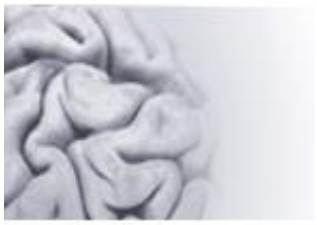
Universiteit Leiden

Faculteit der Sociale Wetenschappen



Reading between the lines

(Preventing) the Effects of Implicit Gender Bias in Selection Procedures

Margot Korteschiel

In collaboration with Ayres Algera, Charlotte Keuning and Marie Nitzgen

Master thesis Psychology, specialization Social and Organizational Psychology
Institute of Psychology
Faculty of Social and Behavioral Sciences – Leiden University

Date: 2017, March 20th
Student number: 1219197
First examiner of the university: Dr. R. A. van der Lee
Second examiner of the university: M. Schutter

Abstract

Women currently remain underrepresented in (academic) top positions. Research has shown that gender stereotypes cause implicit gender bias, which influences organizational behaviour. This helps to maintain gender inequality. The current study focused on evaluator's gender bias against women in selection procedures and possible solutions. Effects of feedback (none versus personal) on implicit gender bias were examined, just like gendered wording (masculine versus balanced) and type of wording (traits versus behaviours) in evaluation forms. Results showed that implicit gender bias negatively influences gender equality perceptions. Also, personal feedback on implicit gender bias increased stereotype activation, which lead to more perceived applicant suitability when male stereotype trait descriptions were used. Without feedback, evaluators were hypothesized to perceive lower applicant suitability when the ideal applicant was described in traits compared to behavioural descriptions. Results did not support this hypothesis. Taken together, measures preventing effects of implicit gender bias in selection procedures were uncovered, contributing to gender inequality in (academic) top positions.

Keywords: Implicit gender bias, gender equality, selection procedures, stereotype activation, gendered wording

Reading Between the Lines:

(Preventing) the Effects of Implicit Gender Bias in Selection Procedures

In January 2016, ten influential companies, including Twitter and Unilever, have disclosed their statistics on gender inequality in their company in 'UN Women's HeForShe Parity Report'. Their statistics show that across these companies, on average 39.7% of employees is female. The higher in the hierarchy, the lower the percentage of women: Men hold between 67% and 89% of senior leadership positions in these ten companies (UN Women, 2016a). When looking at academic gender distributions, women also remain underrepresented in top positions: In Europe, fewer than 40% of academics is female (Grove, 2013), only 18% of women is appointed as a full professor (Vernos, 2013) and only 15.5% of women is employed as Vice Chancellor (Morley, 2014). This shows a decline in the amount of women in academic positions with each step in the career ladder. Despite these facts, when UN Women HeForShe IMPACT Report disclosed statistics on gender inequality from ten influential global universities, including The University of Oxford and Nagoya University, only 40% of these universities had made a commitment to striving for closing the gender gap in academics (UN Women, 2016b). Taken together, these facts demonstrate the extent of gender inequality and the lack of support for gender equality in (academic) top positions.

The current research looks into subtle factors that help to maintain gender inequality in (academic) top positions, like implicit gender bias and gendered wording. The current study therefore aims to identify implicit gender bias in the evaluation process of job applications and its consequences, and also aims to identify possible solutions to the influence of implicit gender bias. It therefore poses the following questions: To what extent does implicit gender bias influence perceptions of gender equality on the work floor? And to what extent does personal feedback about implicit gender bias, gendered wording and type of wording in

evaluation forms of applicants influence the impact that implicit gender bias had on perceived competence of applicants by evaluators?

Gender stereotypes in selection procedures

Striving for gender equality in (academic) top positions is important, as gender equality ensures that women's leadership skills are also optimally used. Also, seeing women in (academic) top positions encourages other women to aspire the same (Sheridan, Fine, Pribbenow, Handelsman & Carnes, 2010). The underrepresentation of women in (academic) top positions is not due to a lack of female potential or a lack of available jobs (Van den Brink, 2010). Stereotypes are involved in this underrepresentation and the maintenance of this inequality. Women are more often associated with 'supporting' qualities, like being communal and more relationally oriented, according to the female stereotype. Men, on the contrast, are more often associated with 'excellent' qualities, like being agentic and leadership-oriented according to the male stereotype (Gaucher et al., 2011; Kaatz, Magua, Zimmerman & Carnes, 2015). In line with these stereotypes, women are stereotypically seen as nicer and more nurturing, but as less competent (Eagly & Mladinic, 1994). Moreover, scientific success and being career oriented are subconsciously linked to male stereotype traits and less to female stereotype traits (Miller, Eagly & Linn, 2015; Nosek et al., 2009; Taris & Bok, 1998). However, when both men and women behave in a way that is not congruent with the stereotype (e.g. a man in a supporting position or a woman in a career-oriented position), they are seen as less likeable and less positive (Heilman, 2001). In other words: In line with gender stereotypes, men are more suitable for jobs in (academic) top positions than women. Applying for a job that is stereotypically seen as more masculine seems more appropriate for men than for women, as it is congruent with the image of the ideal 'masculine' candidate. Therefore, men have a larger chance at obtaining an (academic) top position at work (Taris & Bok, 1998).

Implicit gender bias in organizations

The way these stereotypes influence behaviour within an organization has changed over time. Whereas bias against women and other minorities in an organization used to be rather explicit, organizations nowadays often have the appearance of being unbiased. However, minority stereotypes still lead to implicit bias, which plays a role in the evaluation of job applications (Dovidio & Gaertner, 1996). Implicit bias occurs when evaluators believe they are unprejudiced, yet unconsciously negatively stereotype minority applicants. When judging an applicant, an evaluator has an implicit preference for a majority group member applicant, because (s)he subconsciously believes the majority group member applicant is more suitable for the job based on stereotypes. Due to this implicit process, the evaluator unintentionally pays more attention to possibly disqualifying characteristics of minority applicants, so that the choice for a majority group member applicant is defendable. However, when the majority group member applicant is being evaluated, those same characteristics that disqualified the minority group member applicant, are accepted and not perceived as disqualifying. Thus, stereotypes lead to a biased selection procedure in which majority group members are (implicitly) preferred over minority group members (Dovidio & Gaertner, 1996).

In the current study, gender stereotyping is defined as a subgroup of minority stereotyping. Gender stereotyping leads to implicit gender bias, disadvantaging women in organizations (Rudman & Kilianski, 2000). Like other forms of minority stereotyping, the expression of gender stereotyping has become more concealed compared to several decades ago (Swim, Aikin, Hall & Hunter, 1995). However, implicit gender bias against women is still negatively impacting the hiring process (Dovidio & Gaertner, 1996). Implicit gender bias against women is in line with gender stereotypes: Women are implicitly seen as more suitable for low-authority positions than high-authority positions in an organization, and vice versa for men (Rudman & Kilianski, 2000). Implicit gender bias could influence anyone: People who

explicitly strongly support gender equality may also be biased. Also, both men and women can be affected by implicit gender bias. A way in which this implicit gender bias shows, is through implicit bias, which is sometimes referred to as 'modern sexism' (Swim et al., 1995). Just like implicit minority stereotyping, modern sexism also entails the belief that discrimination based on gender no longer exists, and therefore no support is given for the affirmative action that is needed to help the position of women in organizations (Swim et al., 1995). Experiments by Steinpreis, Anders and Ritzke (1999) and more recently Moss-Racusin, Dovidio, Brescoll, Graham and Handelsman (2012) showed the effects of modern sexism on selection procedures: Male and female academics were lead to believe they could hire a job applicant for a scientific job based on a resume. While the only difference was the applicant's gender on the resume, the willingness to hire the applicant was in favour of men: Both genders preferred to hire male applicants over female applicants. Therefore, this experiment supports the notion that implicit gender bias helps to maintain gender inequality in academic positions, thereby negatively impacting the number of women in (academic) top positions.

Influencing the impact of stereotypes on interpersonal behaviour

Whether people act upon their implicit stereotypical beliefs, is amongst others influenced by the cognitive accessibility of gender stereotypes in a person. The environment can stimulate this accessibility: Activation can take place through implicit reminders of the stereotype; for example, in language. Due to the increased accessibility of stereotype, it is more often unconsciously acted upon as people are implicitly influenced by the stereotype and behave in ways that are congruent with it (Gupta, Turban & Bhawe, 2008; Wheeler & Petty, 2001). Stimulating the accessibility can also be explicit, by overtly reminding a person of gender stereotypes, for example. When stereotypes are made explicit, behaving congruent with the stereotype becomes less likely (Gupta et al., 2008).

The process behind the effects of explicit stereotype activation is a process similar to reactance, called stereotype reactance (Kray, Thompson & Galinsky, 2001). Reactance is described as the reaction people have towards explicit limitations of their freedom. This reaction usually contains a powerful counterstatement: People do not accept these limitations (Dillard & Shen, 2005). Stereotype reactance is a specific form of reactance. When people are told that their own stereotypical beliefs limit themselves in their behaviour, people attempt to demonstrate that those beliefs do not hold them back (Kray et al., 2001). Research by Gupta et al. (2008) illustrates how implicit versus explicit reminders of gender stereotype affect behaviour in different ways through stereotype reactance. In their experiment, business students received information about entrepreneurship that was either implicitly or explicitly connected to either the male stereotypes or female stereotypes, except for a control condition in which no link to a stereotype was made and a condition where both genders were equally emphasized. Then, these students were asked to what extent they had intentions to become an entrepreneur. The outcomes revealed that men had more intentions to become an entrepreneur in the control group, indicating that society already has made an impact on how entrepreneurship is seen by both genders. Also, when both genders were explicitly equally emphasized, the societal impact disappeared. When male stereotypes were explicitly activated, female participants acted against it by displaying more entrepreneurial intentions, while male participants lowered their intentions. When male stereotypes were implicitly activated, the opposite effect was found. No effect was found for either implicit or explicit activation of the female stereotype. In conclusion, implicitly or explicitly activating gender stereotypes leads to different results, because explicit activation elicits stereotype reactance, whereas implicit activation does not (Kray et al., 2001).

