ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis - IBEB

A QUANTITATIVE ANALYSIS OF THE GENDER CREDIBILITY - AUTHORITY GAP

Name student: Clara Martinez
Student ID number: 595841

Supervisor: Dr. Anne Boring
Second assessor: Dr. Josse Delfgaauw

Date final version: 9th of July 2024

Abstract

This study delves into the persistent gender authority and credibility gaps, the disparity in perceived trustworthiness when a woman makes a statement compared to a man. With the use of an experimental study, the paper aims to quantify this disparity. A survey, shared using convenience and snowball sampling methods, is used to collect data. Respondents are quasi-randomly assigned to videos featuring either a man or a woman who both present the same five factual statements. Individuals are then asked to rate the likelihood of each statement to be true. The study finds that, on average, statements made by women are trusted more than those made by men. The research additionally explores how this gap varies depending on the characteristics of the respondents. The latter include age, gender, political affiliation, and education level. It is hypothesised that the disparity: diminishes with age, is consistent across genders, is less pronounced or inverse among right-wing individuals, increases with higher education levels, and varies with the gender association of the subjects discussed. The findings reveal that older and right-wing respondents on average trusted the woman presenter less. However, gender and education level interactions with the treatment did not show statistically significant differences, indicating a need for further investigation. The study also finds that subjects associated with femininity result in higher trust for female presenters, although this difference was not statistically significant. Although the findings are insightful, limitations in the external validity of the analysis such as potential biases from the survey method and the sample's representativeness arise. The methodology nonetheless could be reused for additional research.

Keywords: gender authority gap, gender credibility gap, stereotypes

Table of Contents

1. Introduction

Women have actively advocated for and secured various rights, including the right to work, earn an income, and manage finances, which were previously inaccessible to them. Despite these advancements, disparities persist, notably in the perceived authority and credibility associated with women compared to men. Referred to as the gender authority gap (European Commission, 1998), this disparity often manifests not only in the numerical difference between men and women holding positions of authority but extends to the imbalance of trust and credibility attributed to their voices. This affects everyday interactions where women's voices are often undervalued compared to men's (Sieghart, 2021). The significance of this research lies in the broader societal and economic consequences of this bias, which contribute to gender disparities in multiple spheres, including economic inequality.

Understanding the gender authority gap is crucial because of its profound implications for women's participation and confidence in both professional and personal contexts. Research has shown that the Pygmalion effect, where low expectations can lead to reduced performance and self-efficacy, disproportionately affects women in the workforce (Latu et al., 2015). They argue that biases exhibited by interviewers can impact not only their perceptions of women candidates but also the interviewees themselves, revealing the role of self-fulfilling prophecies. Indeed, in situations where interviewers harbor biases, women interviewees may experience heightened stress, distraction, and lower confidence which can further exacerbate the challenges they face in the employment process. This psychological phenomenon, coupled with confirmation bias, which reinforces pre-existing stereotypes, can result in women being less confident and thus perpetuating disparities in earnings (Kay et al., 2014; Sterling et al., 2020). It stands to reason that when women underestimate their capabilities, they may gravitate towards lower-paying employment opportunities and refrain from negotiating for higher salaries, whereas men, with a tendency to overestimate their competencies, are more inclined to pursue higher-paying positions and advocate for salary increases. Disparities such as the gender authority gap are hence not merely academic but have tangible effects on women's career trajectories and economic opportunities.

The existing literature primarily examines the gender authority gap in terms of numerical representation in positions of power. Blommaert et al. (2020) focus on the relationship between social networks and job authority, finding that women's networks are less diverse and have fewer managerial connections than men's. This lack of access to social capital inhibits women's progress into roles of authority. Stojmenovska et al. (2021) expand on this by differentiating between types of authority positions, showing that the gap is wider in roles involving significant control over organisational resources, thus reinforcing gender biases in decision-making contexts. Though interesting to study, this interpretation of the term is as previously mentioned, limited.

Marry-Ann Sieghart in her book 'The Authority Gap: Why women are still taken less seriously than men, and what we can do about it' (2021) however extends the interpretation. She highlights the skepticism or resistance women face when asserting their authority or expertise, even when they possess the necessary qualifications and experience. To do so, she presents research findings, case studies, and personal anecdotes, she also presents the experiences of women who have had to face and challenge these biases and stereotypes.

This more comprehensive interpretation is closely related to the gender credibility gap. An important part of the literature that studies the gender credibility gap, focuses on work environments and relies heavily on student participants. Wissmath et al. (2008) for instance examined the perceived credibility of newscasters in Switzerland finding that although women's statements were often rated as more credible, female newscasters themselves were perceived as less credible. Similarly, Armstrong and McAdam (2009) found that male-authored blogs are seen as more credible, especially among information seekers. Andsager and Mastin (2003) however, while exploring the credibility of political columnists, find no significant gender differences overall.

The paper expands the current literature by quantifying this discrepancy in value people attach to men's and women's voices outside of workspaces and not only focusing on students. It utilises the more comprehensive definition of the gender authority gap, which overlaps with the concept of the gender credibility gap, by not differentiating between perceived authority and credibility but instead asking for people's trust, thereby averaging perceptions of both. The research furthermore tries to show that it does not

only apply to women in authoritative positions but to anything any woman ever says. This research is particularly important because it highlights the pervasive nature of gender biases that affect women's perceived credibility and authority. By understanding these biases, we can better address the root causes of the gender authority gap and its broader economic implications, such as the gender pay gap.

The question of focus is hence the following: *how does the gender of the speaker influence the level of trust attributed to their statements?*

To test the hypothesis that there is a gender credibility-authority gap, the study utilises an experimental study. In a survey, participants are quasi-randomly assigned to the treatment and control groups. They are respectively shown videos featuring either Charlotte, a woman, or Leo, a man, presenting the same five facts. After having watched the videos, participants are asked to rate out of 10 how likely they believe each statement to be true. This method allows for a simple with-and-without comparison to uncover the average treatment effect on the treated (hereinafter: ATT). Tests of the effectiveness of the quasi-randomization, in Table 1, show proper balance for all observed characteristics except the level of study attained. The latter is hence added to all models to check variability.

The results of the baseline model, found in Table 2, allowed us to quantify the difference in trust people have for women compared to men. They show that, on average, being assigned to the treatment group, and being shown videos of Charlotte, is associated with a significant increase in the rated likelihood of the statement to be true by 0.690 units compared to the control group (p-value<0.01). The paper hence aims to understand if this differs depending on other variables. Detailed more in the literature review part, the following hypotheses are formulated based on existing studies (Armstrong & McAdams, 2009; Brough et al., 2016; Coffe, et al., 2023; Coffman, 2014; Croson & Buchan, 1999; Enloe, 2000; Hammond-Thrasher & Järvikivi, 2023; Radvansky et al., 2008; Radvansky et al., 2010; Stern & Axt, 2021; Garaigordobil, 2013).

Hypothesis 1: The found effect is reduced as participants get older.

Hypothesis 2: The effect does not differ depending on the gender of the respondents.

Hypothesis 3: The effect is reduced or inversed for conservative (right-wing) individuals but not for liberal (left-wing) or centrist individuals.

Hypothesis 4: The effect is smaller when the level of education of the respondent is lower. Hypothesis 5: The effect is larger when the subjects discussed are likely to be unconsciously considered more feminine compared to masculine.

To this end, multiple different models of regression analysis were conducted. Interaction terms between treatment and observed variables of interest can be found in Table 3. The results of the considered more masculine or feminine statements appear in Table 2 and Figure 1, along with the baseline model. Lastly, the regressions of simple with-and-without comparison of each statement made are included in Table 4.

Interaction terms, found in Table 3, expose how the disparity varies depending on observed characteristics. Model 1 of Table 3, reveals that, as respondents get older, the disparity between Leo and Charlotte, in rated likelihood of the statements to be true gets reduced. The association of age and average trust for untreated individuals is -0.025 units per year, statistically significant at level 10%. Additionally, the interaction term, -0.006, although not significant, means that each additional year of age reduces the average trust score by 0.031 units for treated individuals. Model 2 indicates that the average trust for women who are not treated is statistically insignificantly lower than untreated men by -0.009. Treatment however is associated with a significant ($p<0.1$) reduction in trust score of 0.9. Non-binary respondents in the control group exhibited significantly higher trust scores than men, although the small sample size warrants caution. Lastly, the positive statistically significant (p-value<0.01) coefficient of the variable treated indicated that men in the treatment group have an average higher trust score of 1.267 compared to men in the control group. Model 3 exhibits the interaction of political affiliation with treatment. It highlights that compared to centrist individuals in the control group, the ones in the treatment group are associated with a significantly higher trust score of 0.964 (p-value<0.05). In the control group, politically right individuals have an insignificantly higher trust score of 0.453 compared to centrists, while left-leaning individuals exhibit an insignificantly lower trust score of -0.159. In contrast, within the treatment group, right-wing individuals experience a statistically significant large decrease in trust score of 2.674 ($p < 0.05$), whereas left-leaning individuals show an insignificant minor reduction in trust score of -0.164. Education level interactions, found in Model 4, also show disparities in average trust scores however most coefficients are insignificant.

