

**Promoting Gender Equality at Different Stages of the Employee Life Cycle: New and
Cumulated Findings**

Inaugural dissertation for the degree Doctor of Social Sciences at the University of Mannheim

By

Mona Salwender

Submitted to the School of Social Sciences at the University of Mannheim

2023

Dean of the School of Social Sciences:

Prof. Dr. Michael Diehl

Thesis Supervisors:

Prof. Dr. Dagmar Stahlberg

Prof. Dr. Herbert Bless

Thesis Reviewers:

Prof. Dr. Jutta Mata

Prof. Dr. Sabine Sonnentag

Date of Defense:

08/24/2023

Preface

The three manuscripts that form this dissertation are listed below. All manuscripts are published or are ready to be submitted.¹

Salwender, M., & Stahlberg, D. (2023). *Do women only apply when they are 100% qualified, whereas men already apply when they are 60% qualified?* [Manuscript in preparation].

Department of Social Psychology, University of Mannheim.

Salwender, M., Schoel, C., Bless, H., & Stahlberg, D. (2023). The politics hurdle: Joint effect of organizational culture and gender on lack of fit experiences. *Social Psychological and Personality Science*, 14(1), 84–92. <https://doi.org/10.1177/19485506221075898>

Salwender, M., Maschmann, I. T., Heinzl, V. K., Sczesny, S., & Stahlberg, D. (2023). *Masculine generics versus gender-inclusive language: A meta-analysis of the effects of language form on the mental representations of gender* [Manuscript in preparation]. Department of Social Psychology, University of Mannheim.

¹ The data from Study 1a in Manuscript 2 (“The politics hurdle...”) stem from Salwender, M. (2018). *Power through political advancement strategies: Self-concept conflict and fear of backlash mediate gender differences in intention to seek power positions* [Unpublished Master's thesis]. University of Mannheim, Germany.

Table of Contents

Overview	1
Promoting Gender Equality at Different Stages of the Employee Life Cycle	4
Gender as One Core Dimension of Diversity	5
Gender Diversity Along the Employee Life Cycle	6
Hot Topic 1: An Empirical Foundation Is Indispensable for Diversity Management	8
Manuscript 1: Do Women Only Apply When They Are 100% Qualified, Whereas Men Already Apply When They Are 60% Qualified?	10
Hot Topic 2: Stop Fixing Women	12
Manuscript 2: The Politics Hurdle: Joint Effect of Organizational Culture and Gender on Lack of Fit Experiences	13
Hot Topic 3: Gender-Inclusive Language	15
Manuscript 3: Masculine Generics Versus Gender-Inclusive Language: A Meta-Analysis of the Effects of Language Form on the Mental Representations of Gender	16
Summary of the Findings	17
Discussion	18
Theoretical Contributions	18
<i>Empirical Foundation for Diversity Management</i>	21
<i>Fixing Organizational Cultures</i>	21
<i>Mental Representations of Gender in Language</i>	22
Empirical Contributions	23
Open Questions and Limitations	23
Practical Implications	25
<i>Gender Diversity Management Along the Employee Life Cycle</i>	25
<i>Three Manuscripts – Three Main Messages</i>	25
<i>Three Hot Topics</i>	26
Conclusion	26
References	28
Statement of Originality	44
Statements of the Co-Authors	45
Appendix: Manuscripts	47

Overview

There are three main reasons for organizations to promote diversity (Charta der Vielfalt, 2022) and many organizations already actively do so (Dobbin et al., 2011). One good reason for promoting diversity is simply that it is a matter of fairness. Another good reason for promoting diversity is that a diverse workforce can generate competitive and financial advantages (Gilbert et al., 1999; Herring, 2009). A third and related reason is that promoting diversity can be an answer to pressuring needs of organizations such as the demographic change and shortage of qualified staff (Charta der Vielfalt, 2022). Such pressuring needs require organizations to be attractive for a diverse workforce at all stages of the employee life cycle. This dissertation offers novel and cumulated findings relevant for organizational diversity management. The focus of this dissertation lies on gender as one of the core dimensions of diversity. Each of the three manuscripts included in this dissertation addresses a hot topic around the promotion of gender equality in organizations at different stages of the employee life cycle.

Manuscript 1 is concerned with the hiring stage of the employee life cycle, specifically with the question whether women only apply for a job when they are 100% qualified, whereas men already apply when they are qualified to a lower extent. This is a wide-spread claim (Mohr, 2014). However, empirical evidence has been lacking so far. By testing this claim, Manuscript 1 responds to the hotly debated lack of empirical investigations in the diversity management practice. The manuscript presents novel data in seven studies testing whether women and men differ in their application intention depending on qualification fit. The results of the first six studies showed no gender differences in application intention at different levels of qualification fit. At the same time, women robustly indicated that they want to be more prepared for a job compared to men. This speaks for higher psychological hurdles present in women than men, which do not seem to translate into differential application intentions. In

Study 7 a competitive test of 29 potential mediators showed that women wanted to be more prepared for a job due to their higher fear of not being able to meet others' expectations.

Manuscript 2 focuses on the development stage of the employee life cycle. Specifically, company cultures and their role in motivating or hindering women and men to strive for leadership were investigated. By focusing on company cultures, this research followed an emerging call to stop fixing women in the attempt to reduce the gender leadership gap (e.g., by offering trainings for female leaders) and rather address systems (England et al., 2020; Fox, 2017). Novel data in three studies showed that cultures focusing on showing competencies foster less self-concept conflict, less fear of backlash, less concerns about one's skills, and a higher intention to seek power positions compared to cultures focusing on playing politics. This pattern was more pronounced for women than men.

Finally, Manuscript 3 focuses on language and therefore has implications for corporate communication across all stages of the employee life cycle. Following an ongoing, partly heated debate in society, media, politics, as well as organizations (e.g., Kurfer, 2022), this manuscript scrutinizes the mental representation of gender in masculine generics. In a meta-analysis, evidence is accumulated showing that women's mental representation is higher when gender-inclusive language (e.g., chairperson) is used compared to masculine generics (e.g., chairman). This applies, for example, to the language used in job advertisements in the hiring stage but also in a leadership context and therefore has direct implications for researching and promoting gender equality in organizations.

The present dissertation is organized as follows: The next chapter outlines theoretical and empirical background on the promotion of gender equality at different stages of the employee life cycle. Afterwards, three current hot topics in the field of diversity management are highlighted. The three manuscripts that form this dissertation are shortly summarized with reference to the respective corresponding hot topic. Details on theory, methods, results, and

discussions pertaining to the specific manuscripts can be found in each manuscript in the appendix. Finally, a general discussion follows highlighting theoretical and practical implications of this dissertation as well as limitations and open questions.

Promoting Gender Equality at Different Stages of the Employee Life Cycle

In a recent survey on diversity at work about half of the surveyed employees reported that a change towards more diversity has taken place over the past years in their organization (Statistisches Bundesamt, 2023d). In another survey diversity at work was perceived as success factor for the creation of a positive corporate image, for increasing employee motivation, and to deal with the shortage of qualified staff (Statistisches Bundesamt, 2023a). But what exactly is diversity? Diversity in an organizational setting is defined as “similarities and differences in the workforce based on individual personality traits, as well as lifestyles and life plans” (Charta der Vielfalt, 2022, p. 6). Those similarities and differences can stem from visible and invisible characteristics (Charta der Vielfalt, 2022; Yadav & Lenka, 2020). Dimensions of diversity vary between different streams of the literature, have been extended and changed over time. For example, one often-made distinction is between relations-oriented diversity attributes, such as gender and age, and task-related diversity attributes, such as work experience and organizational tenure (Joshi & Roh, 2009; Yadav & Lenka, 2020). The seven dimensions that are regarded as core dimensions of diversity are gender and gender identity, age, ethnic background and nationality, physical and mental abilities, religion and worldview, sexual orientation, and social background (Voß & Würtemberger, 2023).

With the goal to eliminate discrimination based on diversity dimensions, organizations started to implement diversity management in the 1990s (Dobbin et al., 2011; Emmerich & Krell, 2002; Gilbert et al., 1999). During this time, civil rights and women’s movements facilitated norms and legislation for equal opportunities (Dobbin et al., 2011). Besides moral arguments, organizational diversity management has been argued to lead to competitive advantages (Cox & Blake, 1991). Indeed, diversity has been found, for example, to be associated with increased sales revenue, an increase in number of customers, and greater relative profits (Dixon-Fyle et al., 2020; Gilbert et al., 1999; Herring, 2009; Singal, 2014).

