Examining the Role Plausible Deniability as a Shield for Confrontation of Sexism

Merel Destrée

S4358937

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group number: 25

Supervisor: dr. Namkje Koudenburg

Second evaluator: prof. dr. Martijn van Zomeren

In collaboration with: Zoë Richter, Carlotta Zimmeck, Nelly Bakker, Eoghan Kiely and
Jennanna Spannenburg

June 20, 2023

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Abstract

While previous research has primarily focused on strategies for addressing sexist remarks, the factors impeding the effectiveness of such confrontations remain understudied. This study aimed to explore the influence of perceived plausible deniability on the relationship between confronting a sexist joke, the formation of norms, and the recognition of sexism. Plausible deniability refers to an individual's perception that allows the individual to dismiss the content of a remark by emphasizing its presentation rather than its substance. 84 male participants were exposed to two videos featuring four individuals in conversation, followed by one person making a sexist joke. Participants encountered one of six reactions to the joke. They then completed a questionnaire, while imagining having the role of the perpetrator.

Our hypotheses suggested that high levels of perceived plausible deniability would impede the intended effects of confrontation on the perpetrators' norms (H1) and recognition of sexism (H2). The results revealed a significant interaction effect for both H1 and H2, indicating a substantial difference in the impact of the condition on norms and recognition of sexism among individuals with low scores on perceived plausible deniability. Moreover, higher levels of plausible deniability weakened the influence of the condition on recognition of sexism (H1). For H2, the effect was reversed, as higher scores of normality were observed for the sexist joke when confronted compared to the absence of confrontation.

These findings suggest that perceived plausible deniability may serve as a defense mechanism for men confronted with their sexist jokes. However, it is important to acknowledge that the limited diversity of the sample necessitates further research to validate this phenomenon.

Keywords: Confronting sexism, Plausible deniability, Norms, Recognition of sexism

Examining the Role Plausible deniability as a shield for Confrontation of sexism

Not too long ago, members of a student association in Amsterdam (ASC) incorporated blatantly sexist slogans into a speech to the entire association during a dinner. They were enthusiastically cheered on by the audience, the incident was leaked through a video that went viral on social media and subsequently made its way to the news (DPG Media Privacy Gate, n.d., para. 2). While a handful of people walked out of the room, no one confronted the situation. It was only after the video spread throughout the Netherlands and made news headlines that an apology was offered, and some members were suspended. The incident surprised people in the Netherlands, as such remarks are considered outdated and inappropriate in the present time, regardless of whether they were intended seriously or not.

However, sexist comments can still be encountered, even when watching television. For instance, during the show 'Voetbal Insite', the following remark was made: "But we have to see that [woman's football], and then you have to listen to that nonsense [female analysts]. In the past, they used to read women's magazines and sit around knitting" (Johan en René onder vuur om bekritiseren vrouwelijke analisten, 2021). Although such instances of sexist remarks often face criticism, the desired effect is not always achieved.

When examining the response of presenters on 'Voetbal Insite' to criticism regarding the show's sexism, it becomes apparent that confrontation is not effective. Van der Gijp dismisses the sexism as mere "teasing" and claims to be consistently misunderstood ("Derksen en Gijp reageren in Voetbal Inside op kritiek - Joop - BNNVARA," n.d.). However, in other cases, confrontation is acknowledged and addressed. For instance, Dutch Prime Minister Rutte was informed by female colleagues that he treated them differently by frequently interrupting them. This resulted in Rutte acknowledging the issue and committing to paying attention to it ("Ministers Klagen Over Seksisme Rutte - Joop - BNNVARA," n.d.).

The contrasting reactions in these two instances raise questions about the varying acceptance of confrontation. Why do some individuals change their behavior after a confrontation while others do not?

Responding to Sexist Remarks/Jokes

To confront a sexist remark as a victim, one must first perceive that the person making the comment holds sexist beliefs that could potentially be changed through confrontation (Ashburn-Nardo et al., 2008). The decision to confront may therefore be influenced by how the recipient perceives the (sexist) comment. A study suggested that the remark in the form of a joke was perceived as less offensive, resulting in it being seen as less worthy of confrontation than the same content presented as a direct remark (Woodzicka et al., 2015). The combination of a joke that is not taken seriously and the resulting perception that it is not worthy of confrontation can make it challenging to address a sexist joke (Cozpp & Monteith, 2003; Hyers, 2007). Similarly, another study indicates that there can be individual differences in the perception of sexual harassment (Fitzgerald et al., 1988). The study revealed that 50 percent of the participants indirectly reported experiencing sexual harassment when they answered questions about what they had experienced. However, only five percent of this group of women labeled the incident as sexual harassment themselves. The research demonstrates that individuals may often fail to recognize when something is genuinely sexist or can find it challenging to realize it. When someone does not realize that something is sexist, they are less likely to engage in confrontation.

The cognitive dissonance theory (Festinger, 1957) may shed light on the thought processes of individuals who do not respond to a sexist remark. According to this theory, when individuals seek to eliminate the discrepancy between their actions and perception of collective norms (dissonance), they must either adhere to the collective norms or make a change. This change can involve modifying the other's behavior or adjusting the collective

norms they had. In the case of a sexist remark that goes unchallenged, it is difficult for the victim to modify the behavior once the opportunity for confrontation has passed. Consequently, individuals tend to adjust their perception of collective norms to reduce cognitive dissonance. For example, a study suggested that individuals who had a missed opportunity to confront sexism exhibited a decrease in the perceived importance of confrontations, particularly among those who initially placed high importance on them (Rasinski, Geers, & Czopp, 2013). When someone consciously chooses not to react, they adjust their opinion, as opposed to when they are unable to confront the situation. Consequently, they perceive the maker of the remark as less sexist and attribute less importance to confrontation itself.