Like in the research of Gupta et al., (2008), the target of the stereotype activation can be one’s own group membership, called self-stereotype activation. Stereotype activation can

also target group members from another group, which is called other-stereotype activation (Wheeler & Petty, 2001). In the current study 'stereotype activation' refers to other-stereotype activation, as it concerns stereotype activation when evaluating a fictional female applicant. Although most studies about stereotype reactance investigated behavioural consequences of stereotyping of the self and this study investigated the behavioural consequences of stereotyping of another person, the underlying processes might be similar. However, explicit stereotype reactance towards another person has not been widely investigated yet (Wheeler & Petty, 2001). Previous research indicates that this process also takes place when it concerns other-stereotype reactance: It has been demonstrated that providing feedback on society's endorsement of gender stereotypes and implicit gender biases in selection procedures helps to increase the number of women that are hired in academic positions (Sheridan et al., 2010). Although the underlying process of this effect was not investigated, this could be an example of other-stereotype reactance. Thus, both the self and another person can be the target of stereotype activation, although the underlying process of the latter target is still under investigation. In conclusion, both the type of reminder and the target of the stereotype influence the impact of stereotype activation.

Describing the ideal applicant using gendered wording

As previously mentioned, stereotypes can also influence behaviour through linguistics. Gendered wording is such a linguistic influence. When looking at the effects of implicit gender bias on the position of women in (academic) top positions, the effect of linguistics on selection procedures is especially relevant. Words are associated with the male or female stereotype. For example, 'excellent' qualities like the words 'competitive' and 'dominant' are associated with the male stereotype, whereas words about 'supporting' qualities like 'support' and 'understand' are associated with the female stereotype (Kaatz et al., 2015). In evaluation forms for job applicants, similar gendered wording is also present (Van der Lee & Ellemers,

2015). The more male-dominated a career area is, the higher the presence of masculine wording in job advertisements, whereas the amount of feminine wording remains stable, regardless of the career area (Gaucher, Friesen & Kay, 2011). In the Netherlands, 76.2% of gendered written communication directed at applicants by the Netherlands Organization of Scientific Research (NWO) was related to masculine stereotypical traits (Van der Lee & Ellemers, 2015). Thus, gendered wording is present in applicant evaluation forms and its amount is influenced by gender.

Gendered wording in advertisements or applicant evaluation forms lead to expectations about the applicant. Applicants are expected to make an impression that is congruent with the gender of the wording used in the evaluation form or advertisement (Gaucher et al., 2011). Since gendered wording is linked to gender stereotypes, applicants are therefore indirectly expected to behave in a way that is congruent with the gender stereotype. When gendered wording is masculine while a job applicant is female, it leads to an incongruence between the applicant and the description of the ideal applicant. This is in favour of male applicants, since they appear to be a better fit with the male stereotype and therefore a better candidate. Gendered wording in job advertisements and evaluation forms therefore maintains gender inequality in (academic) top positions: Since these positions are dominated by men, the wording in job advertisements and evaluation forms is more often masculine and therefore in favour of other men (Gaucher et al., 2011).

Another factor influencing the impact of implicit bias in selection procedures is the type of wording in either traits or behaviours when describing an applicant. A job application often contains a description of the ideal applicant, which is usually described in terms of traits. A trait is seen as a part of a personality, whereas behaviour is flexible and less dependent on personality. By using behaviours to describe the ideal applicant, the perspective is changed from ‘traits that an applicant has to possess already’ to ‘performances that the

applicant has to be capable of at this job'. Traits are more strongly linked to stereotypes than behaviours, as stereotypes make assumptions about traits of a group of people, but people from that stereotyped group can still behave in a way that is incongruent with the stereotypical traits (Taris & Bok, 1998). In the Netherlands, a job advertisement contains on average 3.5 traits (Taris & Bok, 1998). Most of the traits mentioned in job advertisements are linked to the gender stereotype (Taris & Bok, 1998). The ideal applicant can therefore be described in gender-specific traits or behaviours in a job advertisement (Born & Taris, 2010). An example of description of a specific male stereotype trait versus behaviour would be: 'The applicant is excellent' versus 'the applicant behaves excellently'.

Born and Taris (2010) conducted research where they investigated the inclination of women and men to apply to jobs based on job advertisements, which contained a description in either traits or behaviours in either masculine or feminine wording. Women are more likely to react to a job advertisement for a job when the advertisement contains masculine wording in terms of behaviour than when it is described as traits. Also, women are more likely to react to a job advertisement when the description contains feminine wording in the form of traits, since women are more attracted to descriptions that are more similar to themselves. For men, no similar effects were found. This can be explained by a larger awareness of gender (and thus gender stereotypes) by women than by men (Born & Taris, 2010). This explanation is in line with findings that show that minorities are more aware of the lower status that usually accompanies minorities than majority members (Ellemers, Van Dyck, Hinkle & Jacobs, 2000). Presenting criteria as male stereotype behaviour thus interferes less with the female stereotype than when it is presented as male stereotype traits. Therefore, it elicits more inclined responses only by women (Born & Taris, 2010). Applying these findings to (academic) top positions, describing the ideal applicant in desired traits helps to maintain unequal gender distributions in (academic) top positions. This is because of the higher

frequency of male stereotype traits than female stereotype traits in job advertisements for (academic) top positions. This would increase the inclination of men and lower the inclination of women to respond to such a job advertisement, as it is more congruent with the male stereotype than with the female stereotype. However, describing the ideal applicant in desired behaviours makes gender equality in (academic) top positions more likely, because behaviours are less strongly linked to gender stereotypes (Born & Taris, 2010). Then, the perceived gap between the ideal applicant and the prospective female applicant is less based on gender stereotypes. In conclusion, based on previous research, gendered wording in selection procedures is often favourable for men and unfavourable for women, unless descriptions of the ideal applicant are given in behaviours instead of traits.

Current study

Previously discussed findings validate the impact of implicit gender bias on job application procedures. Both the problem and possible solutions are often not approached from an applicant evaluator's perspective. To gain insight in the impact of implicit gender bias from an evaluator's perspective, the current study focused on the role of gendered language and type of wording in the evaluation forms of job applications and reducing prejudiced behaviours towards women in selection procedures through feedback. In line with previous findings (Gaucher et al., 2011; Rudman & Kilianski, 2000), participants were expected to have a stronger implicit association between male names combined with 'excellent' words and female names combined with 'supporting' words, than with male names combined with 'supporting' words and female names combined with 'excellent' words (*Hypothesis 1*), due to implicit stereotyping. Based on other previous findings (Swim et al., 1995), a significant positive relationship was expected between implicit gender bias and modern sexism, and belief in a just world (*Hypothesis 2a).* Also, a negative relationship was expected between implicit gender bias and support for affirmative action *(Hypothesis 2b*).

Previous research has shown that providing general feedback on the effect of society's implicit endorsement of gender stereotypes on selection procedures promotes gender equality (Sheridan et al., 2010). Therefore, providing personal feedback on implicit gender bias was expected to also effectively diminish its impact on selection procedures, as it would also make the impact of implicit gender bias explicit. Due to the personal nature of the feedback, it would not be possible to think that implicit gender bias is not applicable to oneself. To gain further insight in the impact of providing feedback on one's implicit gender bias, gendered wording and type of wording on perceived overall suitability of the applicant, measures of perceived overall applicant suitability and stereotype activation were included. Previous findings on providing feedback on applicant rating showed that providing feedback affects people's applicant judgements compared to when no feedback was provided (Sheridan et al., 2010). Therefore, in the current study, to gain more insight in the process, personal feedback was provided. Based on research on stereotype activation (Gupta et al., 2008) and stereotype reactance (Kray et al., 2001) a mediation effect was expected between providing personal feedback, stereotype activation and perceived overall applicant suitability: Providing personal feedback was expected to be an explicit way of activating stereotypes, thereby increasing stereotype activation (*Hypothesis 3a*). No increased stereotype activation was expected when no feedback on implicit gender bias was given (*Hypothesis 3b*). Providing personal feedback in turn was expected to lead to stereotype reactance, as participants feel limited in their actions, and they could react against these perceived limitations (Kray et al., 2001). It was expected that stereotype reactance would be displayed by increased ratings on perceived overall applicant suitability when gendered wording is masculine, and the description in traits as traits are more strongly connected to stereotypes than behaviours (Taris & Bok, 1998). Since masculine wording, especially in trait descriptions, is not in line with a female applicant, this creates an incongruence between the applicant and the advertisement.

Therefore, masculine wording in the evaluation form was expected to lead to lower perceived applicant likeability compared to an evaluation form containing balanced wording, in line with previous findings (Heilman, 2001). In conclusion, it was expected that the increased stereotype activation would increase ratings of overall applicant suitability for participants using an evaluation form containing masculine wording in terms of traits (*Hypothesis 4a*). This effect was not expected for an evaluation form using descriptions in terms of behaviour or balanced wording (*Hypothesis 4b*).

Given the previous finding that masculine wording in advertisements creates expectations about traits of the ideal applicant that are in favour of men (Gaucher et al., 2011), a similar effect was expected when gendered language occurs in evaluation forms. Participants using an evaluation form containing masculine wording were expected to rate the applicant lower on overall applicant suitability when the ideal candidate was described in traits than when the ideal candidate was described in behaviours, but only when no feedback was given (*Hypothesis 5a*). No effect of type of wording was expected when balanced wording was used in the evaluation form (*Hypothesis 5b*).

Method

Participants and design

In total, 185 participants completed our experiment. All participants were required to be fluent in Dutch, since the entire experiment was in Dutch. Four participants were excluded from analyses as they had not correctly responded to our manipulation check, resulting in a total of $N = 181$ participants. The average age of the participants was $M = 20.8$ years ($SD = 2.3$) and 23.1% of participants was male, while 76.9% was female. Participants were recruited at Leiden University and the University of Applied Sciences Leiden, or signed up after viewing Facebook-advertisements or flyers. The experiment took about 30 minutes and the reward was one course credit or 3.50 Euros.