Yet, besides positive outcomes a vast amount of research has shown mixed and even negative outcomes of diversity. In their meta-analysis, Joshi and Roh (2009) found a weak negative relationship between relations-oriented diversity (e.g., age) and team performance and a weak positive relationship of task-related diversity (e.g., work experience) and team performance. The simultaneous existence of positive and negative outcomes of diversity has been coined as “double-edged sword of diversity” (Carter & Phillips, 2017) and highlights the importance of extensive research and evaluation in the field of diversity.

Gender as One Core Dimension of Diversity

In this dissertation, I will focus on the core dimension of gender as far-reaching work-related gender gaps exist (World Economic Forum, 2022). Gender is a widely researched dimension with regards to diversity at work. Both biological factors (oftentimes denoted with the term sex; examples for biological factors are X and Y chromosomes, hormones) and sociocultural factors (e.g., norms and expectations) determine the gender of a person (Deaux, 1985; Hyde et al., 2019). Overall, gender is an omnipresent, readily available diversity dimension (Bennett et al., 2000; Ito & Urland, 2003; Macrae & Bodenhausen, 2001). Gender is acquired as concept very early as for example three- to four-month-old infants already distinguish between genders (Quinn et al., 2002; for an overview see Martin & Ruble, 2010). Beyond childhood gender possesses a prominent role throughout life until death (see gender in obituaries, Zehnter et al., 2018). Gender identity is defined as the identification of a person as female, male, or another gender (de Vries et al., 2014). Theories explaining gender differences and similarities exist from evolutionary perspectives (Buss, 2005; Buss et al., 2020; Matlin, 2004) as well as social, cognitive, and developmental psychological perspectives (e.g., Eagly et al., 2000; Eagly & Wood, 2012).

A look at broad, work-related indicators shows a clear pattern of gender differences. Gender gaps exist for example with respect to fewer women than men in leadership positions

(Kulich & Bosak, 2019; World Economic Forum, 2022), women and men applying for different jobs or at a different rate (Fluchtmann et al., 2022; Ley & Hamilton, 2008), and gender gaps in salary (AAUW, 2022). According to the World Economic Forum it will take 151 years to close the gender gap in economic participation and opportunity (World Economic Forum, 2022). Many organizations already work towards closing gender gaps. Considering the employee life cycle is a helpful framework to identify the needs of employees as well as organizations regarding gender diversity initiatives (Gladka et al., 2021).

Gender Diversity Along the Employee Life Cycle

The employee life cycle is defined as “the stages employees go through and the role HR or similar support takes on during those stages” (Saltmarsh, 2017, p. 8). In other words, the employee life cycle describes the different periods of employees’ work life in an organization, for example from being hired, to being promoted, to retiring. The employee life cycle can also be viewed from employers’ perspective. From this perspective, the employee life cycle defines tasks for the Human Resources (HR) department and related departments of an organization such as onboarding of a new employee, facilitating training and learning, to succession planning when the employee retires (Gladka et al., 2021).

Importantly, the employee life cycle, and accordingly the needs of employees and activities of employers, vary depending on diversity dimensions of the employees (Gladka et al., 2021, Heitner et al., 2013). To illustrate, the employee life cycle differs depending on age as older employees typically stay longer within the same organization than younger employees (Becton et al., 2014). With regards to gender one frequently researched topic in the employee life cycle is parenthood. While 93% of fathers were employed in 2019, only 75% of mothers were (Statistisches Bundesamt, 2023b). Parenthood leads to more fragmented employment histories for mothers than fathers (Fenton & Dermott, 2006), therefore to different employee life cycles.

Beyond parenthood, much research on gender diversity along the employee life cycle has examined the hiring stage. Past research has for example found preferential hiring of male over female applicants (Bosak & Sczesny, 2011; Hardy et al., 2022; Koch et al., 2015). A simulation study shows that already a small gender bias in hiring decisions can result in substantial hiring discrimination for female applicants. In addition, such a small gender bias can result in financial losses for organizations as biased hiring practices prevent organizations from hiring the best candidates (Hardy et al., 2022).

Further, plenty of research on gender diversity along the employee life cycle has investigated gender and leadership. Summarizing past research, findings indicate, for example, that men emerge more than women as leaders from initially leaderless groups (Eagly & Karau, 1991). Further, female and male leaders have been found to lead equally effective (Eagly et al., 1995; Paustian-Underdahl et al., 2014). Female leaders even tend to show more of the very effective transformational leadership style than male leaders (Eagly et al., 2003). Beyond that, studies have examined how women and men differ in their aspiration towards leadership positions and found lower leadership aspirations for women than men (Gino et al., 2015; Schuh et al., 2014).

These examples of hiring and leadership show two stages of the employee life cycle where the needs of employees (e.g., need for leadership motivation) differ depending on their gender. At the same time, examining gender equality along the employee life cycle reveals demands for organizational diversity programs and interventions, for example to prevent financial losses due to biased hiring.

In the past years in research as well as practice several hot topics for organizational diversity management have emerged. Three hot topics that are relevant to promoting gender equality at the hiring and leadership stages of the employee life cycle are outlined in the following with reference to the corresponding manuscript of this dissertation.

Hot Topic 1: An Empirical Foundation Is Indispensable for Diversity Management

The first hot topic is the indispensability of an empirical foundation for diversity management. Several examples showcase how important an empirical foundation for diversity management is.

One example are the popular diversity trainings. Such trainings often focus on educating about diversity and on building skills for dealing with diversity (Ehrke et al., 2020; Emmerich & Krell, 2002). Even though diversity trainings are popular, they are accompanied with criticism: Diversity trainings are rarely evaluated (Ehrke et al., 2020; Emmerich & Krell, 2002; McCauley et al., 2000; Schmader et al., 2022), oftentimes only asking for participants' satisfaction (Emmerich & Krell, 2002), but seldomly measuring their effectiveness regarding education and skills. The evaluations that did go beyond participants' satisfaction showed that diversity trainings lead to cognitive learning (e.g., increased knowledge about diversity), and to a smaller extent to behavioral (e.g., change in behavior) and affective learning (e.g., attitude change; Bezrukova et al., 2016; Kalinoski et al., 2013). This lack of empirical evaluation is problematic, because—despite their good intentions and positive outcomes, especially for cognitive learning—recent discussions arose around unintended, negative outcomes of diversity trainings (Ehrke et al., 2020; Kalev & Dobbin, 2020). Indeed, evidence has accumulated showing that diversity trainings can backfire by activating biases (Bigler, 1999; Rudman et al., 2001). This can be explained on the one hand by a focus on differences which can activate stereotypes (Caleo & Heilman, 2019). On the other hand, so-called rebound effects, where the suppression of biases ironically leads to higher accessibility of the very same, can occur (Rudman et al., 2001; Wegner et al., 1993).

Another example showcasing the importance of an empirical foundation for diversity management are recently discovered unintended signals of managing diversity per se. Diversity initiatives can for example unintendedly signal that underrepresented groups are less

competent and need training (Dover et al., 2020). Further, rationalizing diversity initiatives with moral arguments can signal that privileged groups are expected to create inclusive environments (Chaney, 2022), leading to less positive evaluations of diversity initiatives by privileged groups (Starck et al., 2021). Rationalizing diversity initiatives with instrumental arguments (“business case for diversity”) has been found to lower underrepresented groups’ sense of belonging (Georgeac & Rattan, 2023). In a similar manner, recent cross-national research has revealed unexpected findings. Instead of closing gender gaps, higher national gender equality was associated with more pronounced gender differences in psychological traits (“gender equality paradox”; see Berkessel, *Salwender*, et al., 2023).

Such discovered unintended outcomes or correlates, due to training, communication, or other factors, have raised the hot topic of strengthening the empirical foundation of diversity management. Regarding gender, research on the question why women are underrepresented in leadership positions highlights exemplarily how important it is to test hypotheses empirically. One could think of very different explanations why women are underrepresented in leadership. For example, women may lack the skills, women may lack the motivation, or external barriers may hinder women from obtaining leadership positions (Stahlberg, 2021). As outlined in the chapter about gender diversity along the employee life cycle, women lacking skills can be refuted as an explanation for the gender gap in leadership based on past research (Eagly et al., 1995; Paustian-Underdahl et al., 2014). In contrast, evidence for lacking motivation and external barriers has accumulated (e.g., Gino et al., 2015; Morgenroth et al., 2020). Importantly, without this empirical foundation, diversity management initiatives might unfoundedly advertise (and fund) skill-building trainings for female leaders as pathway towards gender equality in leadership. In contrast, basing diversity management on an empirical foundation ensures a sustainable investment of time and monetary organizational resources as well as reaching the intended diversity goals.