The findings can serve as an explanation for why women find sexist jokes less offensive. When a sexist remark is presented as a joke, someone can see plausible deniability as a reason to disregard the content of a joke because of the form it is presented in. For example they can use the phrase "Don't take it so seriously, it was just a joke" to dismiss a confrontation. Consequently, when a woman is not given the opportunity to confront a sexist remark disguised as a joke, because of for example plausible deniability, she may adjusts her collective norms regarding sexist expressions to cope with the joke. This, in turn, could explain why sexual harassment often goes unrecognized, as indicated by other research (Mallett et al., 2019). Thus, a sexist remark presented as a joke can lead to a modification of someone's perception of the collective norms concerning sexism. Because of this sexism in jokes is seen as more in line with what is normal, to the disadvantage of women. Confrontation sexism therefore is of importance.

Sexist Remarks as Jokes

In most cases, we encounter sexist remarks in our daily lives in the form of jokes and less in direct instances. The justification-suppression model (JSM) (Crandall & Eshelmans,

2003) may provide an explanation for this phenomenon. Due to contemporary egalitarian norms, direct sexist remarks are rarely heard in everyday life. However, when an individual holds strong biases against a particular group, they may unconsciously seek ways to no longer suppress these biases. As stated before, using plausible deniability, a joke can serve as a prime example where a person can say that the remark is funny because of the form and the content should not be taken seriously. Essentially, this affords an individual the opportunity to evoke a suppressed bias, rationalized by the premise that the joke is intended to be non-serious in nature.

This lightheartedness that is expected when a joke is made, may be explained by the Prejudiced Norm Theory (Berry & Ferguson, 2004). Which posits that when a sexist remark is presented as a joke, it is less likely to be perceived as driven by prejudice. This is because the recipient assumes a non-critical mindset, which is associated to a lighter reception of the remark. The article argues that this can be dangerous, especially for individuals with strong biases, as positive reactions may normalize these prejudices. Research confirms that people are less likely to confront a sexist remark presented in the form of a joke (Swim et al., 2005; Mallett et al., 2016). Given the results above, a confrontation may be less effective if the sexism is presented as a joke, even though it may be even more important to do so.

The confrontation

Upon engaging in confrontation, different findings have been observed. Some research shows that when confronted, individuals acknowledge their gender bias and take steps to address it (Parker et al., 2018). However, this acknowledgement is significantly lower when no evidence is presented. In most real-life cases there is no evidence and the accusation is based on someone's experience, so we can assume the acknowledgement here could be lower. Furthermore, when a woman confronts a man, this may be associated with a more friendly

and cooperative interaction, ultimately leading to mutual sympathy (Mallett & Wagner, 2011).

Despite these positive findings, there is also research that show that confrontation can be costly for the confronter (Gervais & Hillard, 2014). For instance, women who confront a man on a sexist remark tend to be more respected and liked by other women in the group. However, among men, the level of respect tends to remain the same while likability decreases compared to when they do not confront (Dodd et al., 2001). For men and women confrontation may also be seen as complaining which has consequences for the confronter (Kowalski, 1996). Is it than worth the trouble to confront when it's 'just a joke'?

Current study

The studies indicate that people are less likely to engage in confrontation when a sexist joke is made, for various reasons. Even when confronting a sexist joke might even be of more importance than direct sexism, because the reaction to sexism in the form of a joke can change people's norms. Confronting sexism does not always only have positive influences, people may perceive the confronter as less likable for instance. Is confronting a sexist joke then really worth the trouble? In other words, will confronting a sexist joke work? With this in mind, the current study will focus on examining the extent to which norm change and recognition of sexism occur when a sexist joke is confronted. The underlying idea is that when someone makes a sexist remark in the form of a joke they can in some cases build a barrier (called plausible deniability) to confrontation as they expect the remark to be received with a lighthearted attitude. When this expectation is not met and there is a confrontation, we expect two different outcomes.

Firstly, when the person receiving the confrontation perceives low plausible deniability we expect them to engage in the confrontation and adapt their behavior accordingly. Alternatively, for those that perceive high plausible deniability we expect that

they may resent the other person for taking the joke too seriously, consequently they will not change their behavior accordingly. Plausible deniability refers to an individual's perception that a confrontation to a sexist remark can be denied due to its humorous form.

Hypothesis 1. Participants with perceived low plausible deniability will adapt their norms accordingly (find the joke less normal) when confronted compared to no confrontation. With high perceived plausible deniability, confrontation will not have an effect on the participants, meaning there will not be an effect of conditions on norms.

Hypothesis 2. Low perceived plausible deniability will strengthen the influence of condition on recognition of sexism, where confrontation will lead to higher recognition and vice versa. High perceived plausible deniability we expect will ensure a non-significant result on the relation between condition an recognition of sexism.

Methods

Participants

The participants in this study were all men above the age of sixteen. Women, other gender identities, and participants younger than the age of sixteen were excluded from the study. Participants were assured that their participation was voluntary and that their responses would be kept confidential. All participants provided informed consent prior to participation in the study. Our initial raw data sample consisted of a convenience sample of 179 participants. Participants that did not answer the last question concerning the current study were excluded from the data, as were the participants that saw a humorous confrontation (further elaboration on this will be provided later in the article). This left us with a total of 84 participants. The attrition rate was 29%, with 52 participants not reaching the last question of the study. The nationality of the participants was predominantly Dutch (n = 53) and German (n = 20). The age of the respondents was between 17 and 70 years old with a mean of 25 years of age (SD = 9.51). The participants were recruited through the online forum; SONA, a

platform for students from the first year psychology bachelor where they have to earn a certain amount of credits for their curriculum (1st-year SONA-practicum Pool / 1e-jaars SONA-practicum Pool (2022-2023), z.d.). Next to SONA we recruited trough snowballing via social media, platforms, and our personal network. We compensated the participants from SONA with SONA-credits for the completion of the survey and the possibility of winning a 15 Euro bol.com voucher for the participants we got through snowballing.