The paper additionally uncovers that, statements linked to feminine traits, when presented by Charlotte, are more likely to be deemed true compared to when presented by Leo, with an average increase of 0.735 in trust scores ($p<0.01$). Conversely, for statements associated with masculinity, being part of the treatment group is associated with a statistically insignificant increase of 0.566 in trust score compared to being in the control group. Finally, specific regression analyses for each statement in Table 4 and Figure2 show positive coefficients for all, ranging from 0.375 to 1.079, with some statements being statistically significant at the 1% and 5% levels. Controlling for observed characteristics does not substantially alter any of these results.

2. Related Literature

The gender authority gap is, as explained earlier, often studied only as a representation disparity. Referring then to the numerical difference between how many men and women hold positions of authority.

For instance, Blommaert et al. (2020) examine the role of social network resources in explaining the gap in job authority between men and women. Utilising data from the Netherlands, the authors employ regression analysis on survey data to understand how access to managerial and diverse networks, which are crucial for job authority, differ depending on one's gender. The study finds that women's networks are less diverse and contain fewer managerial connections compared to men's, which partially explains the gender gap in job authority. Their research highlights the importance of social networks in perpetuating the gender authority gap and suggests enhancing women's access to diverse and managerial networks could help this disparity. Throughout the study, the term 'position of authority' is defined by the authors as a job that requires the supervision of others. They however admit that the latter assumption is a limitation of their study.

In their study, Stojmenovska et al. (2021) use a more comprehensive definition. They define authority positions based on the amount of authority (supervisory vs. non-supervisory) and the type of control over resources (human vs. organisational). They also focus on the Netherlands, using a large dataset of Dutch employees from the National Survey of Working Conditions, linked to administrative records. The researchers found

that the gender gap is more pronounced in positions with higher levels of supervisory tasks and control over organisational resources, such as budgets, compared to those involving control over human resources, like hiring decisions. It is additionally more pronounced in formal authority positions, officially recognised leadership roles, than in informal ones, where influence comes from expertise or relationships. The paper provides a nuanced understanding of how the gender gap in authority varies across different types of positions.

Although interesting to study and assess, the definition of the authority gap used by Stojmenovska et al. (2021) and Blommaert et al. (2020) is limited. A gender gap is defined by the European Commission (1998) as a "gap in any area between women and men in terms of their levels of participation, access, rights, remuneration or benefits." Additionally, authority extends beyond hierarchical arrangements. It also encompasses the recognition of an individual's credibility, mastery, or commanding presence, which allows individuals to influence and gain respect from others. Hence limiting its analysis merely to numerical disparities is insufficient as it fails to capture the full spectrum of gender-based biases and their broader implications. The current research hence contributes to this field by exploring the perception gap of the authority of women compared to men, closely linked to the gender credibility gap. It additionally explores this disparity beyond professional settings. Indeed, very few published studies focused on the disparities in trust and credibility given to women and men in social settings. However, the literature on the gender credibility gap encompasses a large range of empirical and theoretical studies examining this disparity in media or professional settings.

Brann and Himes (2010) for instance, conducted an experimental study to investigate the perceived credibility of television newscasters by gender. It is implied that participants, undergraduate students, were randomly assigned to view either a man or a woman newscaster delivering a weather-related story. After viewing the clips, students had to rate their credibility based on dimensions such as competence, character, sociability, composure, and extroversion. The study controlled for variables such as physical attractiveness and vocal quality to isolate the impact of gender on credibility. The authors found that perceived credibility on dimensions such as competence, composure, and extroversion was higher for male newscasters than for women. No significant differences

were observed in terms of character and sociability. This study underscores the persistence of genre stereotypes that favour men in authoritative roles.

Wissmath et al. (2008) conducted a similar study however focusing on Switzerland. They explore how the age and gender of Swiss newscasters affect their perceived credibility. The diversity of the sample was crucial for the investigation to be representative of the country. However, of 156 participants, 111 were students. These participants were randomly assigned to four categories: old woman, old man, young woman, and young man. They were then asked to rate the credibility of the statements and the speakers themselves. The researchers found that while women's statements were often rated as more credible, the newscasters themselves were perceived as less credible than their male counterparts. This discrepancy highlights the role of gender stereotypes in shaping perceptions of authority.

Andsager and Mastin (2003) studied the differences in the credibility of political columnists by gender and race, employing a robust experimental methodology where participants were shown the same column, whose topic was an account of hate speech, with a randomly assigned gender and race identifier. The identifiers were four mugs shot which were chosen to be as similar as possible in terms of age, dress, hairstyle, expression, and angle, to limit omitted variable bias. The study included students from different regions to account for cultural and regional variations in credibility perceptions. The findings revealed no significant differences in credibility ratings based on gender alone, however exhibiting a slightly higher insignificant rating of credibility for women.

In a study published in 2009, Armstrong and McAdam's explored how gender cues influence perceptions of credibility in the context of informational blogs. Participants, drawn from undergraduate classes at a large university in the United States, were offered extra credit for their participation. The paper involved two separate experimental designs conducted in November 2005 and November 2007. In the former, respondents were exposed to blog posts about rebuilding homes in New Orleans after Hurricane Katrina that were randomly presented as authored by either a man (James Fitzegarld), a woman (Ann Fitzgerald), or gender-neutral pseudonyms (Urbanite). They were then asked to rate the credibility of the blog. The latter was similar, it only added another step in which students were randomly assigned to read one of two blogs, about either bottled

water or the academic rankings of colleges. The authorship was assigned as previously done, randomly and using the same pseudonyms for men and women but Iconoclast for the gender-neutral one. Armstrong and McAdam's (2009) found that make-authored blogs were generally perceived as more credible. This disparity is more pronounced for 'information seekers' who, as they used the blog primarily to find specific information, tended to value accuracy, trustworthiness, and reliability the most. This led them to perceive male authors as more credible. The research hence highlights the ongoing influence of gender biases, which associate authority and expertise with men, in shaping perceptions of credibility and provides insights into how these biases manifest in online environments.

Overall, these studies provide a robust empirical and theoretical foundation for understanding the gender authority-credibility gap. They highlight the pervasive nature of gender biases and their impact on perceptions of authority across different media. This research seeks however to extend the established body of knowledge relating to the analysis of gendered perceptions of credibility. By moving beyond the analysis of media explored by Wissmath et al. (2008), Brann and Himes (2010), Armstrong and McAdam's (2009), and Andsager and Mastin (2003), this work examines how such disparities in credibility and authority perception manifest in diverse everyday interactions. The study designs employ a broader population, transcending the focus on student demographics often found in prior research.

By doing so, the research additionally aims to uncover whether, as respondents' age increases, the credibility they are assigned to videos featuring a woman compared to a man is smaller. Radvansky et al. (2008) in their research suggest that, with age, inhibitory control declines, hence making older adults more inclined to rely on stereotypes. This aligns with the findings of Radvansky et al. (2010), who also found that older individuals exhibit greater difficulty in suppressing stereotypic thoughts. As cognitive resources diminish with age, older adults might struggle more to override automatic, stereotype-consistent thoughts, thereby perpetuating biases more than younger individuals. Most of these studies however solely focus on broader biases. Our paper aims hence to add to this literature by testing whether this disparity between age intervals applies to gender bias too. We hence test the hypothesis that the interaction term between age and being treated is negative.

Research also suggests that men and women do not significantly differ in the patterns of trust they assign to a woman compared to a man. Armstrong and McAdams (2009) as expressed above found that male-authored blogs were generally perceived as more credible, however, this perception did not significantly differ between male and female participants, indicating that both genders share similar biases in evaluating credibility. Coffman (2014) also highlighted that gender biases in trust and credibility are perpetuated by both men and women, further supporting the notion that the tendency to rate male-authored content as more credible is not predominantly influenced by the evaluator's gender. Additionally, Croson and Buchan (1999) demonstrated that both men and women exhibit similar patterns in trust games, suggesting that gender stereotypes in trust and credibility are upheld by both genders. These findings collectively lead to the prediction that no statistically significant differences will be found. We hence test the hypothesis that at least one of the interaction effects of being treated on trust score is different for men, women, and non-binary people.