Taken together, these examples clearly show the importance of a strong empirical foundation for diversity management. Manuscript 1 of this dissertation follows this call. A claim often made but not empirically tested so far was investigated: Do women only apply when they are 100% qualified, whereas men already apply when they are qualified to a lower extent?

Manuscript 1: Do Women Only Apply When They Are 100% Qualified, Whereas Men Already Apply When They Are 60% Qualified?

The first manuscript of this dissertation starts with the employee life cycle at the hiring stage. The main purpose of Manuscript 1 was to test a widespread, but under-researched claim in seven studies, namely whether women only apply when they are 100% qualified, whereas men already apply when they are 60% qualified.

In Studies 1-3 $N = 645$ participants read a CV and a job advertisement imagining the presented CV as their own and indicated their intention to apply for the job. All three studies contained a CV and a job advertisement that matched in 60% of the qualification criteria. In Study 1 this 60% qualification fit was compared to a 100% qualification fit in a between-subjects design. In Study 2 this 60% qualification fit was compared to a 0%, 20%, 40%, 80%, and 100% qualification fit in a within-subjects design. In Study 3, only 60% qualification fit was examined. In all three studies, neither frequentist nor Bayesian analyses found evidence for a gender difference in application intention, for no level of qualification fit.

Based on these first three studies a natural conclusion could be that no gender differences in application intention depending on qualification fit exist. However, the design used in Studies 1-3 may have come with some problems, as imagining a presented CV as one's own might have erased self-relevance of the situation and identification with the CV might have been difficult. In Studies 4-6 we therefore used a different study design reducing the hypotheticality of the situation. Specifically, in Studies 4-6 participants first read a job

advertisement and rated their intention to apply. Afterwards, participants indicated for each qualification criterion whether they fulfilled it or not. An integrative data analysis of $N = 807$ participants showed that, in line with previous research (Van Hooft et al., 2006), higher qualification fit predicted higher application intentions. However, the evidence for the claimed gender difference in application intention depending on qualification fit was inconclusive: On the one hand, the Participant Gender \times Qualification Fit interaction on application intention was significant. On the other hand, this interaction effect was mainly driven by women indicating a higher application intention than men at high levels of qualification fit and the Bayesian analysis spoke against the presence of the predicted interaction. In contrast to this inconclusive result for application intention, we robustly found higher psychological hurdles in application situations present for women compared to men. We examined these higher psychological hurdles in more detail in Study 7 using a machine-learning algorithm to competitively test various explanatory variables and found that women wanted to be more prepared than men regarding their qualification fit when applying for a job mainly due to their higher fear of not being able to meet others' expectations. Taken together, this allows the conclusion that women have to go the extra mile by overcoming higher psychological hurdles but seem willing and capable to do so, as higher psychological hurdles did not translate into differential application intentions.

In sum, Manuscript 1 shows that the matter of gender differences in application settings seems more complex than previously thought (i.e., higher psychological hurdles and at the same time no gender difference in application intention) and more empirical research is needed to reach a final conclusion. With reference to the indispensability of an empirical foundation in the field of diversity management, Manuscript 1 shows that testing claims empirically is very important to ensure that there is a need for action and the right actions are taken before investing time and money for interventions. The implications for future research as well as organizations' pursuit to gender equality are discussed in the discussion section.

Hot Topic 2: Stop Fixing Women

Besides the need for an empirical basis for diversity management another current hot topic is the call to stop fixing women. Past research has predominantly addressed the disadvantaged groups, whereas the role of advantaged groups was neglected (Croft et al., 2015; Ehrke et al., 2020). Further, popular literature such as the book “Lean in” (Sandberg, 2015) has spread the message that women can overcome gender hurdles by increasing their confidence. This focus in both research and popular literature can lead to attributions that women are responsible for creating and solving gender inequalities (Kim et al., 2018). In her 2017 book “Stop fixing women” Catherine Fox outlines that the prevailing narrative of women need to (be) fix(ed) hinders a shift towards fixing the gender imbalanced system as the underlying cause of gender inequalities. Correspondingly, an analysis of the progress towards gender equality shows that progress has stalled, potentially as a shift from changing women to changing men and systems is needed to enable further progress (England et al., 2020).

The call to stop fixing women is reflected in changes of focus in research. One stream of research has addressed the role of men in the pursuit to gender balance, with research covering, for example, backlash for men (Chaney et al., 2019; Moss-Racusin, 2014), and men as allies (Moser & Branscombe, 2022). Another stream of research has addressed the system side. The system approach seems especially crucial, as more and more research reveal backfires of diversity initiatives focusing on training of individuals (see the first hot topic where evidence was presented revealing backfiring of diversity trainings; Ehrke et al., 2020; Kalev & Dobbin, 2020). Further, implicit biases of individuals have been found to reflect biases in the environment (Vuletic & Payne, 2019). This suggests that changes in environments could translate into changes in implicit biases of individuals and therefore changing systems can be effective also on an individual level. In addition, changes in processes seem very effective. Research focusing, for example, on hiring processes has shown that biases preventing gender

balance can be overcome by changing hiring processes (e.g., using structured instead of unstructured interviews; Koch et al., 2015; Prati et al., 2019; Welpé & Peus, 2014).

Besides processes a huge impact variable in organizational systems is organizational culture. Organizational culture is defined as “the pattern of shared values and beliefs that help individuals understand organizational functioning” (Deshpande & Webster, 1989, p. 4). Past research on organizational cultures has, for example, revealed that workplaces are far from being a meritocracy and rather amplify inequalities (van Dijk et al., 2020). Further, an analysis of organizational cultures has shown that they are gendered – they developed around the interests of men (e.g., concerning management style, working hours, informal socializing activities; Rutherford, 2014). Building on the call to investigate systems, Manuscript 2 in the present dissertation focuses on organizational cultures in the context of leadership.

Manuscript 2: The Politics Hurdle: Joint Effect of Organizational Culture and Gender on Lack of Fit Experiences

The second manuscript of this dissertation moves along the employee life cycle from the hiring stage to the development stage. According to a recent survey, women in leadership and equality in promotions are two of the top five diversity topics where organizations must catch up (Statista, 2023c). The main purpose of Manuscript 2 was to test the effect of organizational culture on women’s and men’s intention to seek power positions. In this regard, Manuscript 2 follows recent discussions on taking systems into account when addressing gender inequalities in organizations (England et al., 2020).

Manuscript 2 consists of three studies investigating women’s and men’s lack of fit experiences in two organizational cultures: playing politics versus showing competencies. In politics cultures, for example, deciding strategically on a purely business basis, and not paying attention to the sensitivities of others is an effective behavior to acquire power positions, whereas in competencies cultures, for example, doing excellent work and being competent at

what you do is an effective behavior to acquire power positions. In a pre-study participants ($N = 100$) rated playing politics behaviors (e.g., deciding strategically) as comparatively more dominant compared to showing competencies behaviors (e.g., doing excellent work). In contrast, showing competencies behaviors were rated as comparatively more competent, warm, and moral compared to playing politics behaviors.

In Studies 1a and 1b we then examined women's and men's lack of fit experiences depending on organizational culture. First, participants should imagine starting their career after university in a consulting company. They either learned that it would take playing politics or showing competencies to move up the ranks. Afterwards, participants indicated their lack of fit experiences when imagining starting their career in this company with the respective organizational culture. Specifically, participants reported their self-concept conflict, fear of backlash, concerns about one's skills, and intention to seek power positions. An integrative data analysis across the samples from Study 1a and 1b ($N = 689$ in sum for both Studies 1a and 1b, which were analyzed together as Study 1b was a direct replication of Study 1a) showed that participants indicated more self-concept conflict, more fear of backlash, more concerns about their skills, and less intention to seek power positions in politics compared to competencies cultures. While no gender difference emerged in competencies cultures, women showed higher self-concept conflict, higher fear of backlash, higher concerns about their skills, and lower intention to seek power positions than men in politics cultures.

In sum, Manuscript 2 highlights the importance of addressing organizational culture as one structural variable in the endeavor to close the gender gap in leadership. Based on the current state of research, competencies cultures motivate, and politics cultures demotivate, especially women, regarding leadership. Again, the implications for future research as well as organizations' pursuit to gender equality are discussed in the General Discussion section.