Design and Procedure

The purpose of this experimental design was to establish a controlled environment in which we could investigate the impact of a confrontation on the individual who was confronted. The participants were asked to complete the study online on their device of choice, in their own personal setting (not in a lab). The study was in English and it was recommended to use headphones for optimal sound. In the study, they were asked to watch a video of a group of four young adults, two males and two females, having a conversation about their holidays. Then, one of the two men makes a sexist remark after one of the women mentioned having had a female pilot: "A woman? Most women can't even drive a car, why is she allowed to fly a plane?". The participants were instructed to imagine they were the person making that comment. The video lasted 22 seconds, after which participants were asked to answer questions about the norms and disapproval surrounding the comment and whether they believed the man expressed his true beliefs. The participants were randomly assigned to one of six experimental conditions depending on which they watched the same video with different responses, the six different conditions were distributed in a 2x3 design. Where there was either a direct confrontation, a humorous confrontation or someone changing the topic of the conversation, carried out by either a man or a woman. Changing the topic was used as the control condition. The response video was then followed by a series of questions about their approval, norms, feminism, personality traits, recognition of sexism, and plausible deniability.

The independent variables are the different reactions to the sexist remark, the gender of the confronter, feminism. The current study is part of a bigger study with different research questions, for this research question we will only look at the direct confrontation and the change of topic conditions. For the variables we will only look at the questions about norms, recognition of sexism and plausible deniability. This study was approved by the Ethics Committee of the RUG. Participants provided informed consent prior to participation in the study and were debriefed upon completion.

Tom joking: https://www.youtube.com/watch?v=3SFpOrhFWSE&feature=youtu.be

Tom confronting: https://www.youtube.com/watch?v=iLZbmC-iJnM

Tom ignoring: https://www.youtube.com/watch?v=2X4X2EnkHv0

Lucy joking: https://youtu.be/uohxdl-k91g

Lucy confronting: https://www.youtube.com/watch?v=XjumQM2ZUnI

Lucy ignoring: https://www.youtube.com/watch?v=MqiqWKgJ04o

Measures

The surveys followed a 7-point Likert scale format and yes-no style questions or statements with which they agreed or disagreed to a certain amount. The exact labels for the questions ranged from 'not at all' to 'very'. For the statements they ranged from 'strongly disagree' to 'strongly agree'. We asked the participants about their demographics, such as age, gender, profession, and country of origin.

Norms were measured twice, once before and once after the second video, using the same four scale items consisting of collective and personal norms. The questions were inspired by earlier research regarding responses to sexism (Koudenburg et al., 2020). Collective norms are perceived social expectations or rules regarding what behaviors are approved or disapproved in a given context and personal norms are individual internalized standards or beliefs about what is morally right or wrong. Participants were asked the

questions: "In this group, how normal is it to make this remark?", "in this group, how appropriate is it to make this remark?", "In this group, how funny is it to make this remark?" and "How likely is it that you would make a similar remark in a group like this?". Their response was measured via a 7-point Likert scale (1= Not at all, 7= Very), higher scores on this scale corresponds with a higher level of normativeness of the statement. The scale for norms prior to the reaction had a cronbach`s alpha of $\alpha = 0.847$ and norms post $\alpha = 0.886$, indicating good internal consistency.

We measured Plausible deniability using a self-constructed measure consisting of two statements: "To what extend do you think Paul expressed his true beliefs?" and "Paul his remark was intentionally sexist/harmful.". Their response was measured via a 7-point Likert scale (1= Strongly disagree, 7= Strongly agree), higher scores on this scale corresponds with a lower level of perceived plausible deniability. The spearman correlation was .395, with only two items this indicated that the individual items signal the same direction so we were able to combine the items for our variable.

Recognition of sexism was measured using a self-constructed measure with three scale items asking to what extend there was recognition of sexism after the reaction to the joke was given by the people in the video. We constructed the following statements to measure this variable: "The response caused Paul to become aware or recognize that his behavior was sexist.", "The response let Paul know that sexism is unacceptable.", "In future conversations within this group Paul will likely make similar comments again". Their response was measured via a 7-point Likert scale (1= Strongly disagree, 7= Strongly agree), higher scores on this scale corresponds with a higher level of recognition of sexism except for the last statement so this item was reverse coded to match the other questions. The Cronbach's alpha level was .748, indicating acceptable internal consistency.

The rest of the variables that were asked but not used for this paper where: warmth & competence, social standing, feminism, disapproval and social context. To confirm the effectiveness of our manipulation, a separate scale was employed to check if the remarks and the reactions were perceived the way it was intended. We will compare if the mean for questions are as expected for the conditions the participants were in. This means that for example the question "Tom confronted/objected to the remark made by Paul" should be scored higher in the confrontation condition than in the control condition. Next to this question we had five other questions evaluating the effectiveness of our manipulation.

Analysis Plan

The analysis was conducted using the statistical software IBM SPSS 27. H1 was tested with a regression analyses (using Model 1form Hayes Process Macro) to examine the differences in norm post scores between the confrontation and control groups while controlling for pre scores for norms and using plausible deniability as a moderator. For H2 a regression analysis (again using Model 1 from Hayes Process Marco) was conducted to examine the moderating role of plausible deniability on the relationship between response condition and recognition of sexism. Response condition was categorical and for norms, recognition of sexism and plausible deniability we used the scale mean and calculated them before the analysis. We used the means to examine the main effects of condition and the interaction effect between the condition and norms/recognition of sexism, as mentioned in the hypothesis above. Perceived plausible deniability was considered as a moderator in these analyses. Data preprocessing included anonymizing the data, as well as excluding the participants who did not complete the whole questionnaire or were exposed to the humor condition (not used for the current research question). The significance level used to determine statistical significance is .05.

Results

For the current study, two different hypotheses were tested (H1 and H2). We looked at two different response conditions (confrontation and control) which were combined of male and female responders. This was because there was no difference between the gender conditions within the response conditions. We examined this difference with an independent samples t-test to compare the difference between genders within the response conditions for the two dependent variables. For norms post we found no significant difference $t(82) = -.21$, $p = .84$ in scores for man ($M = 1.87$, $SD = 1.39$) and woman ($M = 1.93$, $SD = 1.21$). The magnitude of the differences in the means was small ($M_{difference} = -0.06$, 95% Cl [-0.62, 0.51]). Recognition of sexism also had no significant difference $t(82) = -1.21$, $p = .23$ in scores for man ($M = 3.84$, $SD = 0.87$) and woman ($M = 4.07$, $SD = 0.87$). The magnitude of the differences in the means was also small ($M_{difference} = -0.23$, 95% Cl [-0.61, 0.15]). Because there is no significant difference, from now on when we refer to condition, we refer to reaction condition only (not gender condition). Both conditions had the same number of participants ($N = 42$). Descriptive statistics of all variables used are in appendix A. Correlations between all variables are shown in appendix B.