Political affiliation is furthermore often found to be correlated with the perpetuation of gender biases, with right-wing individuals often exhibiting stronger biases compared to their left-wing or centrist counterparts. The study by Stern and Axt (2021) underscores that conservative individuals tend to maintain more rigid gender stereotypes, largely due to their preference for stability and resistance to change. Hammond-Thrasher and Järvikivi (2023) further support this by demonstrating that liberal individuals, who score higher on traits such as openness and empathy, are more sensitive to gender stereotype violations and less reliant on cognitive biases. We hence predict that conservative (right-wing) individuals see women as less credible than men compared to liberal (left-wing) or centrist individuals. Having available data on respondents' political affiliation hence allows the paper to test the hypothesis that the interaction effects of being treated on average trust score are the same for right-wing, left-wing, and centrist individuals.

Predictions concerning the role of education levels are also made. Coffe et al. (2023) while exploring the relationship between masculinity, sexism, and support for populist radical right parties, specifically in Spain, found that education negatively correlated with the likelihood of voting for VOX and higher educational levels were linked to lower levels of sexist attitudes. Garaigordobil et al. (2013) also uncovered a relationship between sexism, alexithymia, and level of education. Researchers, as they found that higher levels

were associated with lower levels of hostile, benevolent, and ambivalent sexism, reinforced the idea that education can mitigate sexist beliefs. We hence hypothesised, based on these studies, that as the respondent's level of education increases, the effect of treatment, visioning Charlotte instead of Leo will vary.

Lastly, the paper aims to assess whether the relationship differs depending on whether the subject discussed is more often associated with women or with men. Specifically, it hypothesised that when the statements are perceived as more "feminine", Charlotte will be trusted more compared to when perceived as more "masculine". To this end, statements made by Leo and Charlotte included environmental issues, animal welfare, human rights, botanical classification, and historical events. Based on established cognitive associations between the concepts of "green" or sustainability and femininity as highlighted by Brough et al. (2016), facts 3 and 4 are likely to be (un)consciously perceived as more feminine. Statement 5, relating the shortest war in history is linked (un)consciously to men. Cynthia Enloe (2000) argues that societies often socialise men towards violence hence creating an unconscious association between masculinity and war. The results of the 'masculine' statements are the same as the results relating to the length of the shortest war statement in Table 4 Model 5. It was added in both for comparison purposes. The last posited hypothesis is hence that the coefficients of being treated in the models including only 'feminine' or 'masculine' facts are significantly different from one another.

The methods used to test these hypotheses as well as the regressions they yield are further discussed in the hypothesis part 3.4.2.

3. Empirical Strategy and Data

3.1 Survey

To quantify the gender credibility-authority gap, a small-scale survey, available in two languages (English and French) was conducted in May 2024. It comprises two distinct blocks of questions. The first block aimed to gather demographic information from respondents, encompassing factors such as gender, age, educational attainment, nationality, number of children, wage, occupation, and political ideology. This

demographic data provides essential context for understanding the perspectives of the participants. The second block, the focal point of the study, presents participants with a series of videos featuring either Charlotte Capellini or Leo Leitenberger. In these videos, the respective individuals convey factual statements, and participants are prompted to rate their perceived likelihood of each statement being true on a scale from 1 to 10. The number of videos shown however had to be limited to keep it short and not discourage respondents from finishing the survey. The facts were the following:

(1) 'Bananasareclassimiedasberries,whereasstrawberriesarenot.'(McVean&Lee, 2017)
(2) 'Approximately90thousandwomenandgirlsaremurderedeachyear,withhalfof those murders being perpetrated by a family member.' (U.N., 2022)
(3) 'Everyyear,around7millionpeopledieduetoairpollution.'(WHO,2023)
(4) 'More than 1 trillion mishes and 92 billion farm animals are killed by humans annually.' (The Humane Society of the United States, 2023) (FishCount, 2019)
(5) 'Theshortestwarlastedonly38minutes.'(HistoricUK,2015)

At the start of the survey and between the two blocks, respondents were reminded of the importance of being genuine and giving their best guess. This aimed at ensuring that they were focused and answered truthfully. Uncoherent answers were moreover excluded from the survey. This only happened once, with the respondent being 9 years old, from Niger, having a doctorate, being unemployed but earning more than four thousand a month.

3.2 Data

Upon the closing of the survey, 529 people opened the link, 322 started it, 208 finished it, and 207 responses were kept. Of these 207 respondents, 87 surveyed in French and 120 in English. Concerning the respondents' gender, 121, specifically, 58.45% were women, 84, hence 40.58% were men, and 2, 0.97% were non-binary. The low number of observations in the last category in addition to both being assigned to the control group makes it complicated to draw robust conclusions regarding non-binary individuals. The age of the respondents varies from 16 to 75 with the median at 21 and the mean marginally above 29 years of age. A large part, 76.33%, of individuals

additionally do not have kids. Of the ones with children, 28 have two, 12 have only one, eight have three and only one person has four.

More than half of respondents feel the most affiliated with left-wing political parties, specifically 113, 54.59%. 72 or 34.78% are centrist, and 22, or 10.63% are right-wing. The highest level of education obtained or in the process of obtaining are as follows: six individuals have no diploma or a brevet (level 0), 64 have a high school diploma (level 1), three have vocational training (level 2), 24 have a baccalaureate plus two years of further study (level 3), 73 have a bachelor's degree or equivalent (level 4), 30 have a master's degree or equivalent (level 5), and seven have a doctorate or PhD (level 6).

A large part of respondents are students, specifically, 113 or 54.59%. Of the rest, 12 are retired, 70 are employed of which 22 are part-time workers, 12 are unemployed with 3 of them actively looking. The last observed variable is monthly income. Although respondents could, if wished, not answer the question, all did. 132 of them have an income below 1001€, 27 between 1001€ and 2500€, 25 between 2501€ and 4000€, and lastly, 23 earned strictly more than 4000€.

3.3 Quasi-randomisation

After answering demographic questions in the first block, participants were quasi-randomly assigned to one of two paths within the second block as the survey platform lacked the capability for true randomisation. This means that people are assigned to treatment based on a non-random criterion that is assumed to be uncorrelated with the outcome of interest. Before moving on to block two, they were asked if their date of birth was an odd or even number. If the person was born on the 5th of June for instance, they had to select odd and were then shown videos featuring either Leo. If on the other hand, they answered even, they were shown videos featuring Charlotte. The treatment group consists of individuals that were shown videos of Charlotte whilst individuals in the control group were shown videos of Leo. This method minimises biases, ensuring that differences in perceived trustworthiness between speakers could be more reliably attributed to their gender.

To ensure the effectiveness of the quasi-random assignment, tests for balance across all observable characteristics between the treated and control groups: age, gender, language

chosen, level of study, income, political affiliation, occupation, and the number of children the participant has, were undertaken. The results of the t-tests conducted are to be found in Table 1. The p-value of the difference in means between the control and treatment groups of the aforementioned variables can be found in the last column. A value below 5% suggests a significant disparity.

All balance tests show negligible differences. Indeed, all p-values but one are larger than 10%. Apart from the lowest, p-values range from 0.134 to 0.934, hence revealing no statistically significant differences. Specifically, for age, p-value = 0.934, for language chosen, p-value = 0.138, for political affiliation, $0.134 \leq$ p-values ≤ 0.811, for gender, $0.155 \leq$ p-values ≤ 0.535, for number of children, p-value = 0.786, for income, $0.178 \leq$ p-values ≤ 0.817, for occupation, $0.526 \leq$ p-values ≤ 0.566, and lastly, for education level, $0.053 \leq$ p-values ≤ 0.674.

The only significant p-value can be found in the last row of the table. It concerns the highest study level obtained or in the process of obtaining, PhD. With a p-value equal to 0.053, it is however only statistically significant at level 10%. The results hence indicate, given the overall balance achieved, that the quasi-randomisation can be considered to have been effective.

3.4 Methodology

By employing this quasi-random design, the study aims to investigate potential biases and variations in perception based on the gender of the speaker, independent of other demographic factors. It is a powerful method that helps enhance internal validity. It minimises endogeneity issues and selection bias by respectively using an instrument unrelated to the outcome variable, if the day of birth is even or odd, and approximating a random assignment.

3.4.1 Research Question

The effectiveness of the quasi-random assignment to treatment and control groups hence implies that there should be little to no selection bias during assignment to treatment in the survey. This in turn would signify that the outcome without treatment for the treated and the outcome without treatment for the non-treated are the same. A simple with-and-

without comparison can hence, as previously mentioned, be used to uncover the ATT which is represented by the parameter ρ in the following regression of interest:

$$(1)\ Trust_i = \alpha + \rho * Treated_i + \varepsilon_i$$

We call trust the likelihood that a respondent believes a statement to be true, also referred to as trust score in the study. When part of the control group, respondents are associated with a trust score equal to alpha. The variable *Treated* equals one if the respondent is part of the treatment group and zero if they are in the control group, respectively either shown videos featuring Charlotte or Leo. Its effect on the amount of trust one has for someone, represented by the variable *Trust* which is the average answer people gave to all five statements in block two, is the coefficient ρ.