Hot Topic 3: Gender-Inclusive Language

The third hot topic within the field of diversity management that this dissertation takes up is gender-inclusive language. This topic has been hotly debated in society, media, politics, science, and organizations again and again. Gender-inclusive language means to communicate in such a way that everyone feels they were included (Charta der Vielfalt, 2023) and stands in contrast to masculine generics which are criticized to elicit a male bias (Guentherodt et al., 1980). The criticism against masculine generics (i.e., using grammatically masculine forms such as “he” or “firemen” generically for all genders) contains that—although masculine generics are meant to include other genders—they go along with a lower mental representation of female and other genders. Based on this feminist language critique, alternative language forms (e.g., he/she, they, firefighter) have been proposed to enhance the mental representation of female and other genders.

Corporate language, specifically gender-inclusive language versus masculine generics, has recently received a lot of attention as the example of an AUDI employee in Germany going to court because he did not want to be addressed by gender-inclusive forms shows (Eidlyn, 2022). In the past few years many organizations have changed their communication and now use gender-inclusive language to address their employees and customers (Statistisches Bundesamt, 2021). In Germany, for example, a law for equal treatment (“Allgemeines Gleichbehandlungsgesetz”) dictates equal treatment in application processes for all genders. This implies that job advertisements need to be formulated in gender-inclusive language (IHK Wiesbaden, 2023). Correspondingly, an increased use in gender-inclusive language can prominently be seen in job advertisements (Haas & Vetter, 2021). This has been met with criticism until today. Repeatedly, the quantitative evidence was doubted, citing single findings that show no difference in the mental representation of women in gender-inclusive language versus masculine generics (Kurfer, 2022). From a scientific standpoint such doubts can best be

answered with summarizing all quantitative evidence available. The results of such a summary are presented in Manuscript 3.

Manuscript 3: Masculine Generics Versus Gender-Inclusive Language: A Meta-Analysis of the Effects of Language Form on the Mental Representations of Gender

Regarding the employee life cycle, Manuscript 3 is relevant to both the hiring and development stages (e.g., language used in job advertisements), but also beyond for corporate communication across all stages of the employee life cycle. The main purpose of Manuscript 3 was to cumulate research on the mental representation of women (the mental representation of other genders was not included as research on this is very recent and therefore limited in number) when gender-inclusive language compared to masculine generics is used. Cumulating 357 effect sizes from 95 independent samples from 19,582 individuals in a meta-analysis provides evidence for a higher mental representation of women in reaction to gender-inclusive language compared to masculine generics.

To come to this summary result, we first searched for relevant studies in a massive effort to comprehensively study all research on the topic. Then a screening, eligibility check, and coding process as well as the calculation of effect sizes followed. Multi-level random effects models with robust variance estimation were calculated to account for the multi-level nature and dependence existing in the data.

Different strategies for gender-inclusive language (Lindqvist et al., 2019) were compared to test for moderating effects. Multi-gendering strategies highlighting a non-binary understanding of gender (e.g., Swedish “hen”, German gender-asterisk *) showed the largest difference in mental representation of women between gender-inclusive language and masculine generics, followed by feminization strategies (e.g., pair form such as he/she) and neutralization strategies (e.g., chairperson).

In sum, Manuscript 3 highlights the difference language can make for the mental representation of women bearing direct implications for corporate communication. A statistical summary of past research on gender-inclusive language versus masculine generics showed that women are more mentally represented in gender-inclusive language compared to masculine generics. Again, the implications for future research as well as organizations' pursuit to gender equality are discussed in the next section.

Summary of the Findings

To conclude, in this dissertation three hot topics from the field of diversity management were addressed: (1) the need for an empirical basis in diversity management, (2) the call to stop fixing women, and (3) the hot topic of gender-inclusive language. Taken together the three manuscripts provide new and cumulated findings on the promotion of gender equality at different stages of the employee life cycle. The stages range from hiring to development and the topics addressed range from the question whether women only apply for jobs when they are 100% qualified, over the influence of organizational culture on leadership motivation (especially for women) to the difference language can make in corporate language for the mental representation of women. The methodological spectrum ranges from experiments to surveys, from classic frequentist to cutting-edge Bayesian and machine-learning analyses, and even incorporates a multi-level meta-analysis.

The general discussion following in the next chapter builds on the findings of the three manuscripts presented above. The theoretical, empirical, and practical considerations of the following general discussion contain novel and continuative ideas beyond the specific manuscripts and integrate the three research projects into a larger framework showcasing learnings and next steps for promoting gender equality in organizations with reference to the employee life cycle and the three hot topics outlined above.

Discussion

Theoretical Contributions

The first theoretical contribution this dissertation makes is highlighting psychological hurdles. Such psychological hurdles can be found in Manuscripts 1 and 2: No gender difference in application intention emerged, whereas women robustly wanted to be more prepared than men in application situations due to their stronger fear of not being able to meet expectations. A similar pattern emerged in the leadership area. Gender differences in intention to seek power positions were very small, whereas gender differences in psychological hurdles such as fear of backlash and self-concept conflict were very pronounced and robust. This pattern allows the conclusion that even though gender differences in a result or behavior are small or non-existent, an inequality in psychological hurdles may still be present and should therefore not be overseen.

The presence of psychological hurdles found in Manuscripts 1 and 2 may be part of a larger phenomenon. At the one hand, many gender gaps, for example in leadership and politics, have become smaller over the past years, as (among other contributing factors) the relative number of female leaders and female politicians has increased (World Economic Forum, 2022). At the other hand, persistent and increasing gender gaps in mental health can be found. Women have, for example, a higher lifetime risk of depression and anxiety (Faravelli et al., 2013; Moreno-Agostino et al., 2021; Piccinelly & Wilkinson, 2000). Further, anecdotal evidence is accumulating that especially for working mothers “women can have it all” seems a myth (Hennig, 2019). While on the surface it seems like women can have it all (which refers to having children and working) in most parts of the Western world (where the research presented in this dissertation has been conducted), it seems that this costs their mental health and happiness (Arena et al., 2023). The Covid-19 pandemic that has started in 2020 served as a magnifying glass highlighting higher psychological hurdles in women than men: When child-care broke down, for example due to lockdowns, this higher demand was mainly shouldered

by women. This has led to an intensification of working mothers' "double shifts" (i.e., being full-time employed and doing a major part of household chores and care work; Huang et al., 2021). Recent research already documents severe, long-term effects on their health (Alon et al., 2020; Yaish et al., 2021).

This has important implications. First, the finding of psychological hurdles could hint towards (gender) inequalities becoming more subtle. Researchers and practitioners are therefore well-advised to observe changes in phenomena closely. Further, this implies that researchers should include assessments of psychological hurdles in their research to learn more about gendered psychological hurdles. Organizations should take the presence and impact of psychological hurdles at work into account. Using the research presented in Manuscript 1 exemplarily, organizations could work on removing psychological hurdles, for example, by revising job advertisements and application processes towards transparent communication of expectations regarding qualification fit. Further, organizations could examine their hiring process concerning gender biases with reference to expectations.

The second theoretical contribution this dissertation makes is highlighting the importance to fix systems. This contribution can be found in Manuscripts 2 and 3: Politics versus competencies cultures make a difference for employees', especially women's, intention to seek power positions. The usage of gender-inclusive versus masculine generic language makes a difference for the mental representation of women. This allows the conclusion that systems such as corporate culture and communication contribute to existing inequalities and can be changed to promote gender equality.

Besides culture and communication, processes have been found to be important, as for example using structured interviews has been identified as an effective measure to debias selection processes (Winkler, 2015). Organizations therefore should be aware of how these aspects shape employee journeys. Focusing on addressing systems instead of individuals is

important as it questions past narratives that women need to be fixed. Further, it circumvents unintended consequences that past research has uncovered (e.g., negative effects of anti-bias trainings). Moving beyond this important implication of this dissertation, the next step could be to question whether anything at all needs to be fixed on the side of individuals towards embracing differences. This shift seems to be the next level for organizations beyond diversity management towards an inclusive organization (O'Donovan, 2017). In an inclusive workplace culture employees feel that they are a fully accepted part of the organizational system (Mor Barak et al., 2016). This pertains to employees with diverse backgrounds and mindsets and stretches from feelings that one's voice is heard to feelings that everyone is encouraged to meaningfully contribute to the organizations' goals (Pless & Maak, 2004). A meta-analysis on inclusive workplace cultures has shown that the perception that an organization encourages an inclusion climate is related to more beneficial (e.g., organizational commitment) and less detrimental (e.g., intention to leave) work outcomes (Mor Barak et al., 2016). An inclusive culture can be created via fostering strong identification with one's team, supporting long-lasting, stable relationships, fostering respectful behavior, and having a management who is a role model for inclusive behavior (O'Donovan, 2017).

