interviewers follow an interview script. Confirmation bias was expected to be especially strong when strong evidence is presented compared to weak evidence, because interviewers are more confident about the suspect's guilt pre-interview (Ask & Granhag, 2005). The current study showed strength of evidence makes individuals more likely to assume guilt and be more confident about it both in pre-and post-measure. However, there is no indication arguments against this evidence are more or less effective depending on the initial strength of evidence because the reduction in guilt judgments was more or less the same in both evidence conditions. But strength of evidence is still very important because a higher assumed guilt pre-interview means people are still ultimately more likely to think an innocent suspect is guilty because interviewers are more confident in their baseline guilt judgements.

When considering the practical implications of this, it becomes clear that police officers (and other judges of guilt) can and do change their minds about the suspect's guilt. However, the majority exhibits a confirmation bias when judging guilt as seen in this study and existing literature (Ask & Granhag, 2005; Hill et al., 2008). While individuals become less confident in their judgments after talking to the innocent suspect, it was only the minority altering from judging the suspects as guilty to innocent. The more confident individuals are in their judgements before the interview, the stronger their confirmation bias is and the more likely they judge the innocent suspect to be guilty (Rollwage et al., 2020). Individuals are especially confident in their judgments pre-interview when strong evidence is presented leading to higher baseline guilt assumptions. This in turn strengthens the interviewers' biases resulting in higher guilt judgments.

However, not all individuals show a confirmation bias or at least exhibit it to a lesser extent. For example, in the sample of Kerstholt's et al. (2010) study individuals were not prone to a confirmation bias when judging suspects. As outlined, individuals exhibit confirmation biases to a

different extent and the specific conditions in which individuals show a confirmation bias are unclear (Rassin, 2008). Ask and Granhag (2005) hypothesized investigators with a high need for cognitive closure showed a stronger confirmation bias and, therefore, were more likely to judge suspects guilty. However, they only found mixed support for their hypotheses (Ask & Granhag, 2005). Generally, literature about individuals judging suspects and the associated confirmation bias is mixed as some studies found evidence for (e.g., Ask & Granhag, 2005) and some studies evidence against police officers exhibiting a confirmation bias (e.g., Kerstholt et al., 2010) and the specific reasons for this remain unclear. Therefore, it should be researched under which circumstances individuals presuming guilt exhibit a confirmation bias, and what prevents individuals from it (Rassin, 2008). The importance of conducting further research in this area is highlighted once again by the fact that interviewers making ill-founded hypotheses can present serious threats to not only the security of the innocent suspects but also to the effectiveness of the law-enforcement system in general (Ask & Granhag, 2005).

### ***Rapport & Trust***

It was expected presuming guilt leads to lower trust and rapport ratings in both interviewer and suspect. However, there were no statistically significant effects of guilt presumptions on trust and rapport. In fact, the mean ratings of trust and rapport were almost identical between the different conditions which is not a surprise given the manipulation of guilt failed. However, there was a significant negative relationship between interviewers' guilt judgments after the interview and their trust and rapport ratings. The more confident interviewers were in their guilt judgments post-interview, the lower they perceived trust and rapport with the participant.

When there is strong evidence and additionally the interviewer presumes guilt the expectation was rapport and trust are lowest. However, evidence strength did not affect the trust

and rapport ratings of both interviewer and interviewee except for the effect of evidence strength on interviewers' rapport ratings which was marginally significant. Interviewers presented with weak evidence rated the established rapport with the suspect to be higher than interviewers presented with strong evidence. A possible explanation for this is strong evidence resulting in interviewers being more confident the suspect is guilty resulting in less rapport-establishing behaviours such as being friendly, respectful and attentive to the suspect and verbal behaviours (e.g., active listening, head nods, turn-taking) (Weiher, 2020). Consequently, the interviewer perceives less rapport.

Literature showed an agreement about the level of rapport between the suspect and a neutral observer (Richardson & Nash, 2022). However, these parties' judgments did not align with the interviewer's rapport rating. Richardson & Nash (2022) hypothesize the weak correlation between interviewers' and interviewees' rapport ratings is due to the interviewers' high cognitive demands (Richardson & Nash, 2022). Moreover, rapport is hypothesized to be an individual state rather than a shared perception (DePaulo & Bell, 1990; Weiher, 2020). Therefore, it was expected interviewers generally rate the suspect to be less trustworthy and perceive less rapport compared to the interviewee. The results supported this and showed interviewers rated trust and rapport lower than suspects. Thus, these findings were in line with the explored literature. According to Richardson & Nash (2022), this difference occurs because the demands of the interview interfere with the interviewer's rapport-building efforts resulting in them omitting cues others take into account. The study at hand contributed to these results by showing no correlation between interviewers' and interviewees' rapport ratings despite the reduced cognitive load of the interviewer due to interviewers simply reading out questions from an interview script in this study. This suggests a different explanation is needed to explain the non-existent correlation between interviewers' and



interviewees' ratings of rapport. Weiher (2020) explained this lack of correlation is due to the measure (RS3i) used. The interviewers' version of RS3i measures how the interviewer perceives the source's rating of rapport (e.g., "The Source probably thinks that I am generally trustworthy.", Duke, 2013) instead of measuring how the interviewers themselves perceive rapport. However, this explanation is not valid for the current study as this study utilized an altered form of the interviewee version of the RS3i measuring how the interviewers themselves perceived rapport. Therefore, future studies should investigate the lack of correlation between interviewers' and interviewees' rapport.

Next to the rapport ratings, this study has also focused on the interpersonal concept of trust. Interviewers who are confident in their guilt ratings after the interview were more likely to judge the established trust as low. To establish trust, individuals must increase their vulnerability to each other (Pearce, 1974). When the interviewers presume suspects are guilty, they might perceive a high-risk situation and, thus, decrease their vulnerability to each other leading to low perceived trust. However, as for rapport ratings, the reasons why this effect only occurs for the interviewers' ratings are unclear and should be investigated by future studies.

To summarize, participants presented with strong evidence against the suspect are more confident in their guilt judgements. This results in participants rating the established rapport as lower compared to participants presented with weak evidence. Moreover, when participants are confident in their guilt judgments post-interview, they perceive less trust because they might assess the situation as risky. Interviewers' rapport and trust ratings do not align with the interviewees' ratings. It is advised future studies investigate this lack of correlation.

### **Strengths**

When exploring literature about investigative interviews, most studies do not utilize an interview script, but participants devise their own questions or select from a pool of questions (Hill et al., 2008; Kassin et al., 2003). It could be argued participants experience the interview to be more realistic when selecting their own questions. However, individuals with no expertise as interviewers could be overwhelmed by the task of generating questions and might be more concentrated on their questions rather than listening to the suspects' answers. Moreover, it would not be possible to isolate the effect of guilt presumptions and evidence strength from the differences in questioning. Therefore, this study provided insights into whether guilt presumptions bias the processing of otherwise identical information. Moreover, the cognitive load of the interviewers was reduced and enabled them to concentrate more on the interviewees' stories. When using an interview script, one-third of interviewers changed their minds after listening to the suspects' argumentations demonstrating their innocence.

### **Limitations, Suggestions for Improvements and Future Research**

After considering the implications of the results and the study's strengths, it is essential to look critically at the results again and inspect to what extent they are generalisable.