Manipulation Check

We conducted a manipulation check to see if the remarks and the reactions were perceived the way they were intended by the participants. We compared the means of the questions we created for the manipulation check, to determine if the scores for specific questions aligned with our expectations for the corresponding conditions in which participants were placed. This confirmed the effectiveness of our manipulation. For instance, in order to assess the impact of the confrontation condition, we asked "Tom confronted/objected to the remark made by Paul". The results revealed the highest mean response in the 'confrontation' condition ($M_{confrontation} = 5.58$), while the response for the other condition showed a lower mean ($M_{control} = 3.22$) .

Assumptions

The assumptions for the hypothesis are as follows. The normality of the residuals was assessed through a Shapiro-Wilk test. H1: the Shapiro-Wilk tests indicated a significant deviation from normality in both different groups (confronting condition $W(42) = 0.30$, $p <$ .001 and control condition $W(42) = .26$, $p < .001$). With both p-values below .05 this assumption is violated. Previous research has shown that ANCOVA is robust to violations of normality or heteroscedasticity (Olejnik & Algina, 1984). Therefore, despite the violated assumption of normality, we proceeded with the ANCOVA analysis and interpreted the results. H2: normality was not violated the Shapiro-Wilk test showed p values above .05 (confrontation condition $W(42) = 0.96$, $p = .18$, and control condition $W(42) = 0.97$, $p = .33$). Linearity was examined through a scatterplot, which showed a linear relationship between the dependent variable and covariate in the conditions (H1: appendix C). The scatterplot also suggested that the regression slopes were approximately equal, and an F-test indicated no interaction between the condition and covariate pre norms (H1: $F(1,80) = 1.30$, $p = .258$). For H2 there was also linearity ($F(1,124) = 6.97$, $p = .01$). The assumption of homogeneity of variance was tested using Levene's test, which revealed no significant difference in variance between the groups (H1: $F(1,82) = 0.60$, $p = .44$ and H2: $F(2,124) = 1.73$, $p = .18$) indicating that this assumption was met. The linear association between the covariate and the dependent variable were examined through a scatterplot and Pearson's correlation coefficient. A significant positive correlation was found between the covariate and the dependent variable see appendix B, satisfying the last assumption.

Hypothesis 1

A regression analysis using Model 1 from Hayes Process Marco was conducted to examine the moderating role of Plausible Deniability on the effect of condition on the Norms Post variable, while controlling for the covariate Norms Pre variable.

Condition significantly influenced norms post (β = 1.94, 95% *Cl* [0.52, 3.36], *t* = 2.72, *p* = .008), indicating a significant main effect where confrontation has higher normality in comparison to control. Plausible deniability had also had a significant main effect on norms post (β = 0.28, 95% *Cl* [0.06, 0.50], *t* = 2.51, *p* = .014), indicating higher plausible deniability was related to higher norm scores. Furthermore, the overall model was significant, suggesting condition and Norms post were significantly moderated by plausible deniability while controlled for norms pre ($F(4,79)$ = 12.29, *p* < .001). The model explained 38.36% of the variance in norms post (R^2 = .38). The negative interaction effect *(β = - 0.46, 95% *Cl* [- 0.752, -0.16], *t* = - 3.05, *p* = .003) indicates that there is an decrease in the magnitude of the effect of condition on the post norm score as the plausible deniability value increases (less perceived plausible deniability), while controlling for the pre norms variable. This is not in line with our hypothesis.

Post hoc analyses examined the conditional effects of condition at different levels of plausible deniability. At a plausible deniability value of 2.3 (high perceived plausible deniability), the effect of condition on norms was significant (β = 0.89, *SE* = .40, *t* = 2.21, *p* = .03). The direction of the effect was positive meaning higher scores of normality where reported when there was a confrontation than without. For plausible deniability value of 4.5 (middle), the effect of condition on norms post was not significant (β = - 0.11, *SE* = .23, *t* = - 0.50, *p* = .62). Indicating no significant difference in scores for norms between conditions. At plausible deniability value 6.00 (low), the effects of condition on norms post was significant (β = - 0.80, *SE* = .32, *t* = - 2.49, *p* = .015). This time we see a negative effect meaning norms were higher in the control condition than the confrontation condition when there is low perceived plausible deniability. The results of the post hoc analysis are in line with our hypothesis. See appendix D for visualization of the results.

Hypothesis 2

A regression analysis using Model 1 from Hayes Process Marco was conducted to examine the moderating role of plausible deniability on the relationship between condition and recognition of sexism. The analysis included condition, plausible deniability, and their interaction term as predictors.

The results indicated that the main effects of condition and plausible deniability were not significant (condition: β = - 0.77, 95% Cl [- 1.89, 0.36], t = - 1.36, p = .177 and plausible deniability: β = - 0.13, 95% Cl [- 0.30, 0.04], t = - 1.56, p = .12). However, a significant interaction effect between condition and plausible deniability was found ($F(3,80)$ = 4.47, p = .006). The model explained 14.36% of the variance in norms post (R^2 = .14). The positive interaction effect $(\beta$ = 0.28, 95% Cl [0.05, 0.52], t = 2.38, p = .02) indicates that there is an increase in the magnitude of the effect of condition on the recognition of sexism score as the plausible deniability value increases (less perceived plausible deniability). This is in line with our second hypothesis.