The statements made by Leo and Charlotte moreover covered a wide range of topics. These included environmental issues, animal welfare, human rights, botanical classification, and historical events. Simple with-and-without regression analyses for each of them individually can be found in Table 4. The large range of topics mentioned hence allows the conclusions to be broad and to apply to credibility given to women outside of workspaces compared to men.

3.4.2 Hypotheses

While the main ATT of speaker gender on trust showed a statistically significant difference, it is important to understand if it varies depending on the observed characteristics of respondents. Including interaction terms allows us to do so. The regressions including different interaction terms which can be found in Table 3. Each of the first four columns include one of the interactions of interest. This allowed to improve the interpretability of interaction effects. To test the first four hypotheses formulated previously regressions including interaction effects are performed. The four regressions yielded are hence the following:

$$(2)\ Trust_i = \alpha + \rho * Treated_i + \beta * Age_i + \theta * (Treated_i \times Age_i) + \varepsilon_i$$

$$(3)\ Trust_i = \alpha + \rho * Treated_i + \beta_1 * Woman_i + \beta_2 * Non\text{-}binary_i + \theta * (Treated_i \times Woman_i) + \varepsilon_i$$

$$(4)\ Trust_i = \alpha + \rho * Treated_i + \beta_1 * Right_i + \beta_2 * Left_i + \theta_1 * (Treated_i \times Right_i) + \theta_2 * (Treated_i \times Left_i) + \varepsilon_i$$

(5) $Trust_i = \alpha + \rho*Treated_i + \beta_1*High\ School_i + \beta_2*Vocational\ Training_i + \beta_3*Bac2_i + \beta_4*Bac3_i + \beta_5*Bac5_i + \beta_6*PhD_i + \theta_1*(Treated_i \times High\ School_i) + \theta_2*(Treated_i \times Vocational\ Training_i) + \theta_3*(Treated_i \times Bac2_i) + \theta_4*(Treated_i \times Bac3_i) + \theta_5*(Treated_i \times Bac5_i) + \theta_6*(Treated_i \times PhD_i) + \varepsilon_i$

Following posited hypotheses and their regressions, the following predictions are made:

- Coefmicients θ and β of regression (2) will be negative.
- The different interactio terms of regresion (3) are not signiSicantly different from one another.
- In regression (4), θ_1 is expected to be smaller than θ_2, as we expect right-wing individuals to trust Charlotte less than centrist or left-wing for whom we expect no significant differences in results.
- Lastly, we expect more educated people to trust charlotte more than people with lower levels of education, hence having θ_1 to θ_6 gradually increase.

Regressions two to five, similarly to the baseline regression include the dummy variable for treatment, a constant and an error term. Terms relating the variables themselves as well as their interaction with the being assigned to the treatment group are however added. A more detailed explanation of how to interpret these coefficients can be found in the results parts.

Moreover, the paper tests the hypothesis that the effect becomes larger if the subjects discussed are associated with women more than men. To this end, in Table 2, Column 3 and 5, the following two additional simple with-and-without regression analysis were performed:

(6) $Trust\ Masculine\ Facts_i = \alpha + \rho_m*Treated_i + \varepsilon_i$

(7) $Trust\ Femine\ Facts_i = \alpha + \rho_f*Treated_i + \varepsilon_i$

Based on established cognitive associations ‘feminine facts’ include statements 3 and 4 whilst ‘masculine facts’ include statement 5 (Brough et al., 2016; Enloe, 2000). The results of the ‘masculine’ statements are the same as the results relating to the length of the shortest war statement in Table 4 Model 5. It was added in both tables for comparison purposes. The other two statements are considered neutral as no papers could be found that study the associations between the subjects and gender.

3.4.3 Robustness

Although the balance tests revealed the effectiveness of the quasi-randomisation, tests are performed to ensure the robustness of our results. The latter include additional regressions that control for all observed variables. These additional steps, which can be found in Tables 2 and 4, help minimise any potential selection bias by ensuring that the groups being compared are similar with respect to the unbalanced characteristics, thereby strengthening the validity of our findings. The robustness of the interaction terms is also tested by checking their variability when adding all interactions together in Model 5 of Table 3. The results of these variability checks are further discussed in the robustness part.

Additional means were used to further elevate internal validity, the extent to which a study accurately establishes a cause-and-effect relationship between variables, without being influenced by external factors. Firstly, factors such as tone of voice, body language, and visual cues were thoroughly assessed to ensure consistency across all videos. Secondly, both actors are cisgender and of white ethnicity to ensure racial parity. This however might not be enough to fully prevent endogeneity issues. Despite the efforts to limit selection and omitted variable biases, Leo and Charlotte have different accents, hair textures, and other characteristics. The ATT may consequently still be subject to bias. This issue is further discussed in the robustness section.

The external validity of the survey is also limited. This is due to several factors. Primarily, the survey was conducted over a short period of only one month, imposing a significant time constraint. This led to reliance on a convenience sampling method, where the survey was initially shared with friends and family. As they then shared it with their network, it transitioned into snowball sampling. While these approaches can be beneficial for quickly gathering a large number of responses, they also reduce the representativeness of the sample. More details on how this sampling method affects the reliability of our results in the discussion. Additionally, as any sound responses were accepted, the respondent pool grew somewhat haphazardly, lacking a targeted demographic focus. This lack of specificity in the sample ultimately limits the generalisability of the survey findings. To achieve a more meaningful research outcome, a large and representative sample size is

crucial. However, given the different constraints faced, restricting the survey to focus on any particular country or demographic group was not an option.

4. Results

The following part presents the results found in Tables 2, 3, and 4. Coefficients are considered significant when they have a p-value smaller than 0,05.

4.1 Baseline Model

The regression analysis of the baseline model in Column 1 of Table 2 reveals that being assigned to the treatment group is associated, on average, with a statistically significant increase in rated likelihood of the statements to be true by 0.690 units compared to being assigned to the control group (p-value<0.01). The analysis however has limited explanatory power, as indicated by the small r-squared, explaining under 5% of the variance. This suggests that respondents tended to attribute higher levels of trust to statements when expressed by Charlotte compared to when expressed by Leo. This difference in trust attribution based on the speaker's gender contradict some prior literature, warranting further investigation into potential factors contributing to this discrepancy (Wissmath et al., 2008; Armstrong et al., 2009). It is additionally important to acknowledge that the results may be subject to bias and hence have limited generalisability. Social desirability bias might have for instance led the respondents to seek for some sort of approval. The use of snowball sampling moreover may lead to bias as respondents, knowing the research or their acquaintances, likely share certain values or have expectations relating what the study is about. The possible sources of endogeneity and bias are further discussed in the limitation part 6.2.

The result of the simple regression analyses specific to each fact are found in Table 4 Columns 1, 3, 5, 7, and 9. All ATT present positive coefficients varying from 0.375 for the fact about violence against girls and women to 1.079 when the statement expressed how many fishes and farm animals are murdered by humans each year. The former, along with statements 1 and 5, are statistically insignificant. The latter and the third fact though are statistically significant at 1% and 5% levels, respectively. The disparity in trust between men and women hence does differ depending on the subject talked about. This variability

suggests that trust disparities between men and women speakers may depend on the topic, indicating that certain subjects may inherently carry gender biases in trust. It highlights that depending on the aim, the spokespersons should either be aligned with or misaligned with the perception of the content they present.

4.2 Hypotheses

Analyses involving interaction terms, found in Table 3, reveals additional insights allowing us to evaluate the hypotheses posited in the introduction.

4.2.1 First Hypothesis

The results of regression (2), which include the interaction term between age and being treated, can be found in Column 1 of Table 3. The coefficient associated with treated indicates that, holding age constant, treated individuals have, on average, a trust score 0.861 units higher than untreated individuals. Age coefficient suggests that for each additional year of age of respondents, average trust score decreases by 0.025 units, holding treatment status constant. Finally, the interaction term of -0.006 indicates that as the age of respondents increase, the disparity in trust between Charlotte and Leo narrows slightly. Indeed, individuals in the treatment group are associated with a decrease in average trust score of -0.031 for each additional year of age whilst in the control group, individuals are associated with a decrease in trust score of 0.025 for each additional year of age. Although the results are in accordance with our predictions, none of the coefficients are statistically significant and the model only explains 9.7% of the variation. We hence cannot reject the null hypothesis as there is insufficient evidence to conclude that the observed interaction is different from zero.