For future research the fixing systems approach implies that boundary conditions and larger contexts should be considered. This can be realized in very different ways as illustrated by our own research (e.g., examining moderators in a meta-analysis or addressing organizational cultures as focal research questions). Beyond the research discussed in this dissertation, I currently work on a research project in which we investigate cross-national differences in gender differences to learn more about the role of cross-national effects with regards to gender equality (Berkessel, *Salwender*, et al., 2023).

In the following short sections, I will highlight theoretical contributions of the individual manuscripts with reference to the three hot topics covered in this dissertation.

Empirical Foundation for Diversity Management

Manuscript 1 tested the claim whether women only apply for a job when they are 100% qualified and men already apply when they are 60% qualified. The results did not provide evidence for gender differences in application intention and at the same time highlighted gender differences in psychological hurdles. With reference to the discussion and call for an empirical foundation for diversity management, this result highlights the importance to test such claims empirically. Without this critical test, the related narrative of women lacking self-confidence (Sandberg, 2015) might spread further feeding a fixing women approach. This is in several ways consequential. First, repeating this unvalidated claim could lead to stereotype threat, ironically leading to worse performance (Spencer et al., 2016). Second, the conclusion is very different for the result found in our studies, as this speaks for fear of not being able to meet expectations in women compared to men. Further investigation is required whether this fear is valid and how to tackle it.

Fixing Organizational Cultures

Manuscript 2 found that women and men are more inclined to seek power positions in competencies cultures compared to politics cultures, and this difference was more pronounced for women. This manuscript highlights the role organizational culture plays with regards to leveling the playing field for female and male leaders. An implication for future research is to address organizational cultures to promote organizational gender equality. This is in line with prior research showcasing the importance of organizational culture as in one exemplary study an increase in female representation in a selection committee only caused an increase in female hires when cultural norms supporting diversity and inclusion were salient (Baron et al., 2021). Yet, with regards to researching culture, investigating moderators seems important as every organization is different, and a one-size-fits-all approach will therefore hardly work (O'Donovan, 2017).

Mental Representations of Gender in Language

The major contribution of Manuscript 3 is that language does make a difference for the mental representation of women. This is also relevant for Manuscripts 1 and 2 as language has been found to make a difference in application situations, for example in job advertisements and also specifically in the leadership context (Hentschel et al., 2018; Horvath & Sczesny, 2016). This intertwining of topics in the current dissertation highlights the usefulness of the employee life cycle as framework to keep the big picture in mind, identify overarching topics and interdependencies (Voß & Würtemberger, 2023).

With regards to mental representations of gender in language it is striking how much reactance towards diversity initiatives at work can occur at several stages of the employee life cycle (Abben et al., 2013). Reactance can be aroused if an employee feels that a freedom they used to have has been limited, for example as they are forced to use gender-inclusive language in their communication. This motivates the employee towards restoring their freedom (Brehm, 1989; Miron & Brehm, 2006), oftentimes accompanied with defensiveness (Abben et al., 2013). Evaluations of diversity initiatives therefore warn that change processes such as implementing gender-inclusive language take time, need to be considered within organizational politics, and resistance to change can be expected (Emmerich & Krell, 2002). Potential resistances include besides denial and avoidance also defiance or manipulation (Dass & Parker, 1999). Therefore, it is critical to understand diversity as a development process (Bierema, 2010; Voß & Würtemberger, 2023; Wiggins-Romesburg & Githens, 2018). This should be considered in future research, for example by investigating employees' reactions to changes in language use or guidelines and identifying best practices for successful diversity interventions.

Empirical Contributions

This dissertation makes several empirical contributions. Across all three manuscripts a variety of social psychological methods is applied, spanning from experiments to surveys to meta-analysis, showcasing different ways to address such very applied topics. Novel data are presented in Manuscripts 1 and 2 which are published open access to encourage re-analysis and assessment of other research questions. Further, Manuscripts 1 and 2 make use of integrative data analyses and small-scale meta-analyses as comprehensive ways to summarize research data. Manuscript 1 further applies brand new machine learning algorithms and cutting-edge Bayesian analyses. Manuscript 3 presents cumulative findings using up-to-date methods, namely multi-level meta-analysis with robust variance estimation. By itself, such a comprehensive summary of past research is a valuable source of information for scientists and practitioners in the field of diversity management.

Open Questions and Limitations

Despite the major contribution to the field of diversity management, the present dissertation comes with its limitations. As outlined in the introduction, gender is one of several diversity dimensions (Voß & Würtemberger, 2023). It is very likely that similar psychological hurdles as identified in Manuscripts 1 and 2 are present in other marginalized groups besides women. This should be the case as socialization processes and power mechanisms have been found to be similar across different marginalized groups (e.g., compare research on social class and gender as in Belmi & Laurin, 2016, and Manuscript 2). Beyond similarities between different marginalized groups, intersectionality could even intensify those hurdles (Di Stasio & Larsen, 2020; Shields, 2008). Further, approaches to fix systems should be directly applicable and also helpful for other diversity dimensions (e.g., using inclusive language beyond gender in communication such as saying that someone sits in a wheelchair instead of saying that someone is confined to a wheelchair). However, though many processes and boundary conditions seem similar, critical differences exist (see for example the stereotype content model,

Fiske et al., 2002). Therefore, the identified topics of psychological hurdles, organizational culture and inclusive language should also be investigated and evaluated focusing on other diversity dimensions as well as intersectionality. A possible way forward is adopting an “inclusion for all” framework where in addition to focusing on marginalized groups an intergroup focus is taken (Brannon et al., 2018).

The research presented in this dissertation has shown that the answer on the research question “depends”, for example on organizational culture or language, and that sometimes the matter seems more complex than at first sight (e.g., psychological hurdles). This calls for more research on boundary conditions of the topics researched and the proposed implications (e.g., What does it need that culture changes will work out? Unintended consequences?) as well as an examination of non-linear relationships.

Third, and very importantly, the present dissertation needs to be complemented by research with different samples, and in organizations. Though the scenarios in the experiments were created as realistic as possible, differences are conceivable and should be investigated. For job advertisements, for example, the qualifications section investigated in Manuscript 1 and the research investigating the language of job titles included in Manuscript 3 are only a part of job advertisements and interact with other parts (e.g., pictures) and information in addition to the job advertisement (e.g., research on the company by the applicant). Context could influence the relevance of the single aspect, and give it a stronger (e.g., high stakes culture) or weaker (e.g., growth mindset statement) influence compared to our experiments. Regarding Manuscript 2 it is conceivable that no company has a purely politics or purely competencies culture, but a mixture or with additional aspects included. Further, assessing real applicants and leaders with more and different experiences and priorities could pose a different picture.

Practical Implications

Gender Diversity Management Along the Employee Life Cycle

The present dissertation includes three manuscripts covering the hiring stage, leadership stage, as well as corporate communication covering all stages of the employee life cycle. As outlined in the introduction, the employee life cycle is a useful framework to identify important areas for research and interventions for diversity management. Taking the employee life cycle into account when thinking about diversity (management) can help organizations to identify relevant areas for action (Voß & Würtemberger, 2023). Further, the employee life cycle is usually backed with an abundance of data (e.g., demographics and performance of employees) that can and should be used for a strategic direction and evaluation of diversity management (O'Donovan, 2017; Inamdar & Abhi, 2020). While providing an overarching, integrative framework by investigating gender diversity along the employee life cycle, at the same time the three manuscripts that build this dissertation pinpoint to three very concrete topics relevant for diversity management and offer insights regarding potential interventions (discussed also in the individual manuscripts, see Appendix). To promote gender balance, the identified differing needs of female and male employees at the different stages of the employee life cycle (i.e., hiring, leadership, and overall) should be considered. This is outlined in further detail in the next two sections.