To investigate the effects of feedback, gendered wording and type of wording, a 2 (feedback: No feedback versus personal feedback on participant's implicit gender bias) X 2 (gendered wording: Masculine versus balanced wording) X 2 (type of wording: Traits versus behaviours) between-subjects experimental design was used. Participants were randomly assigned to one of the eight experimental groups.

Procedure

Participants were welcomed into the lab and signed the informed consent form. Then, each participant was placed in front of a computer in an individual lab cubicle. The participant was assigned to one of the conditions in a double-blind procedure. The experiment started with explaining the cover story: Participants were told that the subject of the study was selection procedures. At the start of the experiment, they were asked to first complete a categorization task. In reality, this was an Implicit Association Test (IAT; Greenwald, McGhee & Schwartz, 1998) to assess the implicit relationship between quality words and gender. This IAT was self-developed and applied to the Dutch context, and contained 7 blocks. The terms used in the IAT were selected from previous research (e.g., Van der Lee & Ellemers, 2015). In block one, participants were asked to categorize Dutch male and female names as fast as they could, like 'Jeroen' as a male name and 'Suzanne' as a female name. They used their left hand to categorize male names, while using their right hand to categorize female names. In block two, they were asked to categorize quality words as fast as possible as 'supporting' or 'excellent'. An example of a 'supporting' word is 'caring', while an example of an 'excellent' word is 'adventurous'. The left hand was used to categorize 'excellent' qualities, while the right hand was used to categorize 'supporting' qualities. Block three was a trial block and block four the actual test block in which participants were asked to categorize words and names congruent with the gender stereotypes: 'Excellent' words and male names had to be categorized by using the left hand, while 'supporting' words and female names had

to be categorized by using the right hand. In block five, participants were asked to categorize reverse-categorize Dutch names, thus categorizing male names with the right key instead of left and vice versa for female names. Block six was a trial block and block seven the actual test block in which names had to be sorted incongruently with gender stereotypes: Male names were categorized with 'supportive' words by using the right hand, while female names with 'excellent' words were categorized by using the left hand.

After completing the IAT, the participant either received no feedback about one's implicit associations concerning gender and careers, or pre-programmed feedback about their 'personal' implicit gender bias. The personal feedback told participants that the IAT showed a bias corresponding with the usual implicit associations between gender and careers: They were faster at making stereotype congruent combinations (male names with excellent qualities and female names with supporting qualities) than at making stereotype incongruent combinations (female names with excellent qualities and male names with excellent qualities). Important is that the so-called results were the same for all participants in the condition where personal feedback was given on participants' implicit gender bias. Participants placed in the condition of receiving no feedback also received the same pre-programmed message as others in that condition: They were only told they had reached the end of the categorization task.

After feedback manipulations, participants were introduced to the second part of the experiment. Participants were told that the experiment was meant to uncover the processes in the selection procedure for an advisor in the advisory board of Leiden University. Participants were then asked to judge a resume of a woman who applied for this position. Although participants were told that the resume was randomly selected from actual recent applications for this position, in reality it was fabricated and similar to the resume used by Moss-Racusin et al. (2012). Appendix A shows the original Dutch version of the resume that was used in the

current study. After reading the resume, participants received evaluation instructions containing descriptions of the ideal applicant in either traits or behaviours (for example, “the candidate is excellent” versus “the candidate displays behaviour that is excellent”). Also, descriptions of the ideal applicant contained either masculine or balanced wording (for example, “the candidate is excellent” versus “the candidate is caring” and “the candidate is organized”). Participants were asked to evaluate the applicant based on the descriptions of the ideal applicant. This way, gendered wording and type of wording manipulations were induced.

Then, participants completed the questionnaire containing the dependent measures. Manipulation checks were included, to check whether participants’ perception of the wording in the experiment matched the actual condition they were assigned to. Participants were asked whether they had rated a male or female applicant, to determine whether participants were aware that they rated a woman. They were also asked whether they had judged the applicant on traits and behaviours, and whether they felt like the application criteria were aimed towards men and women. Also, the participants were asked about demographics containing their gender, age, study (phase) and experience with evaluation procedures. Lastly, participants were debriefed, thanked for their participation and paid.[1]

[1] Participants participated in one of two sessions. In the first session, the experiment was done immediately after participants had already completed another experiment. In the second session, participants first completed our experiment before continuing another experiment, in another laboratory than the participants from session one. As the circumstances in which the experiment was conducted varied based on the session, the participants’ session was considered a possible covariate and therefore tracked for each participant. Results are described in Appendix B.

Dependent variables

Gender equality measures. Several questionnaires were used to measure the relationship between implicit gender bias and participants' views on gender equality. The scale used to measure modern sexism contained nine items (adapted from Swim et al., 1995; e.g., "Discrimination against women is no longer a problem in The Netherlands", "It is easy to understand the rage of women's groups in The Netherlands"; α = .80). The scale used to measure support for affirmative action contained 4 items (adapted from Sidanius, Pratto & Bobo, 1996; e.g., "Affirmative action is unfair towards men", "Affirmative action is needed to reduce the difference in opportunities between men and women"; α = .66). The scale used to measure belief in a just world contained five items (adapted from Stroebe, Postmes, Täuber, Stegeman & John, 2015; α = .27). Due to the low inter-item reliability of these five items, the construct could not be analysed. An analysis of the separate items is shown in Appendix C. These items were: "To what extent do you believe the world currently is a righteous place?", "To what extent do you think the world should be a righteous place?", To what extent do you believe affirmative action is necessary?", "To what extent do you believe people get what they deserve?" and "To what extent do you believe that forms of discrimination are inevitable?".

Overall applicant suitability measures. Also, several questionnaires were used to measure the effects of feedback on implicit gender bias, gendered wording and type of wording on perceived overall applicant suitability. Participants could indicate their answers to items of all scales by answering on a seven-point scale (with 1 meaning 'not at all' and 7 meaning 'totally'). The scale used to measure applicant rating contained 3 items (Moss-Racusin et al., 2012; e.g., "To what extent do you judge the candidate to be suitable?", "To what extent do you judge the candidate to be positive?"; α = .84). The scale used to measure perceived competence of the applicant contained 3 items (Moss-Racusin et al., 2012; e.g., "To

what extent do you judge the applicant to be qualified?", "To what extent do you judge the applicant to be competent?"; α = .86). The scale used to measure willingness to employ the applicant contained two items (Moss-Racusin et al., 2012; e.g., "How likely is it that you would invite this applicant to an interview?", "How likely is it that you would invite the applicant for a job interview for this position?"; α = .93). The scale used to measure perceived fit of the applicant with the job contained 2 items (Eliezer, Major & Mendes, 2010; e.g., "To what extent do you think the applicant would fit with the job?", "To what extent do you feel like the applicant belongs at this job?"; r = .83). The scale used to measure willingness to mentor the applicant contained two items (Moss-Racusin et al., 2012; e.g., "How likely is it that you would aid this applicant if she had difficulty with the job's activities?", "How likely is it that you would encourage the applicant to stay in the advisory board, if the applicant considers leaving?"; α = .65). The scale used to measure the applicant's likeability contained 3 items (Eagly & Mladinic, 1994; e.g., "To what extent do you think you would like the applicant?", "To what extent do you want to get to know the applicant?"; α = .73).

Stereotype activation measures. A logarithm was used to calculate participants' IAT scores. A positive score indicated that 'excellent' qualities were more rapidly combined with male names than with female names, and vice versa for 'supporting qualities (Greenwald et al., 1998). To determine whether our gendered wording has led to stronger connections to stereotypes, a scale was developed to measure stereotype activation based on Steele & Aronson's 'stereotype avoidance' scale, containing 6 items (1995; e.g., "To what extent do you think this applicant enjoys football?", "To what extent do you think this applicant loves to dance?"; α = .67). To measure implicit stereotype activation, a word completion task was also included. Participants were shown six incomplete words, and were asked to complete them. Participants' answers were coded into four categories: Words that are connected to the female stereotype (F), words that are connected to the male stereotype (M), words that are either

gender stereotype neutral or ambiguous in their meaning and/or (N), and non-words (NON). Not all categories were an option for all items. For example, for the task ".US": Participants could either complete this with the female stereotype congruent word "zus", or gender stereotype neutral words such as "bus", while " .OK" could be completed with the female stereotype congruent word "rok", the male stereotype congruent word "bok", or gender stereotype neutral words such as "sok". Appendix D shows the coding for the given answers on all five items of the word completion task.

Results

Data were analysed with ANOVAs treating feedback, gendered wording and type of wording as independent variables.

Manipulation checks

Manipulation checks were conducted to determine whether the manipulations of type of wording and gendered wording were perceived by participants in the way they were intended. No manipulation check for feedback was included, because attention was explicitly drawn to the contents of this manipulation, unlike the other manipulations. First, the manipulation of type of wording was effective for behavioural descriptions, $F(1, 278) = 5.91$, $p = .02$, $\eta_p^2 = .02$. Participants reading descriptions of the ideal applicant in behaviours judged the applicant more on behaviours ($M = 4.83$, $SD = 1.18$) than did participants reading trait descriptions ($M = 4.40$, $SD = 1.54$). However, for the extent to which participants judged the applicant based on traits, the manipulation was not effective, $F(1, 278) = .08$, $p = .78$, $\eta_p^2 < .001$. Whether participants judged the applicant based on traits ($M = 5.19$, $SD = 1.26$) or behaviours ($M = 5.15$, $SD = 1.22$), has not influenced the perceived extent to which participants judged the applicant based on traits. In other words, the manipulation of type of wording was partially successful.