4.2.2 Second Hypothesis

The following interaction terms in Column 2 Table 3 relates to the respondent's gender. Compared to men in the control group, untreated women are, on average, associated with a trust score lower by 0.009, although this coefficient is also not statistically significant. Non-binary people in the control group are associated with a trust score higher by 1.477 compared to untreated men, statistically significant at the 1% level. However, it is important to note that there were only two non-binary respondents, both in the control

group, which may limit the reliability and generalisability of this result. When treated, both men and women are associated with an increase in trust score. The coefficient associated with being a woman, -0.891, signifies that the increase in trust score when treated is lowered to 0.375 for treated women compared to treated men for whom the increase stays at 1.267. Both coefficients are statistically significant, the former at the 10% level whilst the latter at the 1% level. Overall, the results suggest potential differences in trust scores across gender and treatment groups. Men are associated with higher trust when facing Charlotte compared to when facing Leo. Although this applies to women too, compared to men they are associated with lower levels of trust. The p-value for the interaction term is 0.079, which is not statistically significant at the 5% level. Therefore, we fail to reject the null hypothesis that the interaction effect of being treated and being a woman is zero. There are no observations for the interaction term, non-binary in the treatment group, so we cannot test this hypothesis. However, the main effect of being non-binary is significant, indicating a difference from the reference group (men). This is in accordance with our predictions and past literature findings that women and men do not differ in significantly in the perpatuation of gender biases (Armstrong & McAdams, 2009; Coffman, 2014; Croson & Buchan, 1999).

4.2.3 Third Hypothesis

In column 3 of Table 3, the interaction terms between political affiliation of the respondents and whether they were assigned to the treatment group are shown. Compared to politically centrist individuals in the control group, those who are politically right have an average trust score that is higher by 0.453. Left affiliated respondents, who were assigned to the control group, have an average trust score lower by 0.159 compared to the untreated centrists. With p-values larger than 10%, these coefficients are not statistically significant. In the treatment group, politically centrist individuals have an average trust score that is statistically significant, at 5% level, higher by 0.936 points compared to when in the control group. The coefficient of individuals who are politically right in the treatment group is -2.674 and is statistically significant at the 5% level. This signifies that they have on average a trust score lower by 0.964 - 2.674 = -1.710 units compared to when untreated. It additionally means that when right wing individuals are treated their trust score is lower than treated left-wing individuals by on average 1.898 units. The coefficients relating left affiliated respondents are however not statistically

significant, meaning that we cannot reject the null hypothesis that the interaction effects between political affiliation and treated are the same for left-wing and centrist respondents. There is nonetheless enough evidence for us to reject the hypothesis that the interaction effect of right-wing individuals with treatment is equal to the interaction effects of other political affiliations. These results support our predictions that right-wing individuals tend to trust women less than men.

4.2.4 Fourth Hypothesis

The regression analysis in Column 4 of Table 3, which explains 31.6% of the variation in the data, presents the interaction between different levels of education and being treated. The main effect of treatment, 2.15, is not statistically significant. The coefficients for education levels (levels 1 to 6) are also positive and not statistically significant. Specifically, the coefficients are the following: β_1=3.493, β_2=3.4, β_3=3.28, β_4=3.551, β_5=3.1, β_6=3.2. Conversly, although they also all have p-values largely above the 0.05 threshold, the interactions terms are all negative. We hence fail to reject the null hypothesis that the interaction terms are equal. This means that there is no significant evidence to suggest that the effect of treatment varies across different levels of education. The interaction terms additionally do not exihibt any sort of pattern that might support the predictions made.

4.2.5 Fifth Hypothesis

The results of the simple with-and-without comparisons of the average rated likelihood of the statements more likely to be (un)consciously associated with 'masculinity' and 'femininity' to be true are found in Models 3 and 5 of Table 1. The coefficients ρ_m = 0.566 and ρ_f = 0.914, have respectively standard errors equal to 0.388 and 0.298 and p-values at 0.146 and 0.002. In order to test whether the difference between the coefficients is statistically significant we derive the critical t-value, |t| = 0.711. For a two-tailed test at the 0.05 significance level and degree of freedom of 412, the critical t-value is approximately 1.96. As the found critical t-value is smaller than 1.96, we fail to reject the null hypothesis that the effect of the treatment differs significantly between the two models. The results are nonetheless in accordance with our predictions and are intuitive. People would tend to trust more Charlotte on facts 3 and 4 and Leo on fact 5. This could

imply that gender stereotypes influence trust in information. This could underscore the importance of matching spokespersons to how the content they present is perceived, possibly enhancing credibility and trust. It could also imply that to address societal biases, mismatches should be made when sharing information.

5. Robustness

In order to check the robustness of our results and limit selection bias, regression analyses controlling for all observed variables were hence conducted for all models. In Column 2, 4, and 6 of Table 2, the ATT of the baseline model and of the models that only account for 'masculine' and 'feminine' facts, increases barely. Their statistically significance moreover stays constant for Model 2 and 6. The significance of Model 3 only increases to the 10% level. The inclusion of control variables also enhances the extent to which the regressions explain the variability in the data. The r-squared of all three models are more than doubled.

The robustness of the results of Models including the statements individually is additionally assessed by adding all available control variables: age, gender, occupation, income, education level, political affiliation, number of children, and language chosen. The found coefficients and the r-squared, in Models 2, 4, 6, 8, and 10 of Table 4, similarly to previous models, all increase slightly in amplitude. As the coefficients and their significance levels remain largely stable, confidence in the internal validity of the analyses is consequently enhanced.

In the final column of Table 3, Model 5 includes all interaction terms to test the variability of the found coefficients. The interaction of age with treatment shows a slight increase but remains statistically insignificant. Similarly, the coefficient representing the impact of being a woman on the average trust score, compared to being a man, is inverse but stays insignificant. The interaction term between being a woman and being treated increases both in magnitude and significance, suggesting that, contrary to previous findings, the combination of being a woman and being shown videos featuring Charlotte might significantly impact average trust scores differently than for men. All coefficients related to political affiliation and their significant remain consistent. However, for education levels, while the main coefficients stay constant with increased significance,

the interaction terms decrease substantially, almost doubling for each level of education. Despite this variability, the interaction effects are not statistically significant, this variation is thus not concerning.

The overall stability of the findings hints that the internal validity of our model holds. It however does not confirm that the conditional mean independence assumption is not violated by omitted variable biases stemming from the different characteristics of Leo and Charlotte. The findings generalisability may hence still be limited. These limitations are detailed in part 6.2.

6. Discussion

6.1 Takeaways

The main takeaways from our analysis indicate that gender dynamics are associated with variations in trust perception, as respondents attributed higher trust levels to statements made by Charlotte compared to Leo. Specifically, the regression analysis of the baseline model reveals that assignment to the treatment group is associated with a statistically significant increase of 0.690 points in the rated likelihood of the statements being true (p-value<0.01). Alongside this, when performing robustness check and including additional control variables, the coefficient's size is even larger, confirming our initial findings that individuals tend to trust Charlotte more than Leo.

Further analysis of the interaction effects of demographic factors such as age, gender, education levels, and political affiliation provides further insights. Suggesting for instance that the trust disparity between both subjects narrows slightly as respondents' age increases. Although these coefficients are insignificant, hence not allowing to disprove the null hypothesis, they corroborate our predictions. The interaction between respondents' gender and treatment shows varied trust scores, with non-binary respondents exhibiting higher trust scores than men and women, though the small sample size limits the reliability of this result.

The analysis regarding the interaction between education levels and treatment refute our fourth hypothesis that higher levels of education correspond on average to elevated trust scores within the treatment group. Conversely, examinations of political affiliation

interactions support our third hypothesis. Notable disparities between right-wing and centrist or left-wing individuals in the treatment group corroborate our predictions that respondents affiliated with right-wing parties would tend to trust women less.

Simple comparisons of the average rated likelihood of statements associated with masculine and feminine facts indicate that the disparity in trust respondents show is more pronounced when Charlotte makes statements on 'feminine facts compared to 'masculine' facts. However, the lack of statistical significance in the difference between these two coefficients contradict our fifth hypothesis.

6.2 Limitations

Though the results are insightful, they do not allow for causal inference or definitive conclusions regarding policy implications. This is especially true as findings differ for expectation and past literature. A reason for that might be the gathering of data leading to experimenter demand effects and social desirability bias. While quasi-randomisation was effective hence limiting selection bias during assignment, selection bias from recruitment likely persisted. The use of convenience and snowball sampling methods to gather data introduces several limitations that affect the generalisability and validity of the findings. Convenience sampling biases the sample towards easily accessible participants, who are readily available and willing to participate. Snowball sampling further narrows participant diversity as initial respondents refer to others similar to themselves.

The participants, knowing the researcher, are likely to have altered their behavior to align with what they perceive as the experimenter's expectations. This would lead to artificially inflated responses that do not accurately reflect the participants' true attitudes or behaviors, thus skewing the results. Similarly, the social desirability bias, arises when participants provide responses they believe will be viewed favorably by others, rather than their genuine thoughts or feelings. This bias can result in overreporting socially acceptable behaviors or underreporting undesirable ones. Together, these biases can compromise the internal validity of our analysis. Both biases could be expected to be limited since most repondents only viewed an drated videos featuring a single presenter, unaware that other participants might view videos featurign someone else. However,

individuals may unconsciously react to seeing a woman be presenting facts, perceiving the situation as significant. This perception could lead them to rate Charlotte highly to avoid being seen as discriminatory against women. The reliability of the conclusions drawn is consequently diminished, impacting the study's overall contribution to the field and its policy implications.