Three Manuscripts – Three Main Messages

Three main messages emerged for how to be attractive for a diverse workforce at all stages of the employee life cycle: address psychological hurdles in applicants, work on organizational leadership culture, use inclusive corporate language. Specifically, from Manuscript 1 organizations can take away that though gender differences in application intentions are not present, they may be more subtle in psychological hurdles. It can therefore be worthwhile looking for such subtleties and addressing those in the pursuit to hire the best talents. From Manuscript 2 organizations can take away that one avenue to motivate women and men (and

especially women) to pursue leadership is to work on their organizational culture towards focusing on showing competencies instead of playing politics. From Manuscript 3 organizations can take away that language has an impact on mental representations. If organizations want to promote gender diversity in mental representations, the results imply that instead of using masculine generics in corporate communication multi-gendering forms should be used.

Three Hot Topics

The three hot topics outlined in this dissertation can also be viewed from a practical point of view. First, organizations are advised to evaluate and make use of data in their diversity management. Second, organizations are advised to take different perspectives, and focus on changing systems, structures, and processes, to create more gender balance. Third, organizations are advised to critically reflect their communication, also beyond gender.

Conclusion

In this dissertation I described the theoretical and empirical background of gender diversity management along the employee life cycle. Further, three hot topics around gender diversity management in relation to three manuscripts were discussed. At the hiring stage of the employee life cycle in Manuscript 1 the hot topic that an empirical foundation is needed for diversity management is picked up. Specifically, Manuscript 1 revealed no gender difference in application intention depending on qualification fit (which was contrary to a popular claim) and at the same time women robustly wanted to be more prepared for a job than men. At the development stage of the employee life cycle in Manuscript 2 the hot topic that not women need to be fixed but systems and cultures was picked up. Specifically, Manuscript 2 showed that whether a gender difference in leadership aspiration exists depended on organization's culture. In competencies cultures women and men equally aspired to become leaders, whereas in politics cultures women aspired less than men to become leaders. Spanning all stages of the employee life cycle in Manuscript 3 data on gender-inclusive corporate language was

accumulated. Meta-analyzing past research on gender-inclusive vs. masculine generic language forms gave evidence for a higher mental representation of women in reaction to gender-inclusive language compared to masculine generics.

References

- AAUW. (2022). *The simple truth about the gender pay gap*. https://www.aauw.org/app/uploads/2022/12/SimpleTruth_12.22_2.1-002.pdf
- Abben, D. R., Brown, S. G., Graupmann, V., Mockler, S. A., & Fernandes, G. F. (2013). Drawing on social psychology literature to understand and reduce workplace discrimination. *Industrial and Organizational Psychology*, 6(4), 476–479.
- Alon, T., Doepke, M., Olmstaed-Rumsey, J., & Tertilt, M., (2020). *The impact of COVID-19 on gender equality* [Working paper]. Department of Economics, University of California San Diego.
- Arena Jr, D. F., Volpone, S. D., & Jones, K. P. (2023). (Overcoming) maternity bias in the workplace: A systematic review. *Journal of Management*, 49(1), 52–84. <https://doi.org/10.1177/01492063221086243>
- Baron, J. A., Ganglmair, B., Persico, N., Simcoe, T., & Tarantino, E. (2021). Representation is not sufficient for selecting gender diversity (No. w28649). *National Bureau of Economic Research*. <https://www.nber.org/papers/w28649>
- Becton, J. B., Walker, H. J., & Jones-Farmer, A. (2014). Generational differences in workplace behavior. *Journal of Applied Social Psychology*, 44(3), 175–189.
- Belmi, P., & Laurin, K. (2016). Who wants to get to the top? Class and lay theories about power. *Journal of Personality and Social Psychology*, 111(4), 505–529. <https://doi.org/10.1037/pspi0000060>
- Bennett, M., Sani, F., Hopkins, N., Agostini, L., & Malucchi, L. (2000). Children's gender categorization: An investigation of automatic processing. *British Journal of Developmental Psychology*, 18(1), 97–102. <https://doi.org/10.1348/026151000165599>

- Berkessel, J. B., Salwender, M., Gebauer, J. E., Stahlberg, D., Bleidorn, W., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2023). *Gender differences in agency and communion across 137 nations: A competitive test of five rival hypotheses* [Manuscript in preparation]. Department of Social Psychology, University of Mannheim.
- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, 142(11), 1227–1274. <https://doi.org/10.1037/bul0000067>
- Bierema, L. L. (2010). Resisting HRD's resistance to diversity. *Journal of European Industrial Training*, 34(6), 565–576.
- Bigler, R. S. (1999). The use of multicultural curricula and materials to counter racism in children. *Journal of Social Issues*, 55(4), 687–705. <https://doi.org/10.1111/0022-4537.00142>
- Bosak, J., & Sczesny, S. (2011). Gender bias in leader selection? Evidence from a hiring simulation study. *Sex Roles: A Journal of Research*, 65(3-4), 234–242. <https://doi.org/10.1007/s11199-011-0012-7>
- Brannon, T. N., Carter, E. R., Murdock-Perriera, L. A. and Higginbotham, G. D. (2018). From backlash to inclusion for all: Instituting diversity efforts to maximize benefits across group lines. *Social Issues and Policy Review*, 12, 57–90. <https://doi.org/10.1111/sipr.12040>
- Brehm, J. W. (1989). Psychological reactance: Theory and applications. In T. K. Srull (Ed.), *Advances in Consumer Research*, (Vol. 16, pp. 72–75). Association for Consumer Research.
- Buss, D. M. (2005). *The evolutionary psychology handbook*. Wiley.

- Buss, D. M., Durkee, P. K., Shackelford, T. K., Bowdle, B. F., Schmitt, D. P., Brase, G. L., Choe, J. C., & Trofimova, I. (2020). Human status criteria: Sex differences and similarities across 14 nations. *Journal of Personality and Social Psychology*, 119(5), 979–998. <https://doi.org/10.1037/pspa0000206>
- Caleo, S., & Heilman, M. E. (2019). What could go wrong? Some unintended consequences of gender bias interventions. *Archives of Scientific Psychology*, 7(1), 71–80. <https://doi.org/10.1037/arc0000063>
- Carter, A. B., & Phillips, K. W. (2017). The double-edged sword of diversity: Toward a dual pathway model. *Social and Personality Psychology Compass*, 11(5), Article e12313. <https://doi.org/10.1111/spc3.12313>
- Chaney, K. E. (2022). An examination of diversity rationales: How instrumental and moral diversity rationales create minority spotlight. *European Journal of Social Psychology*, 52(5-6), 783–796. <https://doi.org/10.1002/ejsp.2852>
- Chaney, K. E., Rudman, L. A., Fetterolf, J. C., & Young, D. M. (2019). Paying a price for domestic equality: Risk factors for backlash against nontraditional husbands. *Gend. Issues*, 36, 3–22. <https://doi.org/10.1007/s12147-017-9207-8>
- Charta der Vielfalt. (2022). *Diversity factbook*. https://www.charta-der-vielfalt.de/fileadmin/user_upload/Diversity-Tag/2022/Deutscher_Diversity-Tag_2022/Factbook_2022_EN.pdf
- Charta der Vielfalt. (2023). <https://www.charta-der-vielfalt.de/en/for-employers/diversity-dimensions/gender/>
- Croft, A., Schmader, T., & Block, K. (2015). An underexamined inequality: Cultural and psychological barriers to men's engagement with communal roles. *Personality and Social Psychology Review*, 19(4), 343–370. <https://doi.org/10.1177/1088868314564789>

- Cox, T. H., & Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *The Executive*, 5(3), 45–56.
- Dass, P., & Parker, B. (1999). Strategies for managing human resource diversity: From resistance to learning. *Academy of Management Perspectives*, 13(2), 68–80.
- de Vries, A.L.C., Kreukels, B.P.C., Steensma, T.D., McGuire, J.K. (2014). Gender identity development: A biopsychosocial perspective. In B. Kreukels, T. Steensma, A. de Vries (Eds.), *Gender dysphoria and disorders of sex development. Focus on sexuality research*. Springer. https://doi.org/10.1007/978-1-4614-7441-8_3
- Deaux, K. (1985). Sex and gender. *Annual Review of Psychology*, 36(1), 49–81.
- Deshpande, R., & Webster, F. E. (1989). Organizational culture and marketing: Defining the research agenda. *Journal of Marketing*, 53(1), 3–15. <https://doi.org/10.2307/1251521>
- Di Stasio, V., & Larsen, E. N. (2020). The racialized and gendered workplace: applying an intersectional lens to a field experiment on hiring discrimination in five European labor markets. *Social Psychology Quarterly*, 83(3), 229–250.
<https://doi.org/10.1177/0190272520902994>
- Dixon-Fyle, S., Dolan, K., Hunt, D. V., & Prince, S. (2020, May 19). *Diversity wins: How inclusion matters*. <https://www.mckinsey.com/featured-insights/diversity-and-inclusion/diversity-wins-how-inclusion-matters>
- Dobbin, F., Kim, S., & Kalev, A. (2011). You can't always get what you need: Organizational determinants of diversity programs. *American Sociological Review*, 76(3), 386–411.
<https://doi.org/10.1177/0003122411409704>
- Dover, T. L., Kaiser, C. R., & Major, B. (2020). Mixed signals: The unintended effects of diversity initiatives. *Social Issues and Policy Review*, 14(1), 152–181.
<https://doi.org/10.1111/sipr.12059>