For the manipulation of gendered wording, the manipulation was effective for masculine wording, $F(1, 278) = 2.56$, $p = .11$, $\eta_p^2 = .01$. Participants in the 'masculine wording'-condition had not felt to a greater extent like the evaluation criteria were directed towards men ($M = 3.13$, $SD = 1.55$) than participants in the 'balanced wording'-condition ($M = 2.83$, $SD = 1.54$), thereby confirming the implicit nature of the manipulation. Lastly, the manipulation was not effective for balanced wording, $F(1, 278) = 4.15$, $p = .04$, $\eta_p^2 = .02$. Participants in the 'balanced wording'-condition more often felt like the evaluation criteria were directed towards women ($M = 3.77$, $SD = 1.76$) than participants in the 'masculine wording'-condition ($M = 3.34$, $SD = 1.55$). Thus, this manipulation was not implicit. In conclusion, the manipulation of gendered wording was also partially successful. Although following results based on these manipulations should therefore be interpreted with some caution, this does not mean that the manipulations were unsuccessful in manipulating the dependent variable.

Implicit gender bias and gender equality measures

IAT. Due to implicit stereotypes, participants were expected to have a stronger association with male names combined with 'excellent' words and female names combined with 'supporting' words, than with male names combined with 'supporting' words and female names combined with 'excellent' words (*Hypothesis 1*). To investigate this, a *t*-test was computed. The mean differed significantly from 0, $t(280) = 11.05$, $p < .001$, indicating bias. The mean was positive ($M = 0.29$, $SD = 0.43$), which shows that participants were better at making stereotype congruent combinations (male names with excellent qualities and female names with supporting qualities) than at making stereotype incongruent combinations (female names with excellent qualities and male names with supporting qualities). Thus, *Hypothesis 1* is accepted.

Gender equality measures. *Hypothesis 2a* stated that a significant positive relationship was expected between implicit gender bias and modern sexism and belief in a just world. *Hypothesis 2b* stated that a negative relationship was expected between implicit gender bias and support for affirmative action. To test these hypotheses, Pearson correlations were computed between implicit bias and support for affirmative action and modern sexism. A significant negative correlation was found between support for affirmative action and implicit gender bias, $r = -.14$, $p = .02$: The stronger participants' implicit gender bias, the less supportive of affirmative action they were. A significant positive correlation was found between modern sexism and implicit bias, $r = .15$, $p = .01$: The stronger participants' implicit gender bias, the stronger participants displayed modern sexism. In conclusion, *Hypothesis 2a* and *2b* were confirmed for modern sexism and support for affirmative action.

The correlations between implicit bias and modern sexism and support for affirmative action were not expected to be influenced by the manipulations. To confirm this, several 2x2x2 ANOVAs were done with feedback, gendered wording and type of wording as the independent variables, and modern sexism and support for affirmative action as the dependent variables. Results for support for affirmative action showed no significant three-way interaction effect, $F(1, 272) = .68$, $p = .41$, $\eta_p^2 = .003$, just like modern sexism, $F(1, 272) = 1.79$, $p = .18$, $\eta_p^2 = .01$. Thus, the correlations between implicit bias and these dependent variables was not influenced by the manipulations.

Overall applicant suitability measures

A 2x2x2 ANOVA was done on overall applicant suitability measures to investigate the effects of the independent variables feedback, gendered wording and type of wording. Three-way interactions were reported at all times, two-way interactions and main effects were reported if significant. Participants using an evaluation form containing masculine wording were expected to rate the applicant lower on overall applicant suitability when the ideal

candidate was described in traits than when the ideal candidate was described in behaviours, but only when no feedback was given (*Hypothesis 5a*). No effect of type of wording was expected when balanced wording was used in the evaluation form (*Hypothesis 5b*).

No significant three-way interaction effect on applicant rating was found, $F(1, 272) = 1.33$, $p = .25$, $\eta_p^2 = .01$. This disconfirmed *Hypothesis 5a* and confirmed *Hypothesis 5b*.

No significant three-way interaction effect on perceived competence was found, $F(1, 272) = 1.97$, $p = .16$, $\eta_p^2 = .01$. This disconfirmed *Hypothesis 5a* and confirmed *Hypothesis 5b.* Unexpectedly, a marginally significant two-way interaction between feedback and type of wording on perceived competence was found, $F(1, 272) = 2.78$, $p = .10$, $\eta_p^2 = .01$. Bonferroni post hoc tests revealed that perceived competence was higher, $F(1, 272) = 4.07$, $p = .05$, $\eta_p^2 = .02$, when personal feedback was given and the ideal candidate was described in trait descriptions ($M = 5.01$, $SD = 1.00$) than in descriptions in behaviours ($M = 4.63$, $SD = 0.96$).

No significant three-way interaction effect on willingness to employ was found, $F(1, 272) = 1.95$, $p = .16$, $\eta_p^2 = .01$. This disconfirmed *Hypothesis 5a* and confirmed *Hypothesis 5b.* An unexpected, significant two-way interaction effect between feedback and type of wording on willingness to employ was found, $F(1, 272) = 4.34$, $p = .04$, $\eta_p^2 = .02$. Bonferroni post hoc tests revealed that willingness to employ increased, $F(1, 272) = 4.10$, $p = .04$, $\eta_p^2 = .02$, when descriptions in behaviours were given ($M = 4.94$, $SD = 1.16$) than when trait descriptions were given ($M = 4.68$, $SD = 1.05$), but only when no feedback was given. Willingness to employ also increased, $F(1, 272) = 3.20$, $p = .075$, $\eta_p^2 = .01$, when no feedback was given ($M = 4.94$, $SD = 1.10$) than when personal feedback was given ($M = 4.50$, $SD = 1.29$), but only when descriptions were given in terms of behaviour.

A marginally significant three-way interaction effect on perceived fit was found, $F(1, 272) = 3.73$, $p = .06$, $\eta_p^2 = .01$. However, Bonferroni post-hoc test revealed that the expected direction of the three-way interaction effect as described in *Hypothesis 5a* was not found, $F(1,$

272) < 0.01, $p = .99$, $\eta_p^2 < .01$. These results confirmed *Hypothesis 5b*. Bonferroni post hoc tests, however, also revealed that giving personal feedback led to more perceived fit when descriptions were given in traits and masculine wording was used, $F(1, 272) = 4,49$, $p = .04$, $\eta_p^2 = .02$ ($M = 5.06$, $SD = .89$), than when no feedback was given in male stereotype trait descriptions ($M = 4.35$, $SD = .94$).

A marginally significant three-way interaction effect on willingness to mentor was found, $F(1, 272) = 3.14$, $p = .08$, $\eta_p^2 = .01$. Bonferroni post hoc tests revealed that willingness to mentor the applicant increased, $F(1, 272) = 4.87$, $p = .03$, $\eta_p^2 = .02$, when descriptions in behaviours were given ($M = 5.11$, $SD = 0.91$) than when trait descriptions were given ($M = 4.44$, $SD = 1.52$), but only when no feedback was given and masculine wording was used. These results confirm *Hypothesis 5a* and *Hypothesis 5b*. Bonferroni post hoc tests also revealed that giving personal feedback marginal significantly increased willingness to mentor, $F(1, 272) = 3.11$, $p = .08$, $\eta_p^2 = .01$, when descriptions of the ideal candidate were given in traits and masculine wording was used ($M = 4.98$, $SD = 1.13$), compared to descriptions given in male stereotype traits with no feedback ($M = 4.44$, $SD = 1.52$). An unexpected, significant main effect of type of wording on the willingness to mentor was revealed, $F(1, 272) = 3.96$, $p = .05$, $\eta_p^2 = .01$. When descriptions were given in terms of behaviour, willingness to mentor was greater ($M = 5.11$, $SD = 0.94$) than when traits were used to describe the ideal candidate ($M = 4.88$, $SD = 1.19$).

No significant three-way interaction effect on applicant likeability was found, $F(1, 272) = 1.95$, $p = .16$, $\eta_p^2 = .01$. This disconfirmed *Hypothesis 5a* and confirmed *Hypothesis 5b.*

Stereotype activation. An ANOVA was also done for the effect of feedback on stereotype activation. Expected was that providing personal feedback about implicit gender bias increases stereotype activation (*Hypothesis 3a*). No increased stereotype activation was

expected when only no feedback on implicit gender bias was given (*Hypothesis 3b*). No significant three-way interaction effect was found, $F(1, 272) = 1.17$, $p = .28$, $\eta_p^2 = .004$. However, a main effect of feedback was found, $F(1, 272) = 8.69$, $p = .003$, $\eta_p^2 = .03$. When personal feedback was given, stereotype activation was higher ($M = 3.94$, $SD = 0.63$) than when no feedback was given ($M = 3.64$, $SD = 0.79$). This supports *Hypotheses 3a* and *3b*.

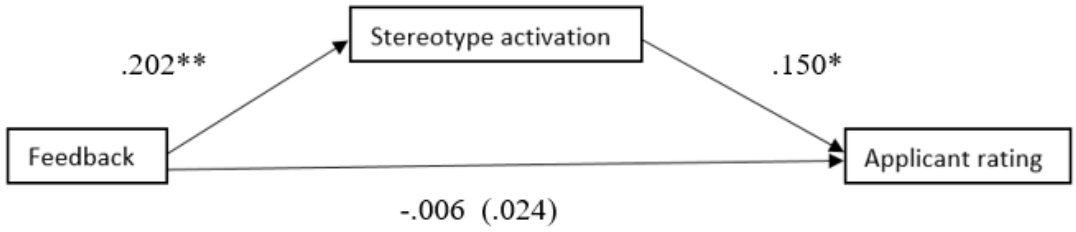
A word completion task was also done to measure the effect of feedback on stereotype activation. The answers to the word completion task was analysed with a chi-square test. The results show that there was a significant effect of feedback on stereotype activation for the first item of the word completion task, $\chi(3) = 10.15$, $p = .02$. More gender stereotype neutral answers were given when participants received no feedback ($n = 117$) than when personal feedback was given ($n = 92$). More masculine stereotyped answers were given when participants received no feedback ($n = 35$) than when personal feedback was given ($n = 19$). More feminine stereotyped answers were given when participants received no feedback ($n = 9$) than when personal feedback was given ($n = 6$). These findings also support *Hypotheses 3a* and *3b*. Unexpected was the finding that gendered wording had a significant effect on the gender stereotyped answers on the second item of the word completion task, $\chi(2) = 6.52$, $p = .04$. More gender stereotype neutral answers were given when participants read a description in balanced wording ($n = 166$), than when a description was given in masculine wording ($n = 91$). Less gender stereotype neutral answers were given when participants read a description in masculine wording ($n = 13$), than when participants read a description in balanced wording ($n = 9$).