The lack of random sampling and the limited amount of data also comprise the external validity of our findings and complicate the estimation of sampling error. It appears, however, that the bias is likely positive. If we for instance consider only French respondents, approximately 40% have a bachelor's, master's, or doctorate, while the national average is around 27%. This overrepresentation of individuals with higher education levels suggests a positive bias, as these levels of education are associated with larger estimates.

Additionally, it has come to my attention that participants did not thoroughly read the instructions under certain questions. For instance, when asked about their highest level of education, many bachelor's students reported their level as "baccalaureate" because they did not notice the note indicating they should include levels they are in the process of obtaining. Therefore, the variable 'study level' is not entirely reliable, and any conclusions drawn from it should be approached with caution. This inaccuracy should however not have significantly affected the baseline model's results, as quasi-randomisation should have averaged out such errors. However, it implies that more individuals fall into the higher education categories than the data suggests, potentially exacerbating the positive bias mentioned above.

Other sources of endogeneity may have introduced bias into the results. Charlotte and Leo differ in several ways that are challenging to control for, potentially leading to omitted variable bias. For instance, variations in speakers' accents, speaking speed, hair texture, and other characteristics may contribute to this bias. Additionally, video angles, camera equipment, and background disparities further complicate the comparison. The direction of these biases remains uncertain, they could either inflate or deflate the estimated coefficients. These biases could however be mitigated by having a single person record all videos and ensure uniformity in angles and backgrounds. Further research could additionally involve speakers twins or individuals who closely resemble

each other, sharing similar accents, intonations, hair textures, and other attributes. This approach would minimise omitted variable bias to the utmost degree.

In conclusion, although the quasi-randomisation employed in this study mitigate some biases, the research remains vulnerable to selection bias arising from the recruitment process and to various sources of endogeneity. The limited number of observations further constrains the generalisability of the findings. These limitations necessitate a cautious interpretation of the findings. However, they could potentially be addressed with sufficient fundings and time. Implementing randomised sampling and conducting the survey on a larger scale with targeted populations would enable the derivation of causal inferences and more robust findings.

6.3 Policy Implications

The findings underscore the importance of considering gender in communication strategies. However, as mentioned, the results are likely biased and are not representative of any population, limiting their generalisability. The findings additionally contradict some prior literature (Wissmath et al., 2008; Armstrong et al., 2009). This further complicates the derivation of robust policy implications. Hence despite the interesting results, the limited explanatory power and potential biases necessitate cautious interpretation. Policymakers should be aware of these limitations and approach the findings as preliminary, warranting further research to validate the results and explore additional factors before making substantive policy recommendations.

7. Conclusion

The methodology employed effectively investigates disparities in trust perception based on gender dynamics. However, to enhance the precision and reliability of such studies, additional resources and improvement in data collection are essential. Future research could benefit significantly from increased sample sizes obtained through randomised sampling methods. This approach would not only yield more conclusive findings but also enable a more comprehensive exploration of trust disparities across various demographic factors. Specifically, the investigation could extend to examining disparities in trust based on ethnicity, skin colour, accent, transgender identity, facial hair, physical attractiveness, hair texture, and other pertinent variables.

Moreover, expanding the scope would enable a more comprehensive analysis of how coefficients and effects vary across subgroups such as countries, age, income levels, and other characteristics. Increased data availability for each subgroup would facilitate statistically significant analyses. This was constrained in the present research due to limited observations, particularly in groups that included non-binary individuals where only two observations were available, and both were part of the control group.

It would be additionally important to uncover what leads to the gender authority-credibility gaps. Does the lack of representation, cognitive biases, or systemic discrimination participate in the perpetuation of these discrepancies? Such studies would play a pivotal role in pinpointing and rectifying systemic biases in trust perception, thereby fostering the development of more equitable policies and targeted interventions.

In conclusion, while this study offers valuable insights into gender-based trust disparities, its limitations highlight the need for continued research with enhanced resources. By leveraging larger, more representative samples and rigorous randomised sampling techniques, future studies can advance our understanding of trust perception across diverse populations and contexts, laying the groundwork for more robust policy recommendations and societal interventions.

References

Armstrong, C. L. & McAdams, M. J. (2009). Blogs of information: How gender cues and individual motivations influence perceptions of credibility. *Journal of Compute Mediated Communication*, *14*(3), 435-456.

Andsager, J. L., & Mastin, T. (2003). Racial and regional differences in readers' evaluations of the credibility of political columnists by race and sex. *Journalism & Mass Communication Quarterly*, *80*(1), 57-72.

Blommaert, L., Meuleman, R., Leenheer, S., & Butkēviča, A. (2020). The gender gap in job authority: Do social network resources matter? *Acta Sociologica*, *63*(4),381-399.

Brann, M., & Himes, K. L. (2010). Perceived credibility of male versus female television newscasters. *Communication ResearchReports*, *27*(3), 243-252.

Brough, A. R., Wilkie, J. E., Ma, J., Isaac, M. S., & Gal, D. (2016). Is eco-friendly unmanly? The greenfeminine stereotype and its effect on sustainable consumption. *Journal of Consumer Research*, *43*(4), 567-582.

Coffe, H., Fraile, M., Alexander, A., Fortin-Rittberger, J., & Banducci, S. (2023). Masculinity, sexism,and thepopulist radical right support. *Frontiers in Political Science*, *5*,1038659.

Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, *129*(4), 1625-1660.

Croson,R., & Buchan, N. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review*, *89*(2), 386-391.

Enloe, C. (2000). *Maneuvers: The international politics of militarizing women's lives*. Univ of California Press.

European Commission (1998). 100 Words for Equality: A Glossary of Terms on Equality between Women and Men. *EU Publication Office*.

FishCount. (2019). Retrieved from https://fishcount.org.uk

Garaigordobil, M. (2013). Sexism and alexithymia: Correlations and differences as a functionof gender, age, and educational level, *Anales de psicología*, *29*(2), 368-377.

Hammond-Thrasher, S., & Järvikivi, J. (2023). Rating gender stereotype violations: The effects of personality and politics. *Frontiers in Communication*, *8*, 1050662.

Historic UK. (2015). The Shortest War in History. Retrieved from https://www.historicuk.com/HistoryUK/HistoryofBritain/The-Shortest-War-in-History/

Kay, K., & Shipman, C. (2014). The confidence gap. *The Atlantic*, *14*(1), 1-18.

Latu, I. M., Mast, M. S., & Stewart, T. L. (2015). Gender biases in (inter) action: The role of interviewers' and applicants'implicit and explicit stereotypes in predicting women's job interview outcomes. *Psychology of Women Quarterly*, *39*(4), 539-552.

McVean, A., Lee, C. (2017). Bananas are berries; raspberries are not. *McGill Office for Science and Society*. Retrieved from https://www.mcgill.ca/oss/article/did-you-know/bananas-areberries-raspberries-are-not

Radvansky, G. A., Lynchard, N. A., & von Hippel, W. (2008). Aging and stereotype suppression. *Aging, Neuropsychology, and Cognition*, *16*(1), 22-32.

Radvansky, G. A., Copeland, D. E., & Von Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of experimental social psychology*, *46*(1), 51-60.

Sieghart, M. A. (2021). The authority gap: Why women are still taken less seriously than men, and what we can do about it. *W. W. Norton & Companny*.

Sterling, A. D., Thompson, M. E., Wang, S., Kusimo, A., Gilmartin, S., & Sheppard, S. (2020). The confidence gap predicts the gender pay gap among STEM graduates. *Proceedings of the National Academy of Sciences, 117*(44), 27213–27219.

Stern, C., & Axt, J. (2021). Ideological differences in race and gender stereotyping. *Social Cognition*, *39*(2), 259-294.

Stojmenovska, D., Steinmetz, S., & Volker, B. (2021). The gender gap in workplace authority: Variation across types of authority positions. *Social Forces*, *100*(2), 599-621.

The Humane Society of the United States. (2023). More animals than ever— 92 billion—are used and killed for food each year. Retrieved from https://www.humanesociety.org/blog/more-animals-ever-922-billion-are-used-andkilled-each-year-food

Weibel, D., Wissmath, B., & Groner, R. (2008). How gender and age affect newscasters' credibility—aninvestigation in Switzerland. *Journal of Broadcasting & Electronic Media*, *52*(3), 466-484.

Whysall, Z. (2018). Cognitive biases in recruitment, selection, and promotion: The risk of subconscious discrimination. *Hidden inequalities in the workplace: A guide to the current challenges, issues, and business solutions*, 215-243.