First, it should be noted some participants felt the study to be moderately realistic which was reflected in their comments and their immersion scores at the end of the study. After excluding the participants with a low immersion score, the results did not change. Nevertheless, the interview was rather artificial and differed from a realistic environment by being conducted in an online environment. Individuals rely on non-verbal behaviour when judging others (Novotny et al., 2018). However, due to the online environment of the study, the access to non-verbal cues was limited.

Therefore, future research should establish a more realistic interaction between interviewer and suspect to obtain more accurate results.

Moreover, the researchers themselves played the role of the suspects. Special care was taken that the researcher playing the suspect was not acquainted with the interviewer. However, most of the interviewers were the same age as the researchers which could have reduced the realism of the study due to group favouritism (Chen & Li, 2009). Due to the similar age, participants could have seen the suspect as an ingroup member and vice versa. This would be problematic as there exists evidence individuals are more altruistic toward ingroup members (e.g., less likely to punish ingroup members for misbehaviour, and more social-welfare-maximizing actions) compared to an outgroup member which would bias the results (Chen & Li, 2009). Furthermore, only two different individuals played the role of the suspect. Theoretically, the interviews did not differ between the two suspects, however, it is possible the interviewers perceived them differently because their style of answering the questions or their body language differed (Vrij, 1993). Furthermore, one of the suspects was female and the other male which possibly influenced the results as well because females are perceived as more trustworthy which could have biased the interviewers' trust ratings of the suspects (Sun et al., 2007). Moreover, it is possible the two researchers playing the suspects generally perceive trust and rapport with other individuals as high which would bias the ratings and would limit generalising results (Abbe & Brandon, 2014). Due to these shortcomings about the two researchers playing the suspect's role, it is advised to conduct the study again with a various sample of suspects meaning uninvolved and a greater quantity of individuals playing the suspect. Another possibility would be analysing the results with the interviewee factored into the analysis as a random effect.

For all participants, the study's language (English) was a foreign language. Some participants gave feedback they struggled with communicating in English during the interview. Moreover, individuals are less emotional when communicating in their foreign language compared to their mother tongue (Caldwell-Harris, 2014). Especially with Guilt, Trust and Rapport ratings, this is crucial as emotions may play an important role. Thus, the language barrier could have reduced the validity and reliability of ratings. Future research should enable participants to conduct the study in their first language to increase the accuracy of the results.

As aforementioned, the guilt manipulation of this study did not work as there was no difference between the guilt presumption and the no guilt presumption condition. Conveniently, there were pre-and post-measures of guilt which could be compared. However, future research should ensure participants presume guilt versus not presume guilt to strengthen the confidence of the results. An example to achieve this is by manipulating participants' expectations through information about base rates. For example, in the study of Kassir et al. (2003) participants in the guilty expectation condition received the information four out of five suspects in the study committed the crime while in the innocent expectation condition this number is one out of five (20%). Another way of manipulating guilt is to introduce the possibility of an alternative suspect. Sjölin Wanner (2020) divided participants in two conditions, namely the suspect guilty condition, where participants were informed about the suspect's potential motive for the crime, and the alternative suspect condition, where another potential suspect was introduced. Thus, there are multiple possibilities to increase the chance guilt manipulations work in future studies.

As there is too little research about the effect of guilt presumptions and strength of evidence on the interviewer's and interviewee's ratings it is advisable to conduct more research about this important societal matter. The relevance of conducting further research in this area is emphasized

once again by the fact that interviewers making ill-founded hypotheses can pose serious threats to both the security of the innocent suspects and the efficiency of the law-enforcement system in general (Ask & Granhag, 2005). To be sure about the specific impact of presuming guilt and facing strong evidence on police officers when interviewing a suspect, the study requires to be replicated in a more realistic environment, with a wider range of suspects, in the participants' first language, and with a functioning guilt presumption manipulation.

### **Conclusion**

This study clarified only around one-third of participants changed their perception from judging the suspects to be guilty to judging them to be not guilty after listening to their story outlining their innocence in an interview. Moreover, participants were less confident about the guilt after listening to the suspect compared to before the interview. If police officers are forced to ask questions not directly implying guilt, the suspects can thoroughly explain themselves and the police officers may be less inclined to exhibit confirmation biases. Interviewers' trust and rapport ratings were generally lower implying interviewers perceive the basis of the interaction differently than suspects.

One practical key implication of this study is to shed light on the high rate of guilt assumptions and how hard individuals change their minds about a suspect's guilt especially when evidence initially seems to be strong.

### References

- Abbe, A., & Brandon, S. E. (2014). Building and maintaining Rapport in investigative interviews. *Police Practice and Research, 15*(3), 207-220. <https://doi.org/10.1080/15614263.2013.827835>
- Adams-Quackenbush, N. M., Horselenberg, R., & Tomas, F. (2019). Detecting guilt presumption in a police-suspect interview: An evaluation of the questions in a Dutch murder case. *Investigative Interviewing: Research and Practice, 10*(1), 37-60. [https://www.researchgate.net/publication/338264524\\_Detecting\\_Guilt\\_Presumption\\_in\\_a\\_Police-Suspect\\_Interview\\_An\\_Evaluation\\_of\\_the\\_Questions\\_in\\_a\\_Dutch\\_Murder\\_Case](https://www.researchgate.net/publication/338264524_Detecting_Guilt_Presumption_in_a_Police-Suspect_Interview_An_Evaluation_of_the_Questions_in_a_Dutch_Murder_Case)
- Alison, L. J., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and Rapport gets results: Observing Rapport-Based Interpersonal Techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy, and Law, 19*(4), 411. <https://psycnet.apa.org/doi/10.1037/a0034564>
- Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of investigative psychology and offender profiling, 2*(1), 43-63. <https://doi.org/10.1002/jip.19>
- Caldwell-Harris, C. L. (2014). Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in psychology, 5*, 1055. <https://doi.org/10.3389/fpsyg.2014.01055>
- Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review, 99*(1), 431-57. <https://doi.org/10.1257/aer.99.1.431>
- Collins, K., & Carthy, N. (2018). No Rapport, no comment: The relationship between Rapport and communication during investigative interviews with suspects. *Journal of Investigative Psychology and Offender Profiling, 16*(1), 18-31. <https://doi.org/10.1002/jip.1517>
- Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of Rapport in forensic interviewing. *Psychiatry, psychology and law, 9*(1), 69-78. <https://doi.org/10.1375/pplt.2002.9.1.69>
- DePaulo, B. M., & Bell, K. L. (1990). Rapport is not so soft anymore. *Psychological Inquiry, 1*(4), 305-308. [https://doi.org/10.1207/s15327965pli0104\\_6](https://doi.org/10.1207/s15327965pli0104_6)
- Duke, M. C. (2013). *The development of the rapport scales for investigative interviews and interrogations*. The University of Texas at El Paso.