Correlations and mediation analyses. It was expected that the increased stereotype activation would in turn increase ratings of overall applicant suitability in participants using an evaluation form containing masculine wording in terms of traits (*Hypothesis 4a*). This effect was not expected for an evaluation form using descriptions in terms of behaviour or

balanced wording (*Hypothesis 4b*). First, to investigate this expected mediation, Pearson correlations between stereotype activation and the overall applicant suitability were computed. Due to the lack of significant effects on most word completion task items, no analysis was done on the effect of feedback on the word completion task. Stereotype activation correlates with applicant rating ($r = .15$, $p < .05$), perceived competence ($r = .18$, $p < .01$), willingness to employ ($r = .13$, $p < .05$), perceived fit ($r = .14$, $p < .05$), willingness to mentor ($r = .23$, $p < .001$) and applicant's likeability ($r = .26$, $p < .001$).

These significant correlations indicate that, as expected, stereotype activation mediates the effect of feedback on the applicant rating, perceived competence, willingness to employ, willingness to mentor, perceived fit and likeability. Step one of the mediation analysis for applicant rating showed that feedback is not a predictor of applicant rating, B = -.006, $t(278) = -.10$, $p = .92$. In step two, the regression analysis performed with feedback as the independent variable and stereotype activation as the dependent variable showed that feedback was a significant predictor of stereotype activation, B = .202, $t(278) = 3.44$, $p = .001$. Step three showed that stereotype activation was shown to be a significant predictor of applicant rating, B = .150, $t(278) = 2.47$, $p = .01$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and applicant rating remained non-significant, B = .024, $t(278) = .400$, $p = .69$. In other words, personal feedback increases stereotype activation, and in turn, stereotype activation led to higher applicant rating. The results of the Sobel test on the unstandardized coefficients were significant, and therefore confirmed the mediation, $z = 2.01$, $p = .04$ (Soper, n.d.). Figure 1 shows a visual representation of this mediation.

Similar mediation analyses were done for perceived competence. It was found that feedback is not a predictor of perceived competence, B = -.013, $t(278) = -.22$, $p = .82$. Step two showed that feedback was again a significant predictor of stereotype

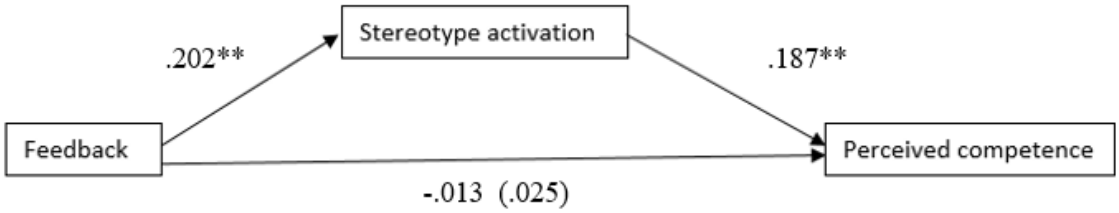


*$p < .005$

**$p < .005$

Figure 1. Standardized regression coefficients for the relationship between feedback and applicant rating, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and applicant rating after correcting for stereotype activation.

activation, B = .202, $t(278) = 3.44$, $p = .001$. In step three, stereotype activation was shown to be a significant predictor of perceived competence, B = .187, $t(278) = 3.10$, $p = .002$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and perceived competence remained non-significant, B = .025, $t(278) = .41$, $p = .68$. In other words, personal feedback increases stereotype activation, and in turn, stereotype

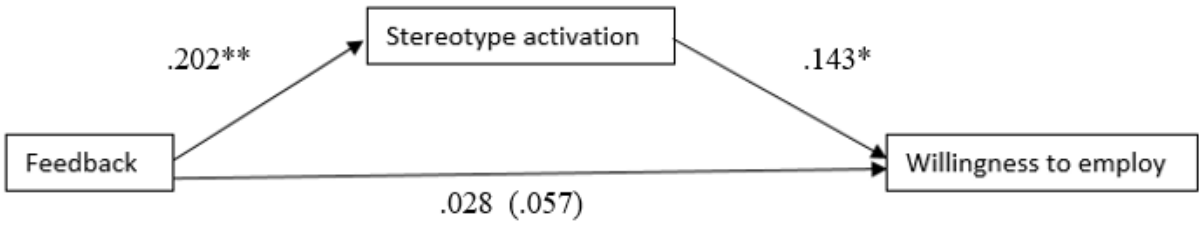


**$p < .005$

Figure 2. Standardized regression coefficients for the relationship between feedback and perceived competence, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and perceived competence after correcting for stereotype activation.

activation led to increased perceived competence of the applicant. The results of the Sobel test on the unstandardized coefficients confirmed the mediation, $z = 2.31$, $p = .02$ (Soper, n.d.). Figure 2 shows a visual representation of this mediation.

Also, a mediation analysis was done for willingness to employ. Step one showed that despite the significant correlations between feedback and willingness to employ, feedback is not a predictor of willingness to employ, B = .028, $t(278) = .47$, $p = .64$. Step two again showed that feedback was a significant predictor of stereotype activation, B = .202, $t(278) = 3.44$, $p = .001$. In step three, stereotype activation was shown to be a significant predictor of willingness to employ, B = .143, $t(278) = 2.358$, $p = .019$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and willingness to employ remained non-significant B = .057, $t(278) = .95$, $p = .35$. In other words, personal feedback increases stereotype activation, and in turn, stereotype activation led to greater willingness to employ the applicant. The results of the Sobel test on the unstandardized coefficients were marginally significant, and therefore confirmed the mediation, $z = 1.95$, $p = .05$ (Soper, n.d.). Figure 3 shows a visual representation of this mediation.

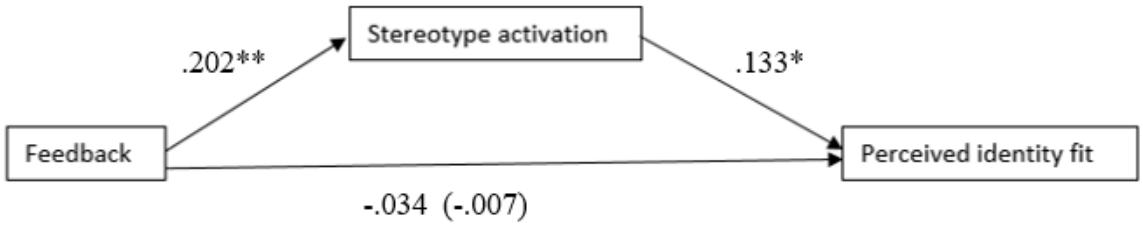


*$p < .005$

**$p < .005$

Figure 3. Standardized regression coefficients for the relationship between feedback and willingness to employ the applicant, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and willingness to employ after correcting for stereotype activation.

A mediation analysis was also done for perceived fit. It was found that despite the significant correlations between feedback and perceived fit, feedback is not a predictor of perceived fit, B = -.034, $t(278) = -.565$, $p = .57$. Feedback was a significant predictor of stereotype activation, B = .202, $t(278) = 3.44$, $p = .001$. Stereotype activation was shown to be a significant predictor of perceived fit, B = .133, $t(278) = 2.19$, $p = .03$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and perceived fit remained non-significant, B = -.007, $t(278) = -.11$, $p = .91$. In other words, personal feedback increases stereotype activation, and in turn, stereotype activation led to increased perceived fit of the applicant. The results of the Sobel test on the unstandardized coefficients were marginally significant, and therefore confirmed the mediation, $z = 1.84$, $p = .07$ (Soper, n.d.). Figure 4 shows a visual representation of this mediation.



*$p < .005$

**$p < .005$

Figure 4. Standardized regression coefficients for the relationship between feedback and perceived fit, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and perceived fit after correcting for stereotype activation.

A similar mediation analysis was done for willingness to mentor. Step one showed that despite the significant correlations between feedback and willingness to mentor, feedback is not a predictor of willingness to mentor, B = -.053, $t(278) = -.89$, $p = .37$. Step two again remains the same in this analysis: The regression analysis performed with feedback as the independent variable and stereotype activation as the dependent variable showed that

feedback was a significant predictor of stereotype activation, B = .202, $t(278) = 3.44$, $p = .001$. In step three, stereotype activation was shown to be a significant predictor of willingness to mentor, B = .232, $t(278) = 3.89$, $p < .001$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and willingness to mentor remained non-significant B = -.007, $t(278) = -.11$, $p = .91$. In other words, feedback increases stereotype activation, and in turn, stereotype activation led to increased willingness to mentor the applicant. The results of the Sobel test on the unstandardized coefficients were significant, and therefore confirmed the mediation, $z = 2.58$, $p = .01$ (Soper, n.d.). Figure 5 shows a visual representation of this mediation.