Women, U. N. (2022). Gender-related killings of women and girls (femicide/feminicide).

World Health Organization. (2023). Household air pollution and health. Retrieved from https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-andhealth#

Appendices

	Means		t-test
	Treated	Not treated	p-values
Gender:			
Women	0.615	0.553	0.368
Men	0.385	0.427	0.535
Non-binary	0	0.019	0.155
Age	29.375	29.553	0.934
Children	0.481	0.447	0.786
Occupation:			
Unemployed	0.048	0.068	0.543
Employed	0.317	0.359	0.526
Retired	0.067	0.049	0.566
Student	0.567	0.524	0.536
Political affiliation:			
Right	0.058	0.117	0.134
Center	0.356	0.340	0.811
Left	0.587	0.544	0.536
Language chosen	0.471	0.369	0.138
Wage interval:			
[0, 1000]	0.462	0.369	0.178
[1001, 2500]	0.125	0.136	0.817
[2501, 4000]	0.106	0.136	0.508
[4001, ∞[	0.096	0.126	0.494
Education level:			
None	0.038	0.019	0.417
High school	0.346	0.272	0.250
Vocational training	0.010	0.019	0.557
Bac +2	0.135	0.097	0.402
Bac +3	0.327	0.379	0.439
Bac +5	0.135	0.155	0.674

PhD	0.010	0.058	0.053

Table 1. T-tests for all Observed Variables

Note: Table 1 includes data on t-tests for each of the observed characteristics of respondents. In the first two columns, the means of the treatment and control groups can be found. The p-values indicate whether the difference in means between the two groups is significantly different from zero. All the coefficients are rounded to 3 decimal places. Significant at a 10 percent level ($p < 0.1$), Significant at a 5 percent level ($p < 0.05$), Significant at a 1 percent level ($p < 0.01$)

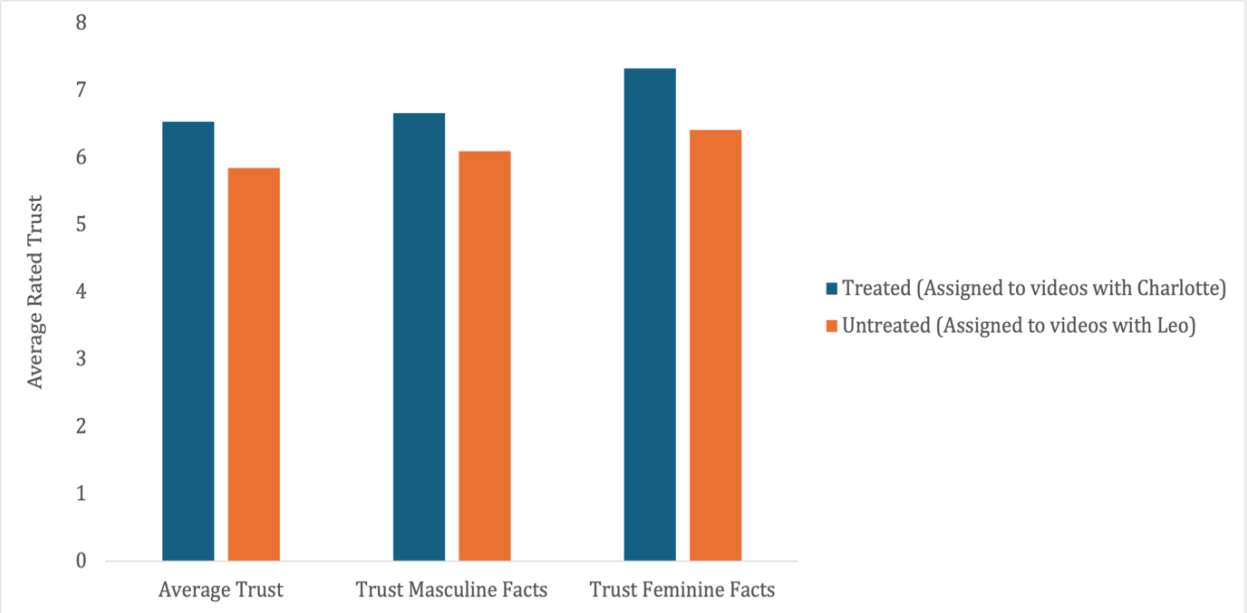

Figure 1. Average Rated Trust for all, Masculine, and Feminine Facts

Note: This figure displays data from three separate regressions analysing trust ratings. The first regression includes the average rated likelihood of all statements being true. The second regression focuses on statements considered masculine. The third regression includes statements considered feminine. The blue bars represent the average trust ratings for the treated group (individuals shown videos featuring Charlotte). The orange bars represent the average trust ratings for the control group (individuals shown videos featuring Leo).

	Average Trust		*Trust Masculine Facts*		*Trust Feminine Facts*	
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Treated	0.690***	0.838***	0.566	0.758*	0.914***	1.006***
	(0.242)	(0.248)	(0.388)	(0.357)	(0.298)	(0.321)
Age		0.023		-0.038		0.042
		(0.028)		(0.034)		(0.036)
Children		-0.349		0.299		-0.662
		(0.340)		(0.400)		(0.407)
Gender:						
Women		-0.540**		-0.911**		0.039
		(0.253)		(0.392)		(0.313)
Non-binary		1.309***		0.590		0.381
		(0.485)		(0.702)		(0.881)
Education level:						
High school		2.134**		2.681**		2.311*
		(0.978)		(1.319)		(1.366)
Vocational training		2.189**		3.063		3.176**
		(1.051)		(2.274)		(1.577)
Bac +2		1.856*		2.452*		1.998
		(0.984)		(1.344)		(1.391)
Bac +3		2.457**		2.422*		2.744*
		(1.010)		(1.385)		(1.414)
Bac +5		2.440**		3.127**		2.317
		(1.026)		(1.385)		(1.440)
PhD		2.927**		3.258**		2.452
		(1.188)		(1.588)		(1.707)
Political affiliation:						
Right		-0.449		0.124		-0.436
		(0.555)		(0.663)		(0.735)
Left		-0.296		-0.447		-0.289
		(0.270)		(0.424)		(0.336)
Occupation:						
Employed		-0.963		-1.361		-1.237
		(0.739)		(1.098)		(0.821)

Retired		-1.949 (1.339)		-0.182 (1.820)		-2.419 (1.725)
Student		-0.163 (0.710)		0.269 (0.984)		-0.572 (0.790)
Wage interval:						
[1001, 2500]		0.658 (0.445)		1.400** (0.575)		0.287 (0.559)
[2501, 4000]		-0.052 (0.619)		-0.507 (0.939)		0.123 (0.790)
[4001, ∞[		-0.410 (0.695)		0.654 (0.973)		0.697 (0.939)
Language chosen		-0.045 (0.277)		-0.713 (0.472)		0.677* (0.350)
Constant	5.847*** (0.188)	4.044*** (1.484)	6.097*** (0.264)	5.593*** (1.721)	6.413*** (0.213)	3.687* (1.912)
Observations	207	207	207	207	207	207
R-squared	0.038	0.208	0.010	0.220	0.044	0.140

Table 2. Regression Analysis of Trust for all Facts Averaged

Note: Table 2 includes data on 6 regressions. For the total average rated likelihood of a statement to be true, a simple regression and one including all observed variables are performed, in models (1) and (2). In the last four columns, similar regressions are performed, including however only data on the average results for 'masculine' facts (fact 5), models (3) and (4), and 'feminine' facts (facts 3 and 4), columns (5) and (6). The standard error for each coefficient is seen in the brackets below. All the coefficients are rounded to 3 decimal places. * Significant at a 10 percent level ($p < 0.1$), ** Significant at a 5 percent level ($p < 0.05$), *** Significant at a 1 percent level ($p < 0.01$)

	Avg Trust	Avg Trust	Avg Trust	Avg Trust	Avg Trust
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Treated	0.861	1.267***	0.974**	2.150	4.984**
	(0.537)	(0.409)	(0.378)	(2.128)	(1.959)
Age	-0.025*				-0.020
	(0.015)				(0.018)
Treated * Age	-0.006				-0.010
	(0.018)				(0.021)
Gender:					
Women		-0.009			0.053
		(0.396)			(0.460)
Non-binary		1.477***			1.475**
		(0.382)			(0.706)
Treated * Women		-0.891*			-1.111**
		(0.504)			(0.541)
Political affiliation:					
Right			0.453		0.355
			(0.657)		(0.711)
Left			-0.159		-0.245
			(0.408)		(0.526)
Treated * Right			-2.674**		-2.687**
			(1.061)		(1.090)
Treated * Left			-0.164		-0.326
			(0.499)		(0.591)
Education level:					
High school				3.493*	3.146**
				(1.871)	(1.540)
Vocational training				3.400*	3.549**
				(1.867)	(1.626)
Bac +2				3.280*	2.976*
				(1.949)	(1.594)
Bac +3				3.551*	3.179**
				(1.854)	(1.530)
Bac +5				3.100	3.067**
				(1.895)	(1.545)
PhD				3.200*	3.271**
				(1.874)	(1.565)
Treated * High school				-1.648	-3.086*
				(2.173)	(1.848)