To further explore the interaction effect, conditional effects of the predictor (condition) at different values of the moderator (plausible deniability) were examined. At a plausible deniability value of 2.3 (high perceived plausible deniability), the effect of condition on recognition of sexism was not significant (β = - 0.12, SE = .32, t = - 0.37, p = .71, 95% CI [- 0.75; 0.52]). However, at plausible deniability values of 4.50 and 6.00 (middle and lower plausible deniability), significant positives effects of condition on recognition of sexism were observed (β = .50, SE = .18, t = 2.77, p = .007, 95% CI [0.14; 0.87] and β = .93, SE = .25, t = 3.65, p < .001, 95% CI [0.42; 1.43]). These findings suggest that the relationship between condition and recognition of sexism is influenced by the level of plausible deniability, with higher values of plausible deniability (lower perceived plausible deniability) strengthening the relation between recognition and condition. This relation entails higher recognition of sexism

in the confrontation condition compared to the control condition. The results of the post hoc analysis are in line with our hypothesis. See appendix E for a visualization.

Discussion

The current study aimed to examine the potential influence of plausible deniability on the relationship between condition, norms change, and recognition of sexism. The research question was divided into two hypotheses, which are discussed below along with their findings.

Hypothesis 1 suggested that perceived plausible deniability would act as a shield against confrontation, thereby not affecting the extent to which individuals adapt their norms in response to the condition they saw. The regression analysis revealed a significant negative interaction effect, indicates that there is an decrease in the magnitude of the effect of condition on the post norm score as the plausible deniability value increases (less perceived plausible deniability). With a significant opposite effect when recognition of plausible deniability was high or low. This partly supports H1; when perceived plausible deniability is low the participants perceived the sexist joke as less normal in the confront condition compared to the control condition, meaning they changed their norms correctly. Interestingly, a significant opposite effect was found when participants had high perceived plausible deniability, with the joke being perceived as more normal in the confrontation condition and less normal in the control condition. We originally expected there to be no significant effect here, but the finding still supports the thought that with high perceived plausible deniability the participant does not change their behavior in the way the confrontation is intended.

It is conceivable that, in the case of high plausible deniability, a confrontation might actually lead to a different effect because the offender seeks self-protection, resulting in perceiving their own remark as more acceptable. This can be linked to the cognitive dissonance theory (Festinger, 1957) mentioned earlier in the introduction. Confrontation

potentially creates a dissonance, which is then resolved by the individual adjusting their collective norms to reassure themselves that they are not wrong and avoid engaging in the confrontation. Further research is needed to explore this phenomenon, to draw conclusions regarding the underlying reasons.

Hypothesis 2 proposed that plausible deniability moderates the effect of the condition on recognition of sexism. Plausible deniability acted as a shield against confrontation, influencing whether the confrontation resulted in an increase in recognition of sexism. The overall model was significant. When we looked at high perceived plausible deniability there was no significant effect of the conditions on recognition of sexism. In cases where there was low perceived plausible deniability, a significant positive effect was observed. Meaning there was a higher level of recognition of sexism following confrontation compared to the control condition. This shows us that H2 was supported.

The results highlight that plausible deniability acts as a shield against confrontation, potentially differentiating the impact of confrontation on sexism. Previous research has focused on how to approach confrontation effectively (Burns & Granz, 2020; Drury & Kaiser, 2014; Dray & Sabat, 2022), but little attention has been given to how confrontation is received and the reasons behind it. By investigating the factors that hinder the desired outcomes of confrontation, new strategies can be developed to combat sexism effectively. The current study demonstrates that plausible deniability may be a reason why confrontation does not always have the desired effects. This aligns with previous theories (Crandall & Eshelmans, 2003; Berry & Ferguson, 2004) indicating that jokes provide a common way of expressing sexist beliefs because their content can be denied afterward. However, these articles also caution that this concept can unwittingly influence norms regarding sexism because of reduced frequency of confronting such jokes. Therefore, while combating sexist jokes can be challenging especially when there is highly perceived plausible deniability, it is

crucial for the other person. Further research is needed to explore the barriers more comprehensively, allowing for the development of new approaches to combat sexism effectively.

Despite shedding light on the complexity of addressing sexism beyond the confrontation itself, the study could have several limitations. Firstly, the scales used in the research, except for the norms scale, were created by ourselves and therefore it is uncertain the intended constructs where measured. Furthermore, the scales consisted of few items, which may limit their reliability (Czerwiński & Atroszko, 2021). It is also uncertain whether participants could fully immerse themselves in the presented scenarios. Although efforts were made, the actors in the videos were not professionals, potentially diminishing the authenticity of the situations and influencing the responses. This would not have implications on the results, but would make the results less generalizable to real life conditions. Due to the influence of social desirability, individuals tend to strive for consistency. This could have implications on the questions that were asked twice (before and after the reaction). Despite this phenomenon in this research, where participants were given the opportunity to recall their answers, there was a noticeable difference in scores of similar questions before and after the different conditions. Which in turn suggests a high likelihood of replicating the effect in future studies.

Limitations could also arise from the use of convenience sampling, particularly through snowballing and the use of SONA. This approach may restrict the results to specific populations, as the study primarily focused on the Western population, predominantly comprising educated men. Different results may be obtained from other societal groups, making it difficult to generalize the current findings. Despite these implications, we have still chosen to proceed in this manner as it represents the only viable option given the available resources and guidelines. The study reveals an effect in the tested population, this could be an

indication to a generalizable result in future research. Lastly, the sample size fell short of the predetermined minimum of 260 participants, with only 127 participants' responses being utilized. This reduced the power of the study below the desired level, meaning there is a limited chance of detecting an effect if it truly exists. Nevertheless, there was still a significant result suggesting that the effect is likely strong enough to be detected even with limited power, which may indicate a substantial effect in the population. However there is also an increased chance of false-positive results (finding an effect when it doesn't actually exist), so a replication with lager power is advised.