- Eagly, A. H., Johannesen-Schmidt, M. C., & Van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, 129(4), 569–591. <https://doi.org/10.1037/0033-2909.129.4.569>
- Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, 60(5), 685–710. <https://doi.org/10.1037/0022-3514.60.5.685>
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, 117(1), 125–145. <https://doi.org/10.1037/0033-2909.117.1.125>
- Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage Publications Ltd. <https://doi.org/10.4135/9781446249222.n49>
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Lawrence Erlbaum Associates Publishers.
- Ehrke, F., Ashoe, A., Steffens, M. C., & Louvet, E. (2020). A brief diversity training: Raising awareness of ingroup privilege to improve attitudes towards disadvantaged outgroups. *International Journal of Psychology*, 55(5), 732–742. <https://doi.org/10.1002/ijop.12665>
- Emmerich, A., & Krell, G. (2002). Diversity-orientierte Trainings und Beurteilungen von Führungskräften [Diversity-oriented training and assessment of managers]. *Wirtschaftspsychologie*, 4(1), 64–69.

- England, P., Levine, A., & Mishel, E. (2020). Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences*, 117(13), 6990–6997. <https://doi.org/10.1073/pnas.1918891117>
- Eidlyn, A. (2022, July 29). Gericht weist Klage gegen gendergerechte Sprache bei Audi ab [Court rejects complaint against gender-equitable language at Audi]. *Zeit Online*. https://www.zeit.de/gesellschaft/zeitgeschehen/2022-07/gender-sprache-klage-audi-vw-ingolstadt?utm_referrer=https%3A%2F%2Fwww.google.com%2F
- Faravelli, C., Scarpato, M. A., Castellini, G., & Sauro, C. L. (2013). Gender differences in depression and anxiety: the role of age. *Psychiatry research*, 210(3), 1301–1303. <https://doi.org/10.1016/j.psychres.2013.09.027>
- Fenton, S., & Dermott, E. (2006). Fragmented careers? Winners and losers in young adult labour markets. *Work, Employment & Society*, 20, 205–221. <https://doi.org/10.1177/0950017006064111>
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. <https://doi.org/10.1037//0022-3514.82.6.878>
- Fluchtmann, J., Glenney, A., Harmon, N. A., & Maibom, J. (2022). The gender application gap: Do men and women apply for the same jobs?. *IZA Discussion Paper*, Article 14906. <http://dx.doi.org/10.2139/ssrn.4114410>
- Fox, C. (2017). *Stop fixing women: Why building fairer workplaces is everybody's business*. Newsouth.
- Georgeac, O. A. M., & Rattan, A. (2023). The business case for diversity backfires: Detrimental effects of organizations' instrumental diversity rhetoric for underrepresented

- group members' sense of belonging. *Journal of Personality and Social Psychology*, 124(1), 69–108. <https://doi.org/10.1037/pspi0000394>
- Gino, F., Wilmut, C. A., & Brooks, A. W. (2015). Compared to men, women view professional advancement as equally attainable, but less desirable. *Proceedings of the National Academy of Sciences*, 112, 12354–12359. <https://doi.org/10.1073/pnas.1502567112>
- Gilbert, J. A., Stead, B. A., & Ivancevich, J. M. (1999). Diversity management: A new organizational paradigm. *Journal of Business Ethics*, 21, 61–76. <https://doi.org/10.1023/A:1005907602028>
- Gladka, O., Fedorova, V., & Dohadailo, Y. (2022). Development of conceptual bases of the employee life cycle within an organization. *Business: Theory and Practice*, 23(1), 39–52. <https://doi.org/10.3846/btp.2022.13510>
- Guentherodt, I., Hellinger, M., Pusch, L. F., & Tromel-Plotz, S. (1980). Richtlinien zur Vermeidung sexistischen Sprachgebrauchs [Guidelines for avoiding sexist language use]. *Linguistische Berichte Braunschweig*, 69, 15–21.
- Haas, C., & Vetter, P. (2021, July 5). Jede dritte deutsche Firma gendert – aber viele nur nach außen [Every third German company uses gender-inclusive languages – however many only in their external communication]. *Welt*. <https://www.welt.de/wirtschaft/article232305221/Gender-Sprache-Jede-dritte-deutsche-Firma-verwendet-neutrale-Sprache.html>
- Hardy III, J. H., Tey, K. S., Cyrus-Lai, W., Martell, R. F., Olstad, A., & Uhlmann, E. L. (2022). Bias in context: Small biases in hiring evaluations have big consequences. *Journal of Management*, 48(3), 657–692. <https://doi.org/10.1177/0149206320982654>

- Heitner, K. L., Kahn, A. E., & Sherman, K. C. (2013). Building consensus on defining success of diversity work in organizations. *Consulting Psychology Journal: Practice and Research*, 65(1), 58–73. <https://doi.org/10.1037/a0032593>
- Hennig, S. (2019, June 12). *It turns out working mothers can't have it all*. <https://medium.com/swlh/it-turns-out-working-mothers-cant-have-it-all-a8501da99f3a>
- Hentschel, T., Horvath, L. K., Peus, C., & Sczesny, S. (2018). Kick-starting female careers. *Journal of Personnel Psychology*, 17(4), 193–203. <https://doi.org/10.1027/1866-5888/a000209>
- Herring, C. (2009). Does diversity pay?: Race, gender, and the business case for diversity. *American Sociological Review*, 74(2), 208–224. <https://doi.org/10.1177/000312240907400203>
- Horvath, L. K., & Sczesny, S. (2016). Reducing women's lack of fit with leadership positions? Effects of the wording of job advertisements. *European Journal of Work and Organizational Psychology*, 25(2), 316–328. <https://doi.org/10.1080/1359432X.2015.1067611>
- Huang, J., Krivkovich, A., Rambachan, I., & Yee, L. (2021, May 5). For mothers in the workplace, a year (and counting) like no other. *McKinsey & Company*. <https://www.mckinsey.com/featured-insights/diversity-and-inclusion/for-mothers-in-the-workplace-a-year-and-counting-like-no-other>
- Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, 74(2), 171–193. <https://doi.org/10.1037/amp0000307>
- IHK Wiesbaden. (2023). *Allgemeines Gleichbehandlungsgesetz (AGG): Kurzer Überblick* [Law for equal treatment: Short overview].

<https://www.ihk.de/wiesbaden/recht/rechtsberatung/personal/allegemeines-gleichbehandlungsgesetz-agg-kurzer-ueberblick-1255692>

- Inamdar, S. P., & Abhi, S. (2020). A strategic approach to HR analytics: A framework for employee life cycle model. *Journal of Management and Research*, 14(1), 1–10.
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, 85(4), 616–626. <https://doi.org/10.1037/0022-3514.85.4.616>
- Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal*, 52(3), 599–627. <https://doi.org/10.5465/AMJ.2009.41331491>
- Kaley, A., & Dobbin, F. (2020, October 20). Companies need to think bigger than diversity training. *Harvard Business Review*. <https://hbr.org/2020/10/companies-need-to-think-bigger-than-diversity-training?fbclid=IwAR25rkMgChiLQdLKL-BuqzFoS9V4KI4Jr863LSwpPmm1U5LqViIQ-SuTKxyE>
- Kalinoski, Z. T., Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. A. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organizational Behavior*, 34(8), 1076–1104. <https://doi.org/10.1002/job.1839>
- Kim, J. Y., Fitzsimons, G. M., & Kay, A. C. (2018). Lean in messages increase attributions of women's responsibility for gender inequality. *Journal of Personality and Social Psychology*, 115(6), 974–1001. <https://doi.org/10.1037/pspa0000129>
- Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, 100(1), 128–161. <https://doi.org/10.1037/a0036734>