Treated * Vocational training				-1.250	-2.558
				(2.159)	(1.876)
Treated * Bac +2				-1.816	-3.226*
				(2.286)	(1.892)
Treated * Bac +3				-1.290	-2.752
				(2.164)	(1.840)
Treated * Bac +5				-1.164	-2.971
				(2.223)	(1.819)
Treated * PhD				-0.050	-1.221
				(2.165)	(1.853)
Constant	6.571***	5.823***	5.880***	2.500	3.393**
	(0.018)	(0.316)	(0.316)	(1.831)	(1.604)
Observations	207	207	207	207	207
R-squared	0.097	0.076	0.316	0.114	0.232

Table 3. Regression Results with Interaction Effects

Note: Table 3 includes data on 5 different regressions that include interaction effects between being treated and various observed variables. In Model (1) a regression including the interaction with age, Model (2) , gender, Model (3) includes the interaction of treatment and political affiliation, and Model (4) education level. Lastly, Model (5) presents all interaction terms in a single regression. The standard error for each coefficient is seen in the brackets below. All the coefficients are rounded to 3 decimal places. * Significant at a 10 percent level ($p < 0.1$), ** Significant at a 5 percent level ($p < 0.05$), *** Significant at a 1 percent level ($p < 0.01$).

Facts	*Berries*		*Women*		*Air Pollution*		*Farm*		*War*	
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)	Model (7)	Model (8)	Model (9)	Model (10)
Treated	0.679	0.974*	0.375	0.481	0.750**	0.888**	1.079***	1.125***	0.566	0.758
	(0.506)	(0.504)	(0.332)	(0.339)	(0.750)	(0.405)	(0.347)	(0.346)	(0.388)	(0.357)
Age		0.043		0.251*		0.040		0.044		-0.038
		(0.055)		(0.128)		(0.041)		(0.047)		(0.034)
Children		-0.101		-0.618*		-0.227		-1.097**		0.299
		(0.670)		(0.357)		(0.514)		(0.494)		(0.400)
Gender:										
Women		-1.511***		-0.358		-0.780		0.630*		-0.911**
		(0.542)		(0.358)		(0.869)		(0.355)		(0.392)
Non-binary		3.102***		2.090**		-0.376		1.062*		0.590
		(0.785)		(0.382)		(0.421)		(0.553)		(0.702)
Education level:										
High school		0.827		2.542*		1.859		2.763*		2.681**
		(1.164)		(1.363)		(1.427)		(1.473)		(1.319)
Vocational training		-0.721		2.252		2.562		3.791**		3.063
		(1.779)		(1.442)		(2.176)		(1.473)		(2.274)
Bac +2		0.507		2.324*		1.680		2.316		2.452*
		(1.225)		(1.402)		(1.475)		(1.504)		(1.344)
Bac +3		1.246		3.128**		2.652*		2.836*		2.422*
		(1.185)		(1.389)		(1.473)		(1.513)		(1.334)

Bac +5	1.741	2.697*	1.488	3.147**	3.127***
	(1.323)	(1.464)	(1.608)	(1.520)	(1.385)
PhD	2.149	4.323***	2.358	2.545	3.258**
	(1.843)	(1.539)	(1.777)	(2.013)	(1.588)
Political affiliation:					
Right	-0.756	-0.740	-0.780	-0.092	0.124
	(1.078)	(0.734)	(0.869)	(0.782)	(0.663)
Left	-0.387	-0.068	-0.376	-0.203	-0.447
	(0.557)	(0.382)	(0.421)	(0.370)	(0.424)
Occupation:					
Employed	-1.231	0.253	-1.245	-1.230	-1.361
	(1.478)	(0.931)	(0.998)	(0.876)	(1.098)
Retired	-3.437	-1.290	-2.573	-2.265	-0.182
	(2.310)	(1.667)	(2.070)	(2.002)	(1.820)
Student	0.197	-0.136	0.018	-1.162	0.269
	(1.390)	(0.918)	(0.932)	(0.834)	(0.984)
Wage interval:					
[1001, 2500]	1.714	-0.400	0.836	-0.262	1.400**
	(1.132)	(0.672)	(0.722)	(0.591)	(0.575)
[2501, 4000]	0.042	-0.040	0.917	-0.0671	-0.507
	(1.163)	(0.795)	(0.960)	(0.883)	(0.939)

[4001, ∞[		-3.439**		-0.657		1.823		-0.430		0.654
		(1.397)		(0.891)		(1.119)		(1.085)		(0.974)
Language chosen		-1.162**		0.298		0.673		0.682*		-0.713
		(0.545)		(0.374)		(0.432)		(1.085)		(0.472)
Constant	3.340***	3.177	6.971***	4.074**	6.010***	3.291	6.816***	4.083*	6.097***	5.593***
	(0.361)	(2.384)	(0.243)	(1.803)	(0.268)	(2.086)	(0.278)	(2.112)	(0.264)	(1.721)
Observations	207	207	207	207	207	207	207	207	207	207
R-squared	0.009	0.197	0.006	0.137	0.019	0.126	0.045	0.172	0.010	0.220

Table 4. Regression Analysis of Trust for all Facts

Note: Table 4 includes data on 10 regressions. For each fact, a simple regression and one including all observed variables are performed. The standard error for each coefficient is seen in the brackets below. All the coefficients are rounded to 3 decimal places. * Significant at a 10 percent level ($p < 0.1$), ** Significant at a 5 percent level ($p < 0.05$), *** Significant at a 1 percent level ($p < 0.01$).

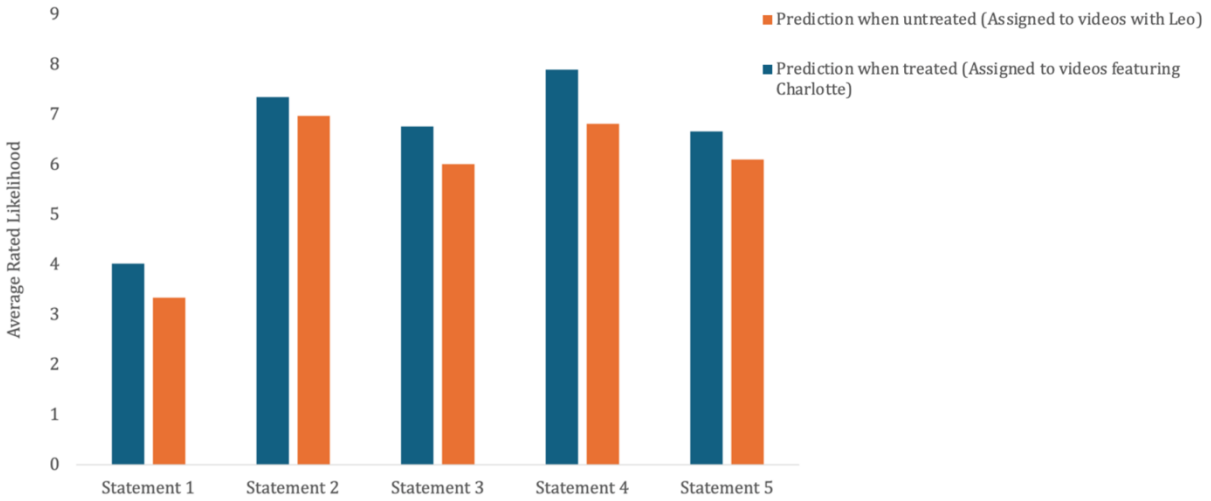

Figure 2. Average Rated Trust for all Statements Individually

Note: This figure displays data from five separate regressions analysing trust ratings, one for each fact presented in block two of the survey. The blue bars represent the average trust ratings for the treated group (individuals shown videos featuring Charlotte). The orange bars represent the average trust ratings for the control group (individuals shown videos featuring Leo). The statements mentioned are the following:

- Statement 1: 'Bananas are classimied as berries, whereas strawberries are not.' (McVean & Lee, 2017)
- Statement 2: 'Approximately 90 thousand women and girls are murdered each year, with half of those murders being perpetrated by a family member.' (U.N., 2022)
- Statement 3: 'Every year, around 7 million people die due to air pollution.' (WHO, 2023)
- Statement 4: 'More than 1 trillion mishes and 92 billion farm animals are killed by humans annually.' (The Humane Society of the United States, 2023) (FishCount, 2019)
- Statement 5: 'The shortest war lasted only 38 minutes.' (Historic UK, 2015)