- <https://www.proquest.com/openview/9e28c4d34519e4ec2031506d9baeea48/1?pq-origsite=gscholar&cbl=18750>
- Duke, M. C., Wood, J. M., Bollin, B., Scullin, M., & LaBianca, J. (2018). Development of the Rapport Scales for Investigative Interviews and Interrogations (RS3i), Interviewee Version. *Psychology, Public Policy, and Law*, 24(1), 64.  
<https://psycnet.apa.org/doi/10.1037/law0000147>
- Gabbert, F., Hope, L., Luther, K., Wright, G., Ng, M., & Oxburgh, G. (2021). Exploring the use of Rapport in professional information-gathering contexts by systematically mapping the Evidence base. *Applied Cognitive Psychology*, 35(2), 329-341.  
<https://doi.org/10.1002/acp.3762>
- Gatenby, E. (2020). The Credibility and Reliability of the Evidence of a Witness with a Mental Condition: An Analysis of New Zealand Evidence Law (Bachelor Thesis). University of Otago, New Zealand.  
<https://www.otago.ac.nz/law/research/journals/otago828528.pdf>
- Grahe, J. E., & Bernieri, F. J. (1999). The importance of nonverbal cues in judging rapport. *Journal of Nonverbal behavior*, 23(4), 253-269.  
<http://dx.doi.org/10.1023%2FA%3A1021698725361>
- Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*, 18(2), 341-355.  
<https://doi.org/10.1111/j.2044-8333.2012.02047.x>
- Granot, Y., Balcetis, E., Feigenson, N., & Tyler, T. (2018). In the eyes of the law: Perception versus reality in appraisals of video evidence. *Psychology, Public Policy, and Law*, 24(1), 93–104. <https://psycnet.apa.org/doi/10.1037/law0000137>
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. John Wiley & Sons. <https://psycnet.apa.org/record/2003-04147-000>
- Gudjonsson, G. H. (2007). Investigative interviewing. In *Handbook of criminal investigation* (pp.466-492). Willian Publishing.  
<https://books.google.nl/books?hl=de&lr=&id=DNb6cyzIT5oC&oi=fnd&pg=PA466&ots=ZempKeFeuD&sig=YPtZXtiTRkMBAB5YggkldhSTHs#v=onepage&q&f=false>
- Hill, C., Memon, A., & McGeorge, P. (2008). The role of confirmation bias in suspect

- interviews: A systematic evaluation. *Legal and Criminological Psychology*, 13(2), 357-371. <https://doi.org/10.1348/135532507X238682>
- Izotovas, A., Kelly, C., & Walsh, D. (2021, 2021/12/01). The Domains of PEACE: Examining Interviews with Suspected Sex Offenders. *Journal of Police and Criminal Psychology*, 36(4), 743-757. <https://doi.org/10.1007/s11896-021-09465-8>
- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003, 2003/04/01). Behavioral Confirmation in the Interrogation Room: On the Dangers of Presuming Guilt. *Law and human behavior*, 27(2), 187-203. <https://doi.org/10.1023/A:1022599230598>
- Kerstholt, J., Eikelboom, A., Dijkman, T., Stoel, R., Hermsen, R., & van Leuven, B. (2010). Does suggestive information cause a confirmation bias in bullet comparisons?. *Forensic Science International*, 198(1-3), 138-142. <https://doi.org/10.1016/j.forsciint.2010.02.007>
- Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on Trust for management: A field quasi-experiment. *Journal of applied psychology*, 84(1), 123. <https://doi.org/10.1037/0021-9010.84.1.123>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational Trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/10.5465/amr.1995.9508080335>
- Milne, B., & Powell, M. (2010). Investigative interviewing. In *The Cambridge handbook of forensic psychology* (pp. 208-214). Cambridge University Press. <https://doi.org/10.1017/CBO9780511730290>
- Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *The British Journal of Criminology*, 32(1), 23-40. <https://doi.org/10.1093/oxfordjournals.bjc.a048178>
- Newburn, T., Williamson, T., & Wright, A. (2012). *Handbook of criminal investigation*. Routledge. <https://www.ojp.gov/ncjrs/virtual-library/abstracts/handbook-criminal-investigation-0>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175-220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Novotny, E., Carr, Z., Frank, M. G., Dietrich, S. B., Shaddock, T., Cardwell, M., & Decker, A. (2018). How people really suspect and discover lies. *Journal of Nonverbal Behavior*, 42(1), 41-52. <https://doi.org/10.1007/s10919-017-0263-2>



The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

- O'Brien, B., & Ellsworth, P. C. (2006). Confirmation bias in criminal investigations. *1st annual conference on empirical legal studies paper*.  
<https://dx.doi.org/10.2139/ssrn.913357>
- Pearce, W. B. (1974). Trust in interpersonal communication. *Speech Monographs*, 41(3), 236-244. <https://doi.org/10.1080/03637757409375842>
- Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, 67(5), 1035-1041.  
<https://psycnet.apa.org/doi/10.1016/j.jbusres.2013.08.010>
- Rassin, E. (2008). Individual differences in the susceptibility to confirmation bias. *Netherlands journal of Psychology*, 64(2), 87-93. <https://doi.org/10.1007/BF0307641>
- Reavey, P., Wilcock, R., Brown, S. D., Batty, R., & Fuller, S. (2016). Legal professionals and witness statements from people with a suspected mental health diagnosis. *International Journal of Law and Psychiatry*, 46, 94-102. <https://doi.org/10.1016/j.ijlp.2016.02.040>
- Richardson, B. H., & Nash, R. A. (2022). 'Rapport myopia' in investigative interviews: Evidence from linguistic and subjective indicators of Rapport. *Legal and Criminological Psychology*, 27(1), 32-47. <https://doi.org/10.1111/lcrp.12193>
- Roberts, K. (2010). Great expectations: Relations of Trust and confidence in police interviews with witnesses of crime. *Policing: A Journal of Policy and Practice*, 4(3), 265-272.  
<https://doi.org/10.1093/police/paq022>
- Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020). Confidence drives a neural confirmation bias. *Nature communications*, 11(1), 1-11.  
<https://dx.doi.org/10.1038/s41467-020-16278-6>
- Schollum, M. (2005). Investigative interviewing: The literature. New Zealand Police Department.  
<https://www.police.govt.nz/resources/2005/investigative-interviewing/investigative-interviewing.pdf>
- Shepherd, E. & Griffiths, A. (2013). *Investigative interviewing: The conversation management approach*. Oxford University Press, USA.  
[https://www.researchgate.net/publication/323666069\\_Investigative\\_InterviewingThe\\_Conversation\\_Management\\_approach](https://www.researchgate.net/publication/323666069_Investigative_InterviewingThe_Conversation_Management_approach)
- Sjölin Wanner, M. (2020). The Effect of Guilt Presumptions on the Perception of Suspect

The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

- Interrogations. <http://hdl.handle.net/2077/65431>
- Stephenson, G. M., & Moston, S. J. (1993). Attitudes and assumptions of police officers when questioning criminal suspects. *Issues in Criminological & Legal Psychology*, 18, 30–36. <https://psycnet.apa.org/record/1993-39704-001>
- Sue, S., Smith, R. E., & Caldwell, C. (1973). Effects of inadmissible Evidence on the decisions of simulated jurors: A moral dilemma. *Journal of Applied Social Psychology*, 3(4), 345-353. <https://doi.org/10.1111/j.1559-1816.1973.tb02401.x>
- Sun, X., Wiedenbeck, S., Chintakovid, T., & Zhang, Q. (2007). The effect of gender on trust perception and performance in computer-mediated virtual environments. *Proceedings of the American Society for Information Science and Technology*, 44(1), 1-14. <https://doi.org/10.1002/meet.1450440211>
- Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *The Journal of social psychology*, 133(5), 601-610. <https://doi.org/10.1080/00224545.1993.9713915>
- Weiherr, L. (2020). *The impact of interview style on the development, maintenance, and transfer of rapport*. Lancaster University. <https://doi.org/10.17635/lancaster/thesis/1147>
- Williamson, T. (2013). *Investigative interviewing* (2<sup>nd</sup> ed.). Routledge. <https://books.google.nl/books?hl=de&lr=&id=HDXB51xHnFMC&oi=fnd&pg=PP1&ots=i0xa3z4ArS&sig=ILYOHHaDu-8EUWiSxRT42W2JMoA#v=onepage&q&f=false>
- Williamson, T. M. (1993). From interrogation to investigative interviewing; strategic trends in police questioning. *Journal of Community & Applied Social Psychology*, 3(2), 89-99. <https://doi.org/10.1002/casp.2450030203>
- Wilson, L., & Walsh, D. (2019). Striving for impartiality: Conflicts of role, Trust and emotion in interpreter-assisted police interviews. *Pragmatics and Society*, 10(1), 122-151. <https://doi.org/10.1075/ps.00018.wil>

## Appendix A

### Instructions before Interview

#### Information Sheet

Thank you for your interest in taking part in our current study. Please read the following information carefully.