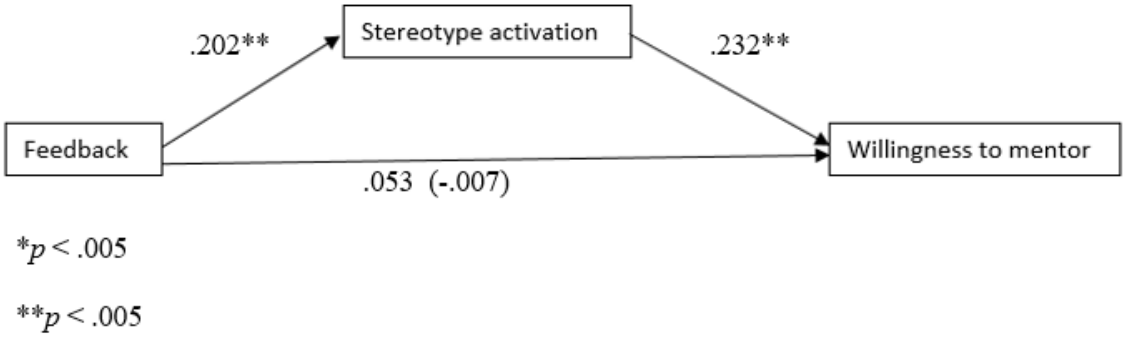
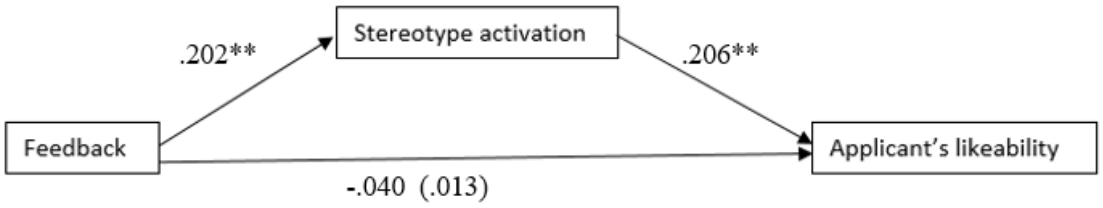


Figure 5. Standardized regression coefficients for the relationship between feedback and willingness to mentor the applicant, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and willingness to mentor after correcting for stereotype activation.

Lastly, a mediation analysis was done for applicant's likeability. Step one showed that despite the significant correlations between feedback and applicant's likeability, feedback is not a predictor of applicant likeability, B = -.040, $t(278) = -.67$, $p = .51$. Again, step two showed that feedback was a significant predictor of stereotype activation, B = .202, $t(278) = 3.44$, $p = .001$. In step three, stereotype activation was shown to be a significant predictor of applicant likeability, B = .206, $t(278) = 4.40$, $p < .001$. Step four indicated that when controlled for stereotype activation, the relationship between feedback and applicant

likeability remained non-significant B = .013, $t(278) = .21$, $p = .83$. In other words, feedback increases stereotype activation, and in turn, stereotype activation led to increased perceived applicant likeability. The results of the Sobel test on the unstandardized coefficients were significant, and therefore confirmed the mediation, $z = 2.71$, $p = .01$ (Soper, n.d.). Figure 6 shows a visual representation of this mediation. In conclusion, providing personal feedback lead to increased overall applicant suitability, as mediated by stereotype activation.



*$p < .005$

**$p < .005$

Figure 6. Standardized regression coefficients for the relationship between feedback and the applicant's likeability, mediated by stereotype activation. Shown in parentheses is the standardized regression coefficient between feedback and the applicant's likeability after correcting for stereotype activation.

Discussion

Women currently remain underrepresented in (academic) top positions (UN Women, 2016a/2016b). Implicit gender bias influences behaviour on the work floor, for example during selection procedures. This helps to maintain gender inequality (Moss-Racusin et al., 2012; Steinpreis et al., 1999), while gender equality is needed to optimally employ women's leadership capacities and ambitions (Steinpreis et al., 1999). The current study assessed the relationship between implicit gender bias and support for gender equality, and the effects of feedback about implicit gender bias, gendered wording and type of wording on implicit gender bias in selection procedures. The current study used literature on stereotype activation (Gupta et al., 2008; Wheeler & Petty, 2001), stereotype reactance (Kray et al., 2001) and

feedback (Sheridan et al., 2010) to find a way to decrease the extent to which implicit gender bias influences selection procedures in (academic) top positions. Previous literature on gendered wording (Gaucher et al., 2011; Kaatz et al., 2015; Van der Lee & Ellemers, 2015) and type of wording (Born & Taris, 2010; Taris & Bok, 1998) was applied to selection procedures in the current study, but was also used to uncover other solutions in battling implicit gender bias in selection procedures.

Implicit gender bias

The IAT on gender bias confirms that implicit gender bias influences selection procedures: Due to implicit stereotypes, individuals have a stronger association with male names combined with 'excellent' words and female names combined with 'supporting' words, than with male names combined with 'supporting' words and female names combined with 'excellent' words (*Hypothesis 1).* Implicit gender bias has consequences for the way the world is perceived: Implicit gender bias is found to be negatively correlated with support for affirmative action (*Hypothesis 2b*), while modern sexism is positively correlated to implicit gender bias (*Hypothesis 2a*).

Thus, the current study demonstrates the need for an intervention in selection procedures, as the relationship between implicit gender bias, support for affirmative action and modern sexism could help to maintain gender inequality in (academic) top positions: Lower support for affirmative action prevents developments in selection procedures promoting the presence of women in (academic) top positions, while modern sexism helps to maintain covert discrimination against women in (academic) top positions. The current findings are in line with previous literature which has shown that implicit gender bias influences selection procedures: Both men and women are more likely to hire a male applicant than female applicant and perceive male applicants as more competent than female applicants (Moss-Racusin et al., 2012; Steinpreis et al., 1999). Evaluators endorse modern

sexism by explicitly rejecting sexism, but implicitly giving a preferential treatment to men when this is justifiable by ascribing it to another factor than gender (Dovidio & Gaertner, 1996). Affirmative action is needed to create gender equality in (academic) top positions, as currently this is far from being achieved (Swim et al., 1995; UN Women, 2016a/2016b). Also, given the relationship between implicit gender bias and affirmative action, increasing support for affirmative action also entails a weaker perceived association between men and excellent qualities, and women and supporting qualities. For modern sexism, the effect is the other way around: Decreasing modern sexism entails a weaker implicit gender bias. Thus, increasing support for affirmative action and diminishing modern sexism is beneficial for women in (academic) top positions in a twofold way.

One way through which the influence of implicit gender bias can be reduced, is through providing personal feedback on implicit gender bias: It is found that providing individuals with personal feedback increases stereotype activation (*Hypotheses 3a* and *3b*). In turn, stereotype activation mediates between providing personal feedback and overall applicant suitability when an evaluation form containing masculine wording in terms of traits is used (*Hypotheses 4a* and *4b*), meaning that personal feedback leads to increased stereotype activation, which in turn leads to increased overall applicant suitability.

Explicit personal feedback on implicit gender bias increases stereotype activation, because feedback has made the gender stereotypes underlying implicit gender bias explicit, resulting in increased awareness of gender stereotypes. Then, people seemingly react against the feedback, possibly wanting to demonstrate that their implicit gender bias does not influence their behaviour. Current findings on explicit other-stereotype activation are similar to the reactions given in other research investigating explicit self-stereotyping (Gupta et al., 2008; Kray et al., 2001; Wheeler & Petty, 2001). Because of the focus on other-stereotype activation by providing feedback and other-stereotype reactance instead of self-stereotype

activation and self-stereotype reactance like most previous research has done, the current study is an important contribution to the research on stereotype activation and stereotype reactance (Gupta et al., 2008; Kray et al., 2001; Wheeler & Petty, 2001). Positively influencing the applicant evaluation of female applicants in the way that is currently demonstrated, is a good start in removing the impact of implicit gender bias from selection procedures from the evaluator's perspective, instead of trying to influence (prospective) applicants and the way they view an advertisement (Born & Taris, 2010). The impact of feedback also has important practical implications: Increasing perceived applicant rating and competence, and willingness to employ female applicants is crucial in achieving gender equality in (academic) top positions. Gender equality benefits society as a whole, since women's leadership skills and ambitions are then optimally used and this may inspire other women to aspire an (academic) top position as well.

Gendered wording

Regarding gendered wording and type of wording, it was hypothesised that individuals using an evaluation form containing masculine wording would rate the applicant lower on overall applicant suitability when the ideal candidate was described in traits than when the ideal candidate was described in behaviours, but only when no feedback was given (*Hypothesis 5a*). Findings only support this hypothesis for willingness to mentor. There is no effect of type of wording when balanced wording is used in the evaluation form (*Hypothesis 5b*). The limited confirmation of expectations about the effect of gendered wording is not in line with previous research which has shown that 76.2% of gendered wording used in communications targeted at (prospective) applicants at the Netherlands Organization of Scientific Research (NWO), is related to male stereotype traits (Van der Lee & Ellemers, 2015). Women's applications are therefore more likely to be perceived as incongruent with the ideal candidate than men's applications (Taris & Bok, 1998).

Also, the effect of gendered wording was expected to differ when descriptions in traits versus behaviours were used. Previous research has demonstrated that using behavioural descriptions of the ideal candidate instead of trait descriptions in a job advertisement, with the description being in masculine wording, increased the inclination of women to respond to that job advertisement. The same research found no effect of descriptions in traits versus behaviours on application inclination when feminine wording was used (Born & Taris, 2010). A possible explanation for the limited confirmation of expectations of the effect of gendered wording and type of wording in the current study could be that the processes preceding application intentions and evaluating are not comparable after all. During previous research, type of wording influenced application intentions (Born & Taris, 2010), while during the current research the type of wording was used to evaluate an applicant. Based on the current findings, it is speculated that judging your own perceived overall suitability when reading a job advertisement is more directly influenced by type of wording than evaluating an applicant's perceived overall suitability.