Despite these limitations, the study provides indications of an effect of plausible deniability on the relationship between confrontation, norms change, and recognition of sexism. Further research is warranted to confirm the role of plausible deniability in the effectiveness of confrontation and explore appropriate ways to address this phenomenon. This necessitates a larger and more diverse sample that encompasses different cultures and levels of education to enhance the generalizability of the results. Additionally, validated scales that effectively measure the variables should be employed. The video stimuli, although helpful in understanding the perspective of the joke-maker, may benefit from longer videos to provide deeper insight. Also future research could look into the potentially counterproductive effect confrontation has on norms when there is high plausible deniability. In the current study, this effect was found; however, a scientific explanation still needs to be determined.

When examining Johan Derksen's response to the confrontation he faced regarding his remark about women's football mentioned in the introduction, a clear instance of high perceived plausible deniability is observed. Following a confrontation by the editor-in-chief of Libelle, Hilmar Mulder, who stated, "We have moved with the times. Johan Derksen cannot say the same," Derksen responded by accusing Mulder of immaturity, stating, "Yes, that was a rather immature reaction for someone who calls themselves the editor-in-chief of

Libelle" (Vandaag Inside, 2021). Derksen deflects the confrontation by criticizing Mulder and implying that she is the one in the wrong, creating a barrier for addressing the initial confrontation. When the role of plausible deniability in the reception of confrontation is better comprehended, further research can focus on developing effective strategies to address sexism. Ultimately, this research aims to prevent sexist jokes by tackling them more effectively. To eventually even make the presenters of 'voetbal insight' realize that what their saying is not 'just a joke'.

References

1st-year SONA-practicum Pool / 1e-jaars SONA-practicum Pool (2022-2023). (z.d.).

https://rug.sonasystems.com/Default.aspx?ReturnUrl=%2fexp_info.aspx%3fexperime

nt_id%3d510&experiment_id=510

Ashburn-Nardo, L., Morris, K. A., & Goodwin, S. A. (2008). The Confronting Prejudiced

Responses (CPR) Model: Applying CPR in Organizations. *Academy of Management

Learning and Education*, *7*(3), 332–342. https://doi.org/10.5465/amle.2008.34251671

Berry, C., & Ferguson, M. K. (2004). Social Consequences of Disparagement Humor: A

Prejudiced Norm Theory. *Personality and Social Psychology Review*, *8*(1), 79–94.

https://doi.org/10.1207/s15327957pspr0801_4

Burns, M. D., & Granz, E. L. (2020). Confronting Sexism: Promoting Confrontation

Acceptance and Reducing Stereotyping through Stereotype Framing. *Sex Roles*, *84*(9–

10), 503–521. https://doi.org/10.1007/s11199-020-01183-5

Crandall, C. S., & Eshleman, A. J. (2003). A justification-suppression model of the

expression and experience of prejudice. *Psychological Bulletin*, *129*(3), 414–446.

https://doi.org/10.1037/0033-2909.129.3.414

Czerwiński, S. K., & Atroszko, P. (2021). A solution for factorial validity testing of three-

item scales: An example of tau-equivalent strict measurement invariance of three-item

loneliness scale. *Current Psychology*, *42*(2), 1652–1664.

https://doi.org/10.1007/s12144-021-01554-5

Czopp, A. M., & Monteith, M. J. (2003). Confronting Prejudice (Literally): Reactions to

Confrontations of Racial and Gender Bias. *Personality and Social Psychology

Bulletin*, *29*(4), 532–544. https://doi.org/10.1177/0146167202250923

Derksen en Gijp reageren in Voetbal Inside op kritiek - Joop - BNNVARA. (z.d.). Joop.

https://www.bnnvara.nl/joop/artikelen/derksen-en-gijp-reageren-voetbal-inside-op-

kritiek

DPG Media Privacy Gate. (z.d.). https://www.volkskrant.nl/nieuws-achtergrond/twijfels-

over-%20cultuurverandering-bij-amsterdamse-corps-na-uitspraken-over-hoeren-en-

sperma-emmers~b869fadb/?referrer=https%3A%2F%2Fwww.google.nl%2F

Dray, K. K., & Sabat, I. E. (2022). Confronting sexism: Identifying dimensions and exploring

impact. *Journal of Applied Social Psychology*, *52*(5), 316–340.

https://doi.org/10.1111/jasp.12861

Drury, B., & Kaiser, C. R. (2014). Allies against Sexism: The Role of Men in Confronting

Sexism. *Journal of Social Issues*, *70*(4), 637–652. https://doi.org/10.1111/josi.12083

Festinger, L. (1957). A Theory of Cognitive Dissonance. In *Stanford University Press eBooks*.

https://doi.org/10.1515/9781503620766

Fitzgerald, L., Shullman, S. L., Bailey, N., Richards, M. M., Swecker, J., Gold, Y., Ormerod,

M., & Weitzman, L. M. (1988). The incidence and dimensions of sexual harassment in

academia and the workplace. *Journal of Vocational Behavior*, *32*(2), 152–175.

https://doi.org/10.1016/0001-8791(88)90012-7

Gervais, S. J., & Hillard, A. L. (2014). Confronting Sexism as Persuasion: Effects of a

Confrontation's Recipient, Source, Message, and Context. *Journal of Social Issues*,

70(4), 653–667. https://doi.org/10.1111/josi.12084

Hyers, L. L. (2007). Resisting Prejudice Every Day: Exploring Women's Assertive Responses

to Anti-Black Racism, Anti-Semitism, Heterosexism, and Sexism. *Sex Roles*, *56*(1–2),

1–12. https://doi.org/10.1007/s11199-006-9142-8

Johan en René onder vuur om bekritiseren vrouwelijke analisten. (2021, 19 juni). RTL

Boulevard.https://www.rtlboulevard.nl/entertainment/showbizz/artikel/5237232/johan-

rene-onder-vuur-bekritiseren-vrouwen

Koudenburg, N., Kannegieter, A., Postmes, T., & Kashima, Y. (2020). The subtle spreading

of sexist norms. *Group Processes & Intergroup Relations*, *24*(8), 1467–1485.

https://doi.org/10.1177/1368430220961838

Kowalski, R. M. (1996). Complaints and complaining: Functions, antecedents, and

consequences. *Psychological Bulletin*, *119*(2), 179–196. https://doi.org/10.1037/0033-