#### *Our research*

This research is about investigative interviewing. We will ask you to play the role of a **police interviewer**. You will interview another participant who will play the role of the **suspect**.

We will give you information about the case including the crime which the suspect is accused of and the evidence against the suspect. We will also provide you with the exact list of questions you ask the suspect. Thus, you just have to read the questions out loud (you do not need to learn them by heart) and can pay attention to the suspect's responses. This helps you when deciding if they are guilty of the crime or not. After the interview, we ask you to fill out an online questionnaire about your experience.

#### *Preparation*

You have already made an appointment with the researchers. In case you have to change the date or time, please tell the researchers and you can reschedule, their email addresses can be found below. You can **access the Zoom Meeting via this link:**

<https://utwente-nl.zoom.us/j/9132871476?pwd=anA5a1p3eHF4R00vZ25BSENFN05Kdz09>

Meeting-ID: 913 287 1476

Code: SFE4eK

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

We ask you to prepare for the interview by reading (1) the Vignette and (2) the questions you will ask the suspect. The researcher will also send them to you via email.

### *The Meeting*

When you enter the meeting, only the researcher will be present, and you have time to discuss some organizational issues. You will have time to read the Vignette and the questions one more time. Moreover, you will be asked to fill out some questions and a consent form before the suspect enters the meeting. If you state you are ready, the researcher will leave the meeting, the suspect will enter the meeting and you can directly start to read out from the list of questions (you do not need to learn them by heart). Once you finish the interview, the suspect will leave the meeting and the researcher will enter again. You will be asked to fill out some more questions about your experience. The researcher will stay in the meeting in case you have any questions. The approximate time for the meeting is half an hour.

### *Potential Risks and Discomforts*

Taking part in this study does **not** expose you to a risk, apart from possible slight stress due to needing to role-play. This study was approved by the BMS Ethics Committee of the University of Twente (Request number: 220428)

### *Right to Withdraw and Questions*

Your participation in this study is completely voluntary. You may choose not to take part at all. If you decide to participate, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify. You can withdraw at any point during the study by stating it to the researcher and your data will not be used.

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

If you decide to stop taking part in the study, if you have questions, concerns, or complaints related to the research, you are welcome to say so to the researcher, but you do not have to explain if you do not want to.

### ***Confidentiality***

Your privacy will be protected to the maximum extent allowable by law. The gathered data will be anonymized and will not be used for any other purpose than the study itself and the reports written based on it. Any documentation created or collected as part of this study will be stored in a secure location with the University of Twente for at least ten years in line with the data policy of our faculty. Only the video could be used to identify you and we will not share this with anyone outside the research team except under extraordinary circumstances (such as court order), or unless you give us explicit permission that we can use your video to present the research.

For further information about the study and everything connected to it, you can contact the researchers themselves or the BMS ethics committee of the University of Twente.

### **Contact details for further information:**

**Luca Marie Hülscher; [l.m.hulscher@student.utwente.nl](mailto:l.m.hulscher@student.utwente.nl)**

**Jonah Sauer; [j.a.sauer@student.utwente.nl](mailto:j.a.sauer@student.utwente.nl)**

**Project Supervisor: Dr. Steven Watson; [s.j.watson@student.utwente.nl](mailto:s.j.watson@student.utwente.nl)**

### **Contact Information for Questions about Your Rights as a Research Participant**

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s),

The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by [ethicscommittee-hss@utwente.nl](mailto:ethicscommittee-hss@utwente.nl)

## **Appendix B**

### Vignette

#### **Background information about the case and list of allegations**

Imagine yourself to be a police officer at the police station in your city. Your supervisor asks you to interview the suspect involved in a new case. In the following text, your supervisor gives you some more information about the crime the suspect is accused of. This includes the evidence gathered against the suspect.

#### **Alleged offence:**

On 14/02/2022 the police arrested a woman named Mrs. Brown for dealing drugs. The woman was caught selling different types of drugs in the park of your town. The women dealt Opiates (Heroin, morphine), Hallucinogens (LSD), and Marijuana. She was arrested at 4.30pm by two police officers who were on street patrol in the park. The suspect is alleged to be her accomplice and therefore is also suspected of dealing drugs.

The evidence gathered against the suspect that may indicate they were implicated in the drug dealing offence of the 14/02/2022 is listed here:

---

#### **(When weak evidence condition)**

- An old woman (suffering from schizophrenia) saw someone that looked like the suspect together with the Mrs. Brown in the park, 5 minutes before and after Mrs. Brown dealt the drugs

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

- A 12-year-old child (suffering from autism) saw a man that might have been the suspect driving together with Mrs. Brown to the crime scene, shortly before Mrs. Brown dealt the drugs
- There was one phone call from Mrs. Brown on the suspect's phone on the day of the crime, though the content of this call is unknown
- Mrs. Brown says that she knows the suspect but refused to disclose the nature of their relationship or whether the suspect is directly involved in her drug dealing
- There were traces of marijuana found in the car of the suspect

### **(When strong evidence condition)**

- There is surveillance camera footage that the suspect was together with Mrs. Brown in the park, 5 minutes before and after Mrs. Brown dealt drugs
- There is surveillance camera footage that the suspect drove together in a car with Mrs. Brown to the crime scene, shortly before Mrs. Brown dealt the drugs
- There were 2 specific messages from Mrs. Brown on the suspect's phone, these messages specify the time and place of the drug deal
- The suspect and Mrs. Brown are friends on Facebook and follow each other on Instagram and Mrs. Brown claims the suspect provided her with the drugs she sold
- There were traces of the same drugs (Heroin, morphine, LSD, Marijuana) in the car from exactly the same batch as Mrs. Brown dealt

---

### **(When guilt presumption condition)**

Based on this evidence, your supervisor is **very sure that the suspect is guilty** of drug dealing.



**(When no Guilt Presumption condition)**

Based on this evidence, your supervisor is not sure whether the suspect is guilty of drug dealing.

---

Your task is to question the suspect, who will be played by another participant. To help you, a script has been provided which gives you the questions you should put to the suspect.

You can assume that the introduction part of the interview, where you introduce yourself to the suspect and explain the legal rights to the suspect, is already done. It has also been explained that he is being questioned because of his links to a woman who was arrested for dealing drugs. Now you are only collecting the suspect's version of events. This means you can directly ask the questions we have provided without having to introduce yourself.

Please **read these questions in order**, think carefully about the suspects responses, and afterwards we will ask you questions about your experience in the interview.