Since current findings might indicate that the hypothesized effects of type of wording might not be applicable to evaluation forms, the confirmation of the lack of effect of type of wording when balanced wording is used in the evaluation form should be interpreted with caution. This expected lack of effect could also be the result of another process underlying the applicant evaluations, instead of the hypothesized processes based on previous research. The practical implications of the findings on gendered wording and type of wording are that using descriptions in behaviours instead of trait descriptions when masculine wording is used, might not be the answer to combatting gender inequality in (academic) top positions. Using balanced wording instead of masculine wording might be promising, but the current findings are inconclusive.

Additional ways to reduce gender bias

The current study also reveals unexpected other ways to reduce gender bias. An unexpected finding is that perceived competence of the applicant increases if the evaluator receives personal feedback and the descriptions are in traits, regardless of gendered wording in the evaluation form. An explanation could be that ‘competence’ and ‘qualified’ are masculine words (Eagly & Mladinic, 1994). Therefore, personal feedback might cause a reaction against these words in the questionnaire about perceived competence, but only when traits were previously used, as traits are more strongly connected to stereotypes than behaviours (Taris & Bok, 1998). This in turn increases perceived competence. A practical implication of this finding is that not only descriptions of the ideal applicant should use balanced wording: All language used to judge an applicant should also contain balanced wording instead of masculine wording. This way, behaviour based on gender stereotyping during selection procedures becomes less likely, thereby helping to achieve gender equality in (academic) top positions.

Another unexpected finding is that evaluators are more willing to employ the applicant if the ideal candidate is described in behaviours instead of traits when no feedback is given, regardless of gendered wording. Similarly, willingness to mentor also increases when descriptions in behaviours instead of trait descriptions are used, regardless of feedback and gendered wording. A possible explanation could be that the willingness to employ and mentor are not perceived in the same way as the other measures: These two measures require (either employing or mentoring) behaviour from the evaluator, unlike the other overall applicant suitability measures. When an applicant is hired based on the perceived fit with a description in behaviours, the evaluator judges the sufficiency of the applicant’s behaviour on the work floor. Then, if the applicant’s behaviour is insufficient, the evaluator is more willing to mentor the applicant, as behaviours are perceived to be easier to change than traits (Born & Taris, 2010). Therefore, it also possible that it becomes easier to hire an applicant, because there is

less risk of hiring an unsuitable applicant due to the changeable nature of the applicant's behaviours at work compared to traits. Thus, using descriptions in behaviours in applicant evaluation forms could benefit gender equality in (academic) top positions, as it increases the willingness to employ women.

Lastly, perceived fit unexpectedly increases when personal feedback is given and descriptions are given in traits, regardless of gendered wording. It could be the case that the wording in the personal feedback addresses the bias about the suitability of men and women in careers, leading the evaluator to make an explicit connection between one's bias and perceived fit, then aiming not to have stereotypical traits influence the evaluation process. However, no measures about active commitment to countering implicit gender bias when evaluating the applicant were included in the current study. Therefore, this explanation remains purely speculative and should be investigated by future research. In conclusion, these unexpected findings provide unforeseen, but useful other ways to prevent implicit gender bias in selection procedures.

Future research and limitations

Future research on the effect of feedback could replicate the feedback manipulation, but add a measure about stereotype reactance. This would clarify the thought process after explicit stereotype activation. Future research could base questions on Dillard and Shen's (2005) measure of reactance, but adapt it to measure stereotype reactance. Also, future research could do a two-part study where a break exists between feedback and the overall applicant suitability measures, to measure how long it takes for the effect of feedback on stereotype activation to diminish. This could give insight in the real-life application of providing feedback to evaluators and how frequent that would need to happen to be effective. Lastly, future research concerning the application of feedback could include information about the effects of implicit gender bias on support for affirmative action and modern sexism

in the feedback. This could increase the awareness of the consequences of implicit gender bias, and therefore possibly create a positive attitude towards affirmative action through reactance (Dillard & Shen, 2005). In turn, this would help to achieve gender equality in (academic) top positions.

A limitation of the study is the partially successful manipulation of type of wording. Future research could manipulate the perceptions of traits and behaviours in a different and more successful way. For example, by priming individuals with behaviours before providing descriptions in traits or behaviours, rather than only providing descriptions in behaviours or traits in the evaluation form. Then, more attention is subconsciously drawn to the manipulation. A masked-priming task (Forster & Davis, 1984), where example behaviours (or traits) are masked with a meaningless pattern of characters would be suitable.

The manipulation of gendered wording was also only partially successful and therefore a limitation of this study. Implicit gender bias is implicit, so it was intended that individuals were unaware of gendered wording. Yet, individuals are aware that balanced wording is directed towards women. The partial success of this manipulation could provide more information about the perception of gendered wording by evaluators: The results of the manipulations could indicate that men are also more aware of female than male work floor stereotypes, whereas previous research suggests that only women are more aware of female stereotypes than men of male stereotypes (Born & Taris, 2010; Ellemers et al., 2000). This implies that during future research that investigates effects of implicit gender bias on evaluation procedures, awareness of female stereotypes should be suppressed to create comparable manipulations. Future research could weaken awareness by creating subconscious associations between women and careers, in turn making the association between women and supporting qualities less salient. Creating this new association could be done with a masked-priming task (Forster & Davis, 1984), where combinations of women and careers are masked

with a meaningless pattern of characters. To investigate the effects of this masked-priming task on the experiment, a control condition should be included in which the priming of women and careers is replaced by meaningless words. This could also create insight in the possibility of masked-priming as a way to reduce the impact of implicit gender bias on selection procedures, and thereby thus contribute to gender equality in (academic) top positions. However, these suggestions are initially experimental in nature to discover the psychological possibilities for suppression of awareness of the female stereotype. Only if future experimental research confirms this possibility, practical implications for organizations could be investigated.

Another limitation of this study is the effect of experimental session on participants. Half of the participants had already completed another experiment, whereas the other half of participants had not. Fatigue in participants might have caused the perceived overall applicant suitability to be lower when participants had already completed another experiment compared to participants who had not, as fatigue leads to poor judgement and a bad mood (Occupational Safety and Health Service, 1998). The effect of fatigue on judgement in combination with a more negative outlook on the applicant might have negatively influenced the participant's applicant evaluation process. To be certain that these factors do not influence results, future research should only have one experimental session.

Taken together, this study demonstrates that there is a need for an intervention in the applicant evaluation process, as implicit gender bias holds back the process of achieving gender equality in (academic) top positions. Gender equality is necessary, as it benefits society as a whole. The current study helps to develop such an intervention in the form of feedback. It also contributes valuable insight in other-stereotype activation and other-stereotype reactance in selection procedures, which is a research area that so far was underexposed by previous research. Also, based on this study, changing wording to traits or

behaviours in terms of either masculine wording or balanced wording does not appear to have a similarly large impact on perceived overall applicant suitability for the job compared to feedback. In conclusion, the current research has provided a valuable step in the ladder society needs to climb to achieve gender equality, by supporting women in climbing the ladder to (academic) top positions.

References

Born, M. P., & Taris, T. W. (2010). The impact of the wording of employment advertisements on student's inclination to apply for a job. *Journal of Social Psychology, 150*, 485-502. doi:10.1080/00224540903365422

Dillard, J. P., & Shen, L. (2005). On the Nature of Reactance and its Role in Persuasive Health Communication. *Communication Monographs*, *10,* 144-168.

Dovidio, J. F., & Gaertner, S. L. (1996). Affirmative action, unintentional racial biases and intergroup relations. *Journal of Social Issues, 52*, 51-75.

Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology, 5*, 1-35.

Eliezer, D., Major, W., & Mendes, W. B. (2010). The cost of caring: Gender identification increases threat following exposure to sexism. *Journal of Experimental Social Psychology, 46*, 159-165. doi:10.1016/j.jesp.2009.09.015

Ellemers, N., Van Dyck, C., Hinkle, S., & Jacobs, A. (2000). Intergroup differentiation in social context: Identity needs versus audience constraints. *Social Psychology Quarterly, 63*, 60-74.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680-698.

Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertising exists and sustains gender inequality. *Journal of Personality and Social Psychology, 101*, 109-128. doi:10.1037/a0022530

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Socials Psychology, 74,* 1464-1480.

Grove, J. (2013, May 2). Global Gender Index, 2013. *Times Higher Education.* Retrieved March 10, 2017, from https://www.timeshighereducation.com/features/global-gender-index-2013/2003517.article

Gupta, V. K., Turban, D. B., & Bhawe, M. N. (2008). The Effect of Gender Stereotype Activation on Entrepreneurial Intentions. *Journal of Applied Psychology, 93,* 1053-1061. doi:10.1037/0021-9010.93.5.1053

Heilman, M. E. (2001). Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues, 57,* 657-674.

Kaatz, A., Magua, W., Zimmerman D. R., & Carnes, M. (2015). A quantitative linguistic analysis of National Institutes of Health R01 application critiques from investigators at one institution. *Academic Medicine, 90,* 69-75. doi:10.1097/ACM.0000000000000442

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the Sexes: Gender Stereotype Confirmation and Reactance in Negotiations. *Journal of Personality and Social Psychology, 80,* 942-958.

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology, 107,* 631-644. doi:10.1037/edu0000005

Morley, L. (2014). Lost leaders: Women in the global academy. *Higher Education Research & Development, 33,* 114-128. doi:10.1080/07294360.2013.864611

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *PNAS, 109,* 16474-16479. doi:10.1073/pnas.1211286109

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *PNAS, 106,* 10593-10597. doi:10.1073/pnas.0809921106

Occupational Safety and Health Service. (1998). *Stress and Fatigue: Their impact on health and safety in the Workplace*. Retrieved from http://www.worksafe.govt.nz/worksafe/information-guidance/all-guidance-items/stress-and-fatigue-their-impact-on-health-and-safety-in-the-workplace/stress.pdf

Rudman, L. A., & Kilianski, S. A. (2000). Implicit and Explicit Attitudes Toward Female Authority. *Personality and Social Psychology Bulletin, 26,* 1315-1328.