2909.119.2.179

Mallett, R. K., Berry, C., & Woodzicka, J. A. (2016). What Did He Mean by that? Humor

Decreases Attributions of Sexism and Confrontation of Sexist Jokes. *Sex Roles*, *75*(5–

6), 272–284. https://doi.org/10.1007/s11199-016-0605-2

Mallett, R. K., Berry, C., & Woodzicka, J. A. (2019). Ignoring sexism increases women's

tolerance of sexual harassment. *Self and Identity*, *20*(7), 913–929.

https://doi.org/10.1080/15298868.2019.1678519

Mallett, R. K., & Wagner, D. E. (2011). The unexpectedly positive consequences of

confronting sexism. *Journal of Experimental Social Psychology*, *47*(1), 215–220.

https://doi.org/10.1016/j.jesp.2010.10.001

Ministers klagen over seksisme Rutte - Joop - BNNVARA. (n.d.). Joop.

https://www.bnnvara.nl/joop/artikelen/ministers-klagen-over-seksisme-rutte

Olejnik, S., & Algina, J. (1984). Parametric ANCOVA and the Rank Transform ANCOVA

When the Data are Conditionally Non-Normal and Heteroscedastic. *Journal of

educational statistics*, *9*(2), 129–149. https://doi.org/10.3102/10769986009002129

Parker, L. M., Monteith, M. J., Moss-Racusin, C. A., & Van Camp, A. R. (2018). Promoting

concern about gender bias with evidence-based confrontation. *Journal of

Experimental Social Psychology*, *74*, 8–23. https://doi.org/10.1016/j.jesp.2017.07.009

Rasinski, H. M., Geers, A. L., & Czopp, A. M. (2013). "I Guess What He Said Wasn't That

Bad". *Personality and Social Psychology Bulletin*, *39*(7), 856–869.

https://doi.org/10.1177/0146167213484769

Swim, J. K., Mallett, R. K., Russo-Devosa, Y., & Stangor, C. (2005). Judgments of Sexism: A

Comparison of the Subtlety of Sexism Measures and Sources of Variability in

Judgments of Sexism. *Psychology of Women Quarterly*, *29*(4), 406–411.

https://doi.org/10.1111/j.1471-6402.2005.00240.x

Vandaag Inside. (2021, 19 juni). *Johan krijgt sneer van hoofdredactrice Libelle | DE

ORANJEZOMER* [Video]. YouTube.

https://www.youtube.com/watch?v=5yak0NMIYuo

Woodzicka, J. A., Mallett, R. K., Hendricks, S., & Pruitt, A. V. (2015). It's just a (sexist)

joke: comparing reactions to sexist versus racist communications. *Humor:

International Journal of Humor Research*, *28*(2). https://doi.org/10.1515/humor-2015-

0025

Appendix A

Means and standard deviations of all variables per condition

Condition		Norms pre	Norms post	Recognition of sexism	Plausible deniability
Control	Mean	2,0179	1,8929	3,7063	4,7024
	Std. Deviation	1,24987	1,36053	,97448	1,51842
Confrontation	Mean	2,2857	1,9048	4,2063	4,2976
	Std. Deviation	1,50464	1,24199	,69044	1,55414
Total	Mean	2,1518	1,8988	3,9563	4,5000
	Std. Deviation	1,38136	1,29475	,87625	1,54061

*Note*: all ratings were on a 7-point scale ranging from 1 = *not at all* to 7 = *very*

Appendix B

Correlations

		Condition	Norms pre	Norms post	Rec sexis	Plau den
Condition	Pearson Correlation	1	,098	,005	,287**	-,132
	Sig. (2-tailed)		,377	,967	,008	,231
Norms pre	Pearson Correlation	,098	1	,553**	,024	-,338**
	Sig. (2-tailed)	,377		,000	,830	,002
Norms post	Pearson Correlation	,005	,553**	1	-,107	-,137
	Sig. (2-tailed)	,967	,000		,335	,213
Rec sexis	Pearson Correlation	,287**	,024	-,107	1	-,016
	Sig. (2-tailed)	,008	,830	,335		,883
Plau den	Pearson Correlation	-,132	-,338**	-,137	-,016	1
	Sig. (2-tailed)	,231	,002	,213	,883	

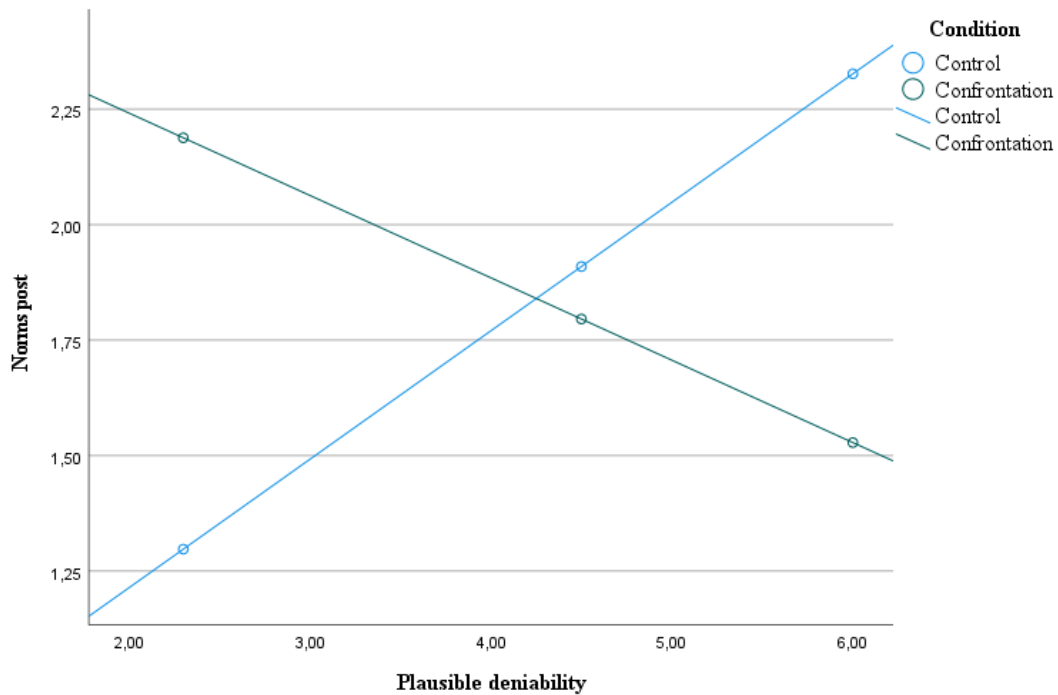**. Correlation is significant at the 0.01 level (2-tailed).