**Appendix C**

## Interview Script

**Interview for weak evidence condition**

**Interviewer:** Can you please tell me your version of events?

**Interviewee:** I don't really have a lot to say. I don't even know this woman and now I am suspected of being involved in a crime with her. I don't really know what to say.

**Interviewer:** Can you tell me what you did on the afternoon of the 14/02/2022.

**Interviewee:** It was a Monday. On Monday afternoons I am usually working. I am a taxi driver in my town, and my work shift is always from 2pm until 8pm. I remember that this day was a really busy day, and I had a lot of clients.

**Interviewer:** Do you remember any of the clients you had that day?

**Interviewee:** I am sorry. I don't remember the clients I had that day, because there were so many. But I can ask my boss if he still has the list of clients I drove on that day.

**Interviewer:** Do you remember anything else?

**Interviewee:** No, sorry. Maybe you could explain why you are questioning me, because at this point I really don't know why I am being suspected of the crime.

**Interviewer:** *An old woman saw that you were together with Mrs Brown in the park, 5 minutes before and after she dealt drugs. So can you explain why you were with Mrs Brown if you do not know her?*

**Interviewee:** Mhh... so ... sometimes during my break, which is usually at around 4pm, I go in the park and walk around a bit. You can ask my colleagues, I have done that since I started working at the taxi company, even when it is a busy day. I am sorry, I forgot to mention that earlier. Sometimes, I feel like talking with other people in the park, so I just ask them how their day was. Most of them are quite friendly, so I have a short chat with them, while walking around in the park. Maybe this Mrs Brown could have been one of the people I talked to.

**Interviewer:** *A 12 year old child saw you together with Mrs Brown in a car, driving to the crime scene, shortly before Mrs Brown dealt the drugs. Can you also explain this?*

**Interviewee:** Probably this woman was a client of mine. Probably she wanted me to drive her to the park, and then I decided to use my break talking to her in the park. I mean, to be honest, I don't remember all the people I drive, even if I have good conversations with them. There are just too many to remember them all. I am sorry.

**Interviewer:** *We know that there was one phone call from Mrs Brown on your phone. Do you know why this is?*

The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

**Interviewee:** My boss does not arrange my clients for me but I schedule my own clients. I have the same phone number for my work and my private stuff. A lot of people contact me outside of my work shift. Most of them are desperately trying to get a taxi, as there are not a lot of taxi drivers around our town. So, when I start my shift in the afternoon, I already have a lot of phone calls and messages from clients wanting me to drive me at a certain time and to a certain location.

**Interviewer:** Earlier you said that you don't know Mrs Brown. *However, she said she knows you, but refused to disclose the nature of your relationship or whether you are directly involved in her drug dealing.* Can you explain this?

**Interviewee:** As I already mentioned, I like to chat with people. I am a social person and a lot of people know me. I mean, as a taxi driver, you get around quite a lot in town and the people just know you. Many people get to know you, but the problem is that I am really bad at remembering their names and faces. I mean, I might have talked with this woman at some point. However, I am certain that I am not involved in any drug business with her, that is how much I can tell you.

**Interviewer:** *Why do you think that there were traces of Marijuana found in your car?*

**Interviewee:** I am telling you, you could basically find traces of drugs in every taxi driving around the city. I mean it is not like all my clients are saints. They certainly like to have fun. Just because there were traces of drugs in my car, it doesn't mean that they belong to me. I would never risk my job by transporting drugs in my taxi. But you said that I drove this woman, this Mrs Brown to the park. I mean if she dealt drugs in the park, she already had drugs on her and

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

this is the reason why you found traces of drugs in my taxi. There is no reason to assume the drugs were from me.

**Interviewer:** Okay, that is the end of the interview.

### **Interview for strong evidence condition**

**Interviewer:** Can you please tell me your version of events?

**Interviewee:** I don't really have a lot to say. I don't even know this woman and now I am suspected of being involved in a crime with her. I don't really know what to say.

**Interviewer:** Can you tell me what you did on the afternoon of the 14/02/2022.

**Interviewee:** It was a Monday. On Monday afternoons I am usually working. I am a taxi driver in my town, and my work shift is always from 2pm until 8pm. I remember that this day was a really busy day, and I had a lot of clients.

**Interviewer:** Do you remember any of the clients you had that day?

**Interviewee:** I am sorry. I don't remember the clients I had that day, because there were so many. But I can ask my boss if he still has the list of clients I drove on that day.

**Interviewer:** Do you remember anything else?

**Interviewee:** No, sorry. Maybe you could explain why you are questioning me, because at this point I really don't know why I am being suspected of the crime.

**Interviewer:** *There is surveillance camera footage that you were together with Mrs Brown in the park, 5 minutes before and after she was confirmed to have dealt drugs. So can you explain why you were with Mrs Brown if you do not know her?*

**Interviewee:** Mhh... so ... sometimes during my break, which is usually at around 4pm, I go in the park and walk around a bit. You can ask my colleagues, I have done that since I started working at the taxi company, even when it is a busy day. I am sorry, I forgot to mention that earlier. Sometimes, I feel like talking with other people in the park, so I just ask them how their day was. Most of them are quite friendly, so I have a short chat with them, while walking around in the park. Maybe this Mrs Brown could have been one of the people I talked to.

**Interviewer:** *There is also surveillance camera footage that shows you together with Mrs Brown in a car, driving to the crime scene, shortly before Mrs Brown dealt the drugs. Can you also explain this?*

**Interviewee:** Probably this woman was a client of mine. Probably she wanted me to drive her to the park, and then I decided to use my break talking to her in the park. I mean, to be honest, I

The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

don't remember all the people I drive, even if I have good conversations with them. There are just too many to remember them all. I am sorry.

**Interviewer:** *We know that Mrs Brown had written you two text messages earlier that day, including the time and place of the drug deal. Do you know why this is?*

**Interviewee:** My boss does not arrange my clients for me but I schedule my own clients. I have the same phone number for my work and my private stuff. A lot of people contact me outside of my work shift. Most of them are desperately trying to get a taxi, as there are not a lot of taxi drivers around our town. So, when I start my shift in the afternoon, I already have a lot of phone calls and messages from clients wanting me to drive me at a certain time and to a certain location.

**Interviewer:** *Earlier you said that you don't know Mrs Brown. However, you are friends with her on facebook and she follows you on instagram, and Mrs Brown claims the suspect provided her with the drugs she sold. Can you explain this?*

**Interviewee:** As I already mentioned, I like to chat with people. I am a social person and a lot of people know me. I mean, as a taxi driver, you get around quite a lot in town and the people just know you. Many people get to know you, but the problem is that I am really bad at remembering their names and faces. I mean, I might have talked with this woman at some point. However, I am certain that I am not involved in any drug business with her, that is how much I can tell you.

**Interviewer:** *Why do you think that there were traces of Heroin, morphine, LSD and Marijuana from the same batch as Mrs Brown had dealt, found in your car?*

**Interviewee:** I am telling you, you could basically find traces of drugs in every taxi driving around the city. I mean it is not like all my clients are saints. They certainly like to have fun. Just because there were traces of drugs in my car, it doesn't mean that they belong to me. I would never risk my job by transporting drugs in my taxi. But you said that I drove this woman, this Mrs Brown to the park. I mean if she dealt drugs in the park, she already had drugs on her and this is the reason why you found traces of drugs in my taxi. There is no reason to assume the drugs were from me.