Sheridan, J. T., Fine, E., Pribbenow, C. M., Handelsman, J., & Carnes, M. (2010). Searching for excellence & diversity: Increasing the hiring of women faculty at one academic medical center. *Academic Medicine: Journal of the Association of American Medical Colleges, 85*, 999-1007. doi:10.1097/ACM.0b013e3181dbf75a

Sidanius, J., Pratto, F., & Bobo, L. (1996). Racism, conservatism, affirmative action, and intellectual sophistication: A matter of principled conservatism or group dominance? *Journal of Personality and Social Psychology, 70,* 476-490.

Soper, D. (n.d.). Sobel Test Calculator for the Significance of Mediation. Retrieved on January 22, 2017, from http://www.danielsoper.com/statcalc/calculator.aspx?id=31

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797-811.

Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study. *Sex Roles, 41,* 509-528.

Stroebe, K., Postmes, T., Täuber, S., Stegeman, A., & John, M. S. (2015). Belief in a Just What? Demystifying Just World Beliefs by Distinguishing Sources of Justice. *PLoS ONE, 10,* 1-30. doi:10.1371/journal.pone.0120145

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995) Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68,* 199-214.

Taris, T. W., & Bok, I. A. (1998). On gender specificity of personality characteristics in personnel advertisements: A study among future applicants. *Journal of Psychology, 132,* 593-610.

UN Women. (2016a). *HeForShe IMPACT 10x10x10 Corporate Parity Report*. Retrieved from http://www.heforshe.org/~/media/heforshe/files/davos/heforshe_impact10x10x10_paritytyreport_corporate_final_2016.pdf

UN Women. (2016b). *HeForShe IMPACT 10x10x10 University Parity Report*. Retrieved from http://www.heforshe.org/-/media/heforshe/files/event/2016/heforshe_impact10x10x10_parityreport_university_2016.pdf

Van den Brink, M. (2010). *Behind the scenes of science: Gender practices in the recruitment and selection of professors in the Netherlands.* Amsterdam: Pallas Publications – Amsterdam University Press.

Van der Lee, R., & Ellemers, N. (2015). Gender contributes to personal research funding success in The Netherlands. *PNAS, 112,* 12349-12353. doi:10.1073/pnas.1510159112

Vernos, I. (2013). Research management: Quotas are questionable. *Nature, 495,* 39. doi:10.1038/495039a

Wheeler, S. C., & Petty, R. E. (2001). The Effects of Stereotype Activation on Behavior: A Review of Possible Mechanisms. *Psychological Bulletin, 127,* 797-826.

APPENDIX A

Personalia

Sollicitant:	#109
Naam:	Laura
Sekse:	Vrouw
Leeftijd:	22

Opleiding

2013 – Heden	Bachelor Rechtsgeleerdheid Universiteit Leiden
2007 – 2013	VWO Rijnlands Lyceum Sassenheim

Werkervaring

2014 – 2015	Lid studiereiscommissie Studievereniging
2013 – Heden	Bijlesddocent Engels Rijnlands Lyceum Sassenheim
2013 – 2014 Stagiair	Advocatenkantoor

Deel van motivatiebrief Laura:

"Ik ben een gemotiveerde student en zou mij volledig inzetten voor deze functie in de adviesraad. Door mijn commissiewerk bij de studievereniging heb ik ervaring met samenwerken met commissieleden. Een mooie volgende stap voor mij zou deze functie in de adviesraad zijn."

Deel van aanbevelingsbrief professor t[redacted] (mentor):

"Ondanks dat Laura niet altijd de hoogste cijfers heeft gehaald, heeft ze indruk op me gemaakt door een enorm gemotiveerde student te zijn. Ze doet erg haar best en ze is ambitieus. Daarnaast zie ik een stijgende lijn in haar cijfers. Ik weet zeker dat ze geschikt is voor deze functie."

APPENDIX B

Multiple one-way ANOVAs to examine the effect of session on the dependent measures showed that session influenced applicant rating, $F(1, 283) = 11.28$, $p = .001$, $\eta_p^2 = .04$. Participants in session one rated the applicant higher on applicant rating scores ($M = 5.38$, $SD = 0.82$) than did participants in session two ($M = 5.03$, $SD = 0.90$). Session also influenced perceived competence, $F(1, 283) = 17.57$, $p < .001$, $\eta_p^2 = .06$. Participants in session rated the applicant higher on perceived competence ($M = 5.08$, $SD = 0.86$) than did participants in session two ($M = 4.61$, $SD = 0.99$). There was an effect of session on willingness to employ, $F(1, 283) = 12.49$, $p < .001$, $\eta_p^2 = .04$. Participants in session one rated the applicant greater on willingness to employ ($M = 4.97$, $SD = 1.14$) than did participants in session two ($M = 4.47$, $SD = 1.23$). Also, session influenced applicant likeability, $F(1, 283) = 8.61$, $p = .004$, $\eta_p^2 = .03$. Participants in session rated the applicant higher on applicant likeability ($M = 5.23$, $SD = 0.88$) than did participants in session two ($M = 4.93$, $SD = 0.83$). Lastly, session influenced perceived fit, $F(1, 283) = 7.07$ $p = .01$, $\eta_p^2 = .03$. Participants in session one rated the applicant higher on perceived fit ($M = 4.87$, $SD = 1.08$) than did participants in session two ($M = 4.52$, $SD = 1.12$). Because of the effect of session on the affected measures, session was added as a covariate when analysing these measures.

APPENDIX C

The items of the belief in a just world scale were: “To what extent do you believe the world currently is a righteous place?”, “To what extent do you think the world should be a righteous place?”, To what extent do you believe affirmative action is necessary?”, “To what extent do you believe people get what they deserve?” and “To what extent do you believe that forms of discrimination are inevitable?”. *Hypothesis 2a* stated that a significant positive relationship was expected between implicit gender bias and modern sexism and belief in a just world.

A marginally significant positive correlation was found between ‘belief in a just world: Currently righteous place’ and implicit bias, $r = .11$, $p = .08$: The stronger participants’ implicit gender bias, the more the participant believed the world currently is a rightful place. A marginally significant positive correlation was found between ‘belief in a just world: Inevitability of discrimination’ and implicit bias, $r = .10$, $p = .09$: The stronger participants’ implicit gender bias, the more they believed that discrimination is inevitable. No significant correlation was found between implicit bias and ‘belief in a just world: Should be righteous place’, $r = -.05$, $p = .43$, implicit bias and ‘belief in a just world: Affirmative action’, $r = .06$, $p = -.35$ or implicit bias and ‘belief in a just world: Deservedness’, $r = -.02$, $p = .74$. Thus, *Hypothesis 2a* was partially confirmed. However, the current findings show that implicit gender bias has consequences for the way the world is perceived.

The correlation between implicit bias and the belief in a just world items were not expected to be influenced by the manipulations. To confirm this, several 2x2x2 ANOVAs were done with feedback, gendered wording and type of wording as the independent variables, and the belief in a just world items as the dependent variables. Results on ‘belief in a just world: Inevitability of discrimination’ revealed a marginally significant three-way interaction effect on perceived inevitability of discrimination, $F(1, 272) = 3.62$, $p = .06$, $\eta_p^2 = .01$. Bonferroni post hoc tests revealed that participants reading trait descriptions thought

the inevitability of discrimination was significantly higher than did participants in the behavioural description-conditions, $F(1, 272) = 4.43$, $p = .04$, $\eta_p^2 = .02$). However, this was only the case when no feedback was given and balanced wording was used. Based on these results, the previously found correlation between this belief in a just world item and implicit bias should be interpreted with caution, as the relationship between this item and implicit bias might have been different if it had not been influenced by the independent variables. Results on 'belief in a just world: Currently righteous place' showed no significant three-way interaction effect, $F(1, 272) = .12$, $p = .74$, $\eta_p^2 < .001$, just like 'belief in a just world: Should be righteous place', $F(1, 272) = .06$, $p = .81$, $\eta_p^2 < .001$, 'belief in a just world: Affirmative action', $F(1, 272) = .01$, $p = .94$, $\eta_p^2 < .001$, and 'belief in a just world: Deservedness' $F(1, 272) = .10$, $p = .75$, $\eta_p^2 < .001$. Thus, these items were not influenced by the manipulations.

APPENDIX D

Table 2.

Answers, translations and gendered coding for word completion task item 1

Completed word	Translation	Gendered coding
Inferieur	Inferior	F
Superieur	Superior	M
Exterieur	Exterior	N
Interieur	Interior	N
-	Non-existent words	NON

Table 3.

Answers, translations and gendered coding for word completion task item 2

Completed word	Translation	Gendered coding
Zus	Sister	F
Bus	Bus / tin	N
Dus	Thus	N
Jus	Gravy / orange juice	N
Kus	Kiss	N
Lus	Loop	N
Mus	Sparrow	N
Pus	Pus	N
Rus	Russian person	N

Table 4.

Answers, translations and gendered coding for word completion task item 3

Completed word	Translation	Gendered coding
Pan	Pan	F
Jan	Male name	M
Man	Man	M
Aan	To	N
Ban	Ban	N
Dan	Then	N
Fan	Fan	N
Kan	Can / jug	N
Van	Of	N
-	Non-existent words	NON

Table 5.

Answers, translations and gendered coding for word completion task item 4

Completed word	Translation	Gendered coding
Rok	Skirt	F
Bok	Male goat	M
Dok	Dock	N
Gok	Guess	N
Hok	Hutch	N
Kok	Cook	N
Lok	Earlock	N
Mok	Mug	N

Mol	Mole	N
Nok	Ridge	N
Ook	Also	N
Sok	Sock	N
Tok	Tock	N
Wok	Wok	N

Table 6.

Answers, translations and gendered coding for word completion task item 5

Completed word	Translation	Gendered coding
Danspas	Dance step	F
Handtas	Handbag	F
Handwas	Hand wash	F
Bankkas	Bank till	N
Bankpas	Bank card	N
Gangpas	Hall pass	N
Landras	Land race	N
-	Non-existent words	NON