Appendix C

Linear relationship between the dependent variable and covariate in all conditions
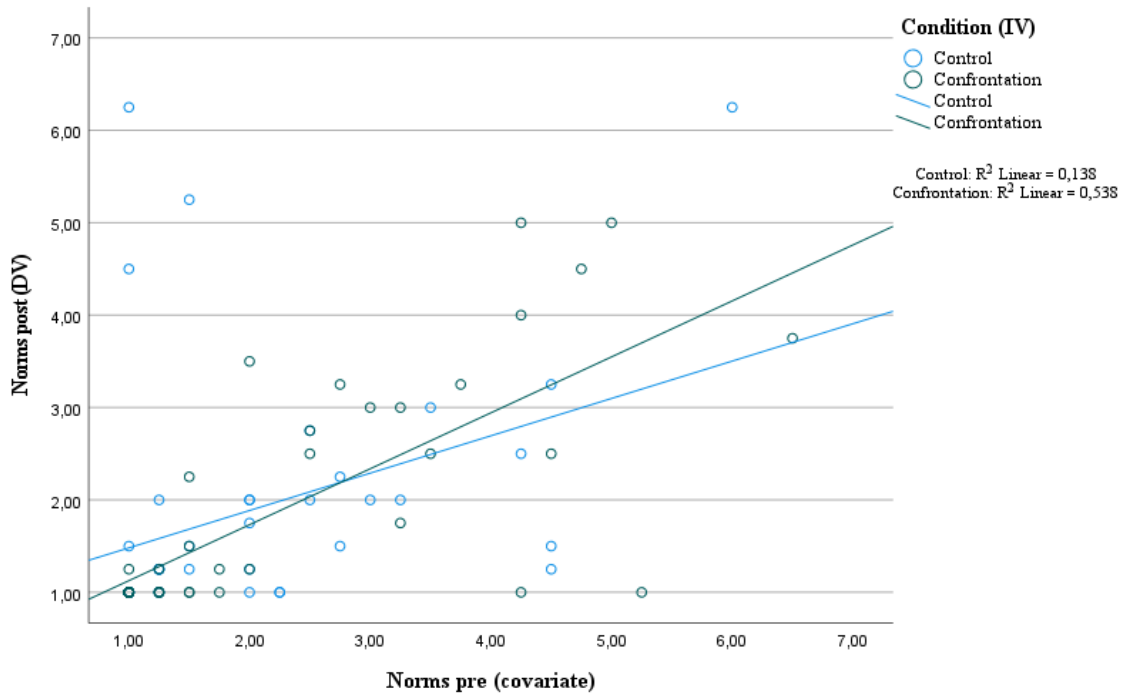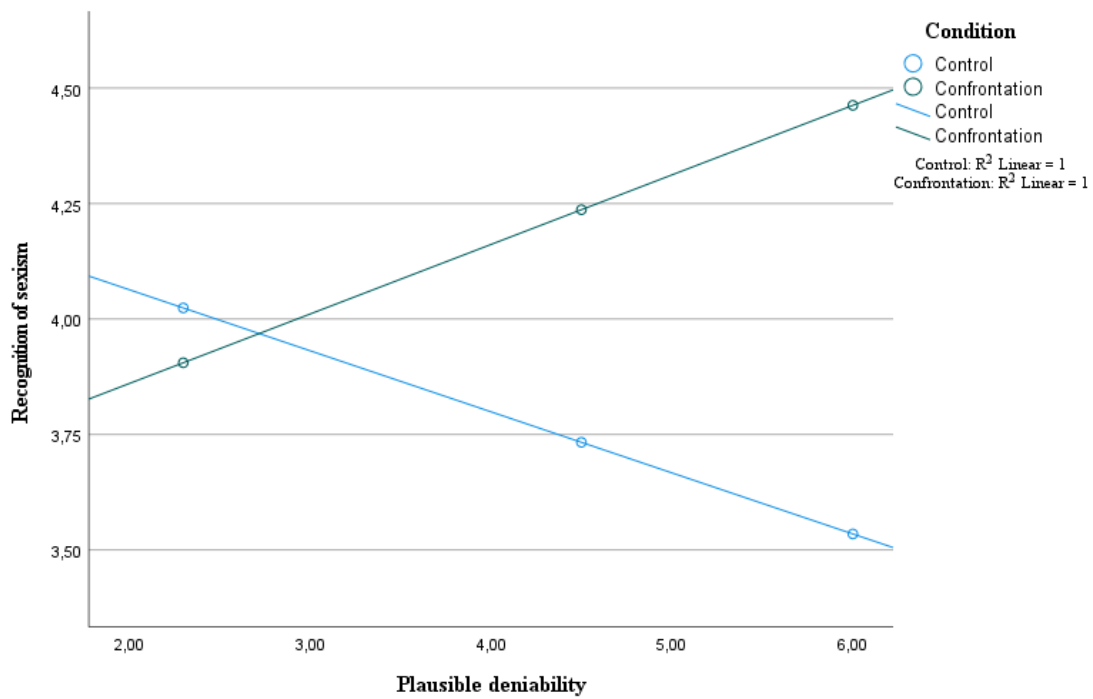

Appendix D

Interaction effect condition and norms with moderator Plausible deniability

Note: confrontation condition = 0, humor condition = 1 and control condition = 2

Appendix E

Interaction effect condition and norms with moderator Plausible deniability

Note: confrontation condition = 0, humor condition = 1 and control condition = 2