**Interviewer:** Okay, that is the end of the interview.



## Appendix D

### Questions Rapport

#### **For interviewee (researchers)**

Please answer the following questions on the 5-point Likert scale with the response options of:

1=strongly agree

2=agree

3=neither agree nor disagree

4=disagree

5=strongly disagree

1. The interviewer really listened to what I had to say
2. The interviewer paid attention to my opinion
3. The interviewer was attentive to me
4. The interviewer was interested in my point of view
5. The interviewer was honest with me
6. The interviewer respects my knowledge
7. The interviewer can generally be trusted to keep the word
8. I can trust the interviewer to keep word to me
9. The interviewer did the job with skill
10. The interviewer performed expertly
11. The interviewer made effort to do a good job
12. The interviewer acted like a professional
13. We have our culture in common
14. The interviewer and I share one ethnicity
15. The Interviewer shares my culture
16. We worked well as a team
17. The communication went smoothly
18. The Interviewer and I got along well

#### **For interviewer (participants)**

Please answer the following questions on the 5-point Likert scale with the response options of:

1=strongly agree

2=agree

3=neither agree nor disagree

4=disagree

5=strongly disagree

1. The interviewee really listened to what I had to say
2. The interviewee paid attention to my opinion
3. The interviewee was attentive to me
4. The interviewee was interested in my point of view

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

5. The interviewee was honest with me
6. The interviewee respects my knowledge
7. The interviewee can generally be trusted to keep the word
8. I can trust the interviewee to keep word to me
9. The interviewee did the job with skill
10. The interviewee performed expertly
11. The interviewee made effort to do a good job
12. The interviewee acted like a professional
13. We have our culture in common
14. The interviewee and I share one ethnicity
15. The Interviewee shares my culture
16. We worked well as a team
17. The communication went smoothly
18. The interviewee and I got along well

## Appendix E

### Questions Trust

#### **For interviewee (researchers)**

Please answer the following questions on the 5-point Likert scale with the response options of:

1=strongly agree

2=agree

3=neither agree nor disagree

4=disagree

5=strongly disagree

1. The interviewer seemed very capable of performing their job
2. I get the impression the interviewer would be successful at the things they try to do
3. The interviewer seemed to have much knowledge about how things should be done
4. I feel very confident about the interviewer's skills
5. The interviewer seemed to have specialised capabilities
6. The interviewer seemed to be well qualified
7. The interviewer seemed very concerned about my welfare
8. My needs and desires seemed very important to the interviewer
9. I got the impression the interviewer would not knowingly do anything to hurt me
10. The interviewer seemed to really look out for what is important to me
11. I got the impression the interviewer would go out of their way to help me
12. The interviewer seemed to have a strong sense of justice
13. I didn't have to wonder whether the interviewer would stick to their word
14. The interviewer seemed to try hard to be fair in dealings with others
15. The interviewer's actions and behaviours were not very consistent
16. I like the interviewer's values
17. Sound principles seemed to guide the interviewer's behaviour

#### **For interviewer (participants)**

Please answer the following questions on the 5-point Likert scale with the response options of:

1=strongly agree

2=agree

3=neither agree nor disagree

4=disagree

5=strongly disagree

1. The interviewee seemed very capable of answering the questions
2. I get the impression the interviewee would be successful at the things they try to do
3. The interviewee seemed to have much knowledge about how to behave appropriately in the interview
4. I feel very confident about the interviewee's skills (apart from this interview)

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

5. The interviewee seemed to have specialised capabilities (apart from this interview)
6. The interviewee seemed to be well qualified (apart from this interview)
7. The interviewee seemed like someone concerned about the welfare of others
8. The needs and desires of others seemed very important to the interviewer
9. I got the impression the interviewee would not knowingly do anything to hurt me
10. The interviewee seemed to really look out for what is important to me
11. I got the impression the interviewee would go out of their way to help me
12. The interviewee seemed to have a strong sense of justice
13. I didn't have to wonder whether the interviewee would stick to their word
14. The interviewee seemed to try hard to be fair in dealings with others
15. The interviewee's actions and behaviours were not very consistent
16. I like the interviewee's values
17. Sound principles seemed to guide the interviewee's behaviour

**Appendix F**

## Questions Immersion

To what extent do you agree/disagree with the following statements?

|  | Strongly disagree     | Somewhat disagree     | Neither agree nor disagree | Somewhat agree        | Strongly agree        |
|--|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|
| I took the task seriously                | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>      | <input type="radio"/> | <input type="radio"/> |
| I could imagine myself as an interviewer | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>      | <input type="radio"/> | <input type="radio"/> |
| The interview was somewhat realistic     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>      | <input type="radio"/> | <input type="radio"/> |
| The suspect played his/her part well     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>      | <input type="radio"/> | <input type="radio"/> |

Do you have anything left to say to the researchers?

## Appendix G

### Debriefing

Thank you for your participation! After this study you might ask yourself what will happen with the data. Our study will help us to understand how presumptions of guilt affect how interviewers interpret explanations provided by suspects.

In our study all the suspects that are interviewed are innocent and provide identical accounts to the interviewer. However, we manipulated prior guilt assumptions and the apparent strength of the evidence against the suspect to determine whether these made the accounts provided by suspects less likely to be believed and whether this might affect investigative decision making such as whether to continue investigating the suspect. We did this by changing the information you received prior to the interview taking place.

We also told you that the suspect was another participant. In truth, they were part of the research team. We apologise for this deception, but it was necessary in order to have participants focus on the narrative provided by the suspect.

In case you would like to know more about the study, the theoretical background or the study findings, feel free to contact any of the researchers. If you have questions or concerns about the study, please do not hesitate to reach out to the researchers.

If you have enjoyed taking part please feel free to share our contact details with your friends, however to maintain the integrity of the study, please do not share with anyone the information in this form about the specific ideas we are testing or how we test them.

**Luca Marie Hülscher;**

**Jonah Sauer;**

**Project Supervisor: Dr. Steven Watson;**

[l.m.hulscher@student.utwente.nl](mailto:l.m.hulscher@student.utwente.nl)

[j.a.sauer@student.utwente.nl](mailto:j.a.sauer@student.utwente.nl)

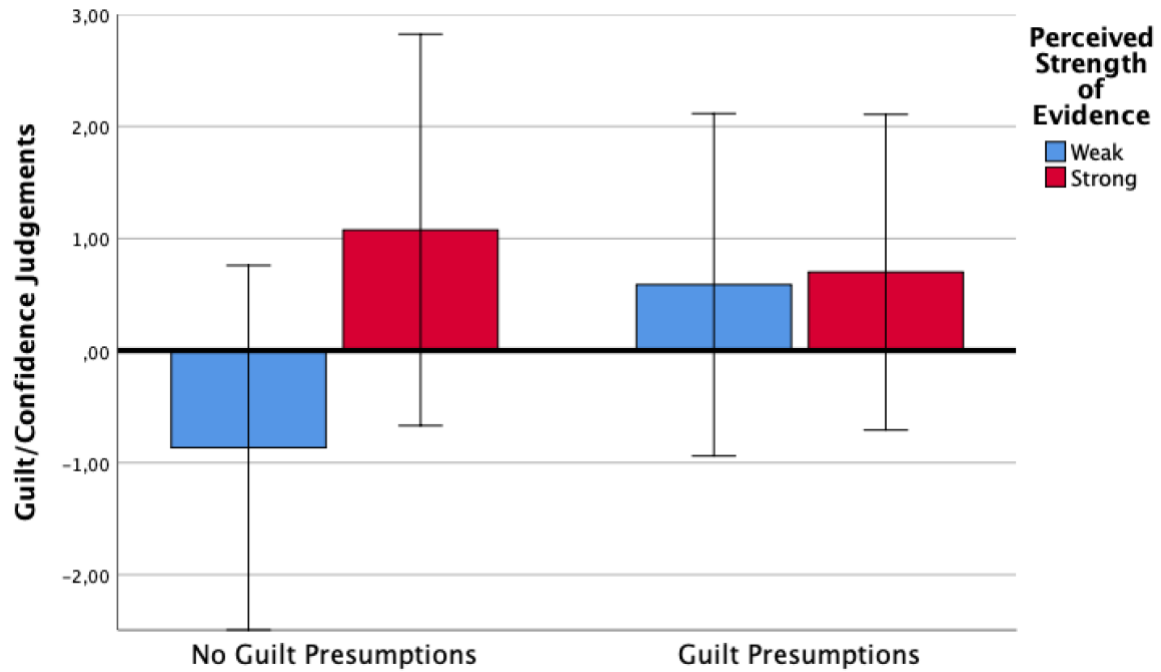
[s.j.watson@student.utwente.nl](mailto:s.j.watson@student.utwente.nl)

## Appendix H

### Originally planned analysis

A two-way ANOVA was conducted with Guilt Judgment as dependent variable and Guilt Presumptions and Evidence Strength as independent variables.

There were no significant main effects for either Guilt Presumption,  $F(1,61) = 0.46, p = .50, \text{partial } \eta^2 = .01$ , or Evidence Strength,  $F(1,61) = 1.69, p = .20, \text{partial } \eta^2 = .03$ . There was no difference in perceived Guilt Judgment between the Guilt Presumption condition versus the No Guilt Presumption condition ( $M = 0.04, SD = 3.04$  versus  $M = 0.65, SD = 3.26$ ) and there was no difference in Guilt Judgment regarding Weak versus Strong Evidence ( $M = -0.09, SD = 2.99$  versus  $M = 0.85, SD = 3.29$ ). There was also no significant interaction effect between Guilt Presumptions and Evidence Strength,  $F(1,61) = 1.34, p = .25, \text{partial } \eta^2 = .02$ . This illustrates there was no difference in Guilt Judgment for participants who presumed Guilt and were presented with Weak Evidence ( $M = 0.58, SD = 2.28$ ), and Strong Evidence ( $M = 0.70, SD = 3.33$ ) and who did not presume Guilt and were presented with Weak Evidence ( $M = -0.86, SD = 2.50$ ), and Strong Evidence ( $M = 1.08, SD = 3.35$ ). Figure 6 visualizes these findings presented by a bar chart.

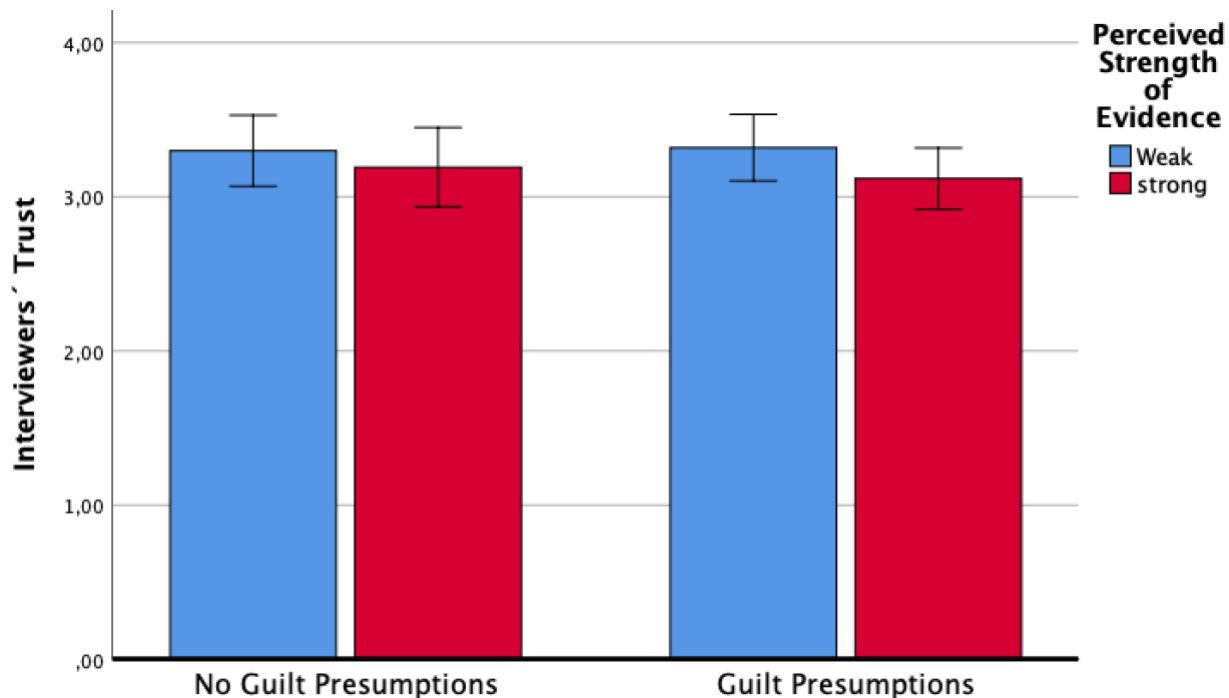
**Figure 6***Effects of Guilt Presumptions and Evidence Strength on Guilt Judgment*



## Appendix I

### Interviewers' and Interviewees' Trust and Rapport Ratings

**Interviewers' Trust Ratings.** For the interviewers' Trust ratings, there were no statistically significant main effects for both Guilt Presumption,  $F(1,60) = 0.06, p = .82$ , partial  $\eta^2 < .01$ , or Evidence Strength,  $F(1,60) = 1.84, p = .18$ , partial  $\eta^2 = .03$ . There was no difference in perceived Trust between the Guilt presumption condition versus the No Guilt presumption condition ( $M = 3.21, SD = 0.38$  versus  $M = 3.25, SD = 0.52$ ) and there was no difference in Trust regarding Weak versus Strong Evidence ( $M = 3.31, SD = 0.44$  versus  $M = 3.15, SD = 0.44$ ). There was also no statistically significant interaction effect between Guilt Presumptions and Evidence Strength,  $F(1,60) = 0.17, p = .68$ , partial  $\eta^2 < .01$ . This illustrates there was no difference in Trust for participants who presumed Guilt and were presented with Weak Evidence ( $M = 3.32, SD = 0.36$ ), and Strong Evidence ( $M = 3.11, SD = 0.37$ ) and who did not presume Guilt and were presented with Weak Evidence ( $M = 3.29, SD = 0.52$ ), and Strong Evidence ( $M = 3.19, SD = 0.55$ ). Figure 7 displays a bar chart visualizing these findings.

**Figure 7***Effects of Guilt Presumptions and Evidence Strength on Interviewers' Trust*

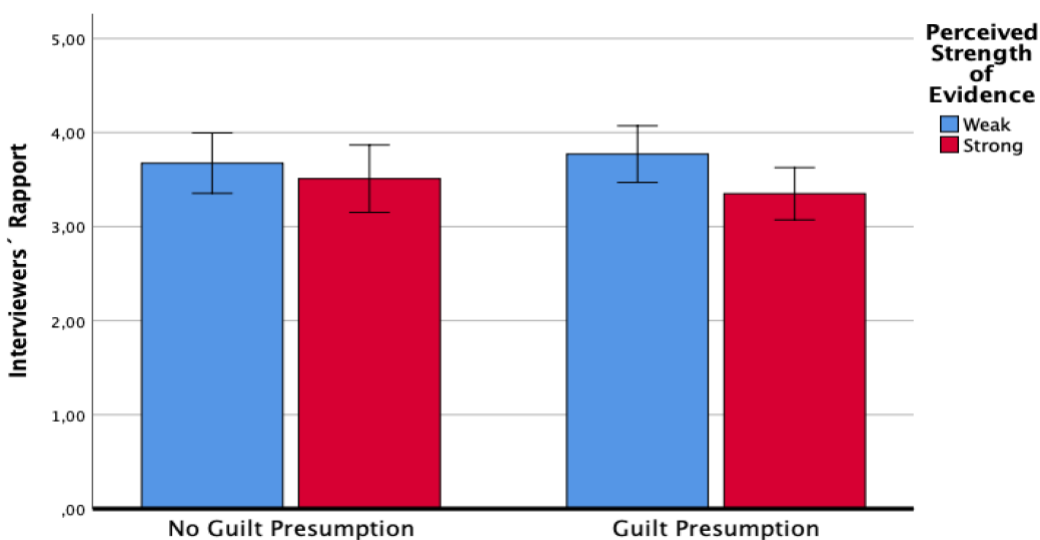
**Interviewers' Rapport Ratings.** For the interviewers' Rapport ratings, there was no statistically significant main effects for Guilt Presumption,  $F(1,60) = 0.04, p = .84, \text{partial } \eta^2 < .01$ . There was no difference in Rapport between the Guilt Presumption condition versus the No Guilt Presumption condition ( $M = 3.54, SD = 0.67$  versus  $M = 3.60, SD = 0.58$ ). There was a marginally significant main effect for Evidence Strength,  $F(1,60) = 3.45, p = .07, \text{partial } \eta^2 = .05$ . There is a possibility that interviewers rated Rapport higher in the Weak versus Strong Evidence condition ( $M = 3.73, SD = 0.61$  versus  $M = 3.41, SD = 0.62$ ). There was no statistically significant interaction effect between Guilt Presumptions and Evidence Strength,  $F(1,60) = 0.66, p = .42, \text{partial } \eta^2 = .01$ . This illustrates there was no difference in Rapport for participants who presumed Guilt and were presented with Weak Evidence ( $M = 3.77, SD = 0.68$ ), and Strong

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

Evidence ( $M = 3.35$ ,  $SD = 0.61$ ) and who did not presume Guilt and were presented with Weak Evidence ( $M = 3.67$ ,  $SD = 0.52$ ), and Strong Evidence ( $M = 3.51$ ,  $SD = 0.64$ ). Figure 8 visualizes these findings presented by a bar chart.

**Figure 8**

*Effects of Guilt Presumptions and Evidence Strength on Interviewers' Rapport*



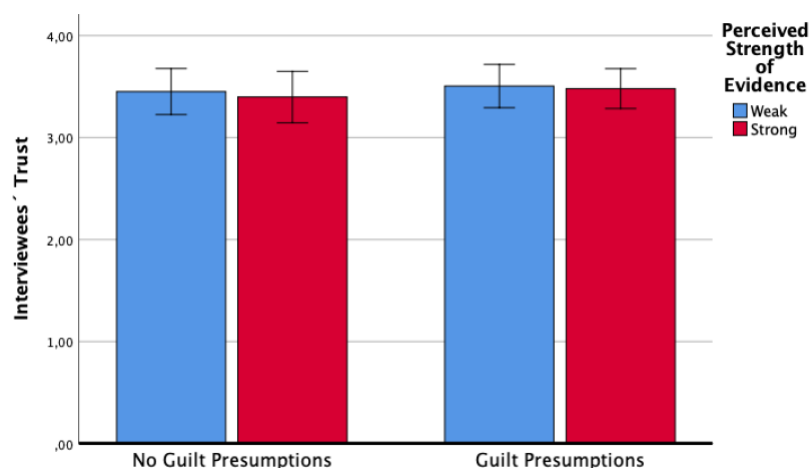
**Interviewers' Trust Ratings.** For the interviewers' Trust ratings, there were no statistically significant main effects for either Guilt Presumption,  $F(1,60) = 0.38$ ,  $p = .54$ , partial  $\eta^2 = 0.01$ , or Evidence Strength,  $F(1,60) = .3$ ,  $p = .72$ , partial  $\eta^2 < .001$ . There was no difference in perceived Trust between the Guilt Presumption condition versus the No Guilt Presumption condition ( $M = 3.43$ ,  $SD = 0.09$  versus  $M = 3.50$ ,  $SD = 0.07$ ) and there was no difference in Rapport regarding Weak versus Strong Evidence ( $M = 3.48$ ,  $SD = .08$  versus  $M = 3.44$ ,  $SD = .08$ ). There was also no statistically significant interaction effect between Guilt Presumptions and Evidence Strength,  $F(1,60) = .02$ ,  $p = .90$ , partial  $\eta^2 < .001$ . This illustrates there was no difference in Trust for participants who presumed Guilt and were presented with Weak Evidence

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

( $M = 3.50$ ,  $SD = .11$ ), and Strong Evidence ( $M = 3.48$ ,  $SD = .10$ ) and who did not presume Guilt and were presented with Weak Evidence ( $M = 3.45$ ,  $SD = .11$ ), and Strong Evidence ( $M = 3.40$ ,  $SD = .13$ ). Figure 9 visualizes these findings presented by a bar chart.

**Figure 9**

*Effects of Guilt Presumptions and Evidence Strength on Interviewees' Trust*



**Interviewees' Rapport Ratings.** For the interviewees' Rapport ratings, there were no statistically significant main effects for either Guilt Presumption,  $F(1,60) = .31$ ,  $p = .58$ , partial  $\eta^2 < .01$ , or Evidence Strength,  $F(1,60) = 0.55$ ,  $p = .46$ , partial  $\eta^2 = .01$ . There was no difference in perceived Rapport between the Guilt Presumption condition versus the No Guilt Presumption condition ( $M = 3.93$ ,  $SD = 0.42$  versus  $M = 3.88$ ,  $SD = 0.41$ ) and there was no difference in Rapport regarding Weak versus Strong Evidence ( $M = 3.95$ ,  $SD = 0.42$  versus  $M = 3.87$ ,  $SD = 0.41$ ). There was also no statistically significant interaction effect between Guilt Presumptions and Evidence Strength,  $F(1,60) = 0.00$ ,  $p = .99$ , partial  $\eta^2 < .001$ . This illustrates there was no difference in Rapport for participants who presumed Guilt and were presented with Weak Evidence ( $M = 3.97$ ,  $SD = 0.44$ ), and Strong Evidence ( $M = 3.89$ ,  $SD = 0.42$ ) and who did not

## The Effect of Guilt Presumptions and Evidence Strength within Police Interviews

presume Guilt and were presented with Weak Evidence ( $M = 3.91$ ,  $SD = 0.41$ ), and Strong Evidence ( $M = 3.83$ ,  $SD = 0.42$ ). Figure 10 visualizes these findings presented by a bar chart.

**Figure 10**

*Effects of Guilt Presumptions and Evidence Strength on Interviewees' Rapport*