- Kulich, C., & Bosak, J. (2019). The persistence of gender inequality in leadership: Still a long way to go? In K. Gangl & E. Kirchler (Eds.), *A research agenda for economic psychology* (pp. 139–153). Edward Elgar Publishing.
<https://doi.org/10.4337/9781788116060.00015>
- Kurfer, T. (2022, July 22). Streit ums Gendern: Nein, die deutsche Sprache diskriminiert Frauen nicht [Controversy over gendering: No, the German language does not discriminate against women]. *Berliner Zeitung*. <https://www.berliner-zeitung.de/open-source/streit-ums-gendern-nein-die-deutsche-sprache-diskriminiert-frauen-nicht-li.246245>
- Ley, T. J., & Hamilton, B. H. (2008). The gender gap in NIH grant applications. *Science*, 322(5907), 1472–1474.
- Lindqvist, A., Renström, E. A., & Gustafsson Sendén, M. (2019). Reducing a male bias in language? Establishing the efficiency of three different gender-fair language strategies. *Sex Roles: A Journal of Research*, 81(1-2), 109–117. <https://doi.org/10.1007/s11199-018-0974-9>
- Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British Journal of Psychology*, 92(1), 239–255.
<https://doi.org/10.1348/000712601162059>
- Martin, C. L., & Ruble, D. N. (2010). Patterns of gender development. *Annual Review of Psychology*, 61, 353–381. <https://doi.org/10.1146/annurev.psych.093008.100511>
- Matlin, M. W. (2004). *The psychology of women* (5th edition). Wadsworth/Thomson Learning.

- McCauley, C., Wright, M., & Harris, M. E. (2000). Diversity workshops on campus: A survey of current practice at U.S. colleges and universities. *College Student Journal*, 34(1), 100–114.
- Miron, A. M., & Brehm, J. W. (2006). Reactance theory-40 years later. *Zeitschrift für Sozialpsychologie*, 37(1), 9–18. <https://doi.org/10.1024/0044-3514.37.1.3>
- Mohr, T. (2014, August 25). Why women don't apply for jobs unless they're 100% qualified. *Harvard Business Review*. <https://hbr.org/2014/08/why-women-dont-apply-for-jobs-unless-theyre-100-qualified>
- Mor Barak, M. E., Lizano, E. L., Kim, A., Duan, L., Rhee, M. K., Hsiao, H. Y., & Brimhall, K. C. (2016). The promise of diversity management for climate of inclusion: A state-of-the-art review and meta-analysis. *Human Service Organizations: Management, Leadership & Governance*, 40(4), 305–333. <https://doi.org/10.1080/23303131.2016.1138915>
- Moreno-Agostino, D., Wu, Y. T., Daskalopoulou, C., Hasan, M. T., Huisman, M., & Prina, M. (2021). Global trends in the prevalence and incidence of depression: A systematic review and meta-analysis. *Journal of affective disorders*, 281, 235–243. <https://doi.org/10.1016/j.jad.2020.12.035>
- Morgenroth, T., Kirby, T. A., Ryan, M. K., & Sudkämper, A. (2020). The who, when, and why of the glass cliff phenomenon: A meta-analysis of appointments to precarious leadership positions. *Psychological Bulletin*, 146(9), 797–829. <https://doi.org/10.1037/bul0000234>
- Moser, C. E., & Branscombe, N. R. (2022). Male allies at work: Gender-equality supportive men reduce negative underrepresentation effects among women. *Social Psychological and Personality Science*, 13(2), 372–381. <https://doi.org/10.1177/19485506211033748>

- Moss-Racusin, C. A. (2014). Male backlash: Penalties for men who violate gender stereotypes. In R. J. Burke & D. A. Major (Eds.), *Gender in organizations: Are men allies or adversaries to women's career advancement?* (pp. 247–269). Edward Elgar Publishing. <https://doi.org/10.4337/9781781955703.00021>
- O'Donovan, D. (2017). Inclusion: Diversity Management 2.0. In C. Machado, J. Davim (Eds.), *Managing Organizational Diversity*. Springer, Cham. https://doi.org/10.1007/978-3-319-54925-5_1
- Paustian-Underdahl, S. C., Walker, L. S., & Woehr, D. J. (2014). Gender and perceptions of leadership effectiveness: A meta-analysis of contextual moderators. *Journal of Applied Psychology*, 99(6), 1129–1145. <https://doi.org/10.1037/a0036751>
- Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *The British Journal of Psychiatry*, 177(6), 486–492. <https://doi.org/10.1192/bjp.177.6.486>
- Pless, N., & Maak, T. (2004). Building an inclusive diversity culture: Principles, processes and practice. *Journal of Business Ethics*, 54, 129–147. <https://doi.org/10.1007/s10551-004-9465-8>
- Prati, F., Menegatti, M., Moscatelli, S., Kana Kenfack, C. S., Pireddu, S., Crocetti, E., Mariani, M. G., & Rubini, M. (2019). Are mixed-gender committees less biased toward female and male candidates? An investigation of competence-, morality-, and sociability-related terms in performance appraisal. *Journal of Language and Social Psychology*, 38(5-6), 586–605. <https://doi.org/10.1177/0261927X19844808>
- Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, 31(9), 1109–1121. <https://doi.org/10.1068/p3331>

- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology*, 81(5), 856–868. <https://doi.org/10.1037/0022-3514.81.5.856>
- Rutherford, S. (2014). Gendered organizational cultures, structures and processes: The cultural exclusion of women in organizations. In R. J. Burke & D. A. Major (Eds.), *Gender in Organizations* (pp. 193–216). Edward Elgar Publishing.
- Saltmarsh, M. (2017). 5 Stages of the employment life cycle in need of improvement. *Hospitality Review*, February, 8–9.
- Sandberg, S. (2015). *Lean in: Women, work, and the will to lead*. WH Allen.
- Schmader, T., Dennehy, T. C., & Baron, A. S. (2022). Why antibias interventions (need not) fail. *Perspectives on Psychological Science*, 17(5), 1381–1403. <https://doi.org/10.1177/17456916211057565>
- Schuh, S. C., Hernandez Bark, A. S., Van Quaquebeke, N., Hossiep, R., Frieg, P., & Van Dick, R. (2014). Gender differences in leadership role occupancy: The mediating role of power motivation. *Journal of Business Ethics*, 120, 363–379. doi: 10.1007/s10551-013-1663-9
- Shields, S. A. (2008). Gender: An intersectionality perspective. *Sex Roles*, 59, 301–311. <https://doi.org/10.1007/s11199-008-9501-8>
- Singal, M. (2014). The business case for diversity management in the hospitality industry. *International Journal of Hospitality Management*, 40, 10–19. <http://dx.doi.org/10.1016/j.ijhm.2014.02.009>
- Spencer, S. J., Logel, C., & Davies, P. G., (2016). Stereotype threat. *Annual Review of Psychology*, 67, 415–437. <https://doi.org/10.1146/annurev-psych-073115-103235>

Stahlberg, D. (2021, October 1). *Immer noch wenige Frauen in Führungspositionen – Warum und was können wir dagegen tun?* [Still few women in leadership positions – Why and what can we do about it?]. DENK FABRIK Chancengleichheit, Freiburg.

Starck, J. G., Sinclair, S., & Shelton, J. N. (2021). How university diversity rationales inform student preferences and outcomes. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 118(16), Article e2013833118.
<https://doi.org/10.1073/pnas.2013833118>

Statistisches Bundesamt. (2021). *Wie stehen Sie als Unternehmen derzeit zum Einsatz geschlechtergerechter Sprache?* [What is your current position as a company on the use of gender-inclusive language?]. <https://de.statista.com/statistik/daten/studie/1235488/umfrage/gendersprache-in-dax-unternehmen/>

Statistisches Bundesamt. (2023a). *Diversity in der Arbeitswelt: Vielfalt ist ein Erfolgsfaktor für...* [Diversity in the world of work: Diversity is a success factor for...]. <https://de.statista.com/statistik/daten/studie/1272094/umfrage/diversity-als-erfolgssfaktor-in-der-arbeitswelt/>

Statistisches Bundesamt. (2023b). *Erwerbstätigenquote von Männern und Frauen mit minderjährigen Kindern im Haushalt in Deutschland von 2009 bis 2019* [Employment rate of men and women with underage children in the household in Germany from 2009 to 2019]. <https://de.statista.com/statistik/daten/studie/972869/umfrage/erwerbstaetigenquote-von-maennern-und-frauen-mit-minderjaehrigen-kindern/>

Statistisches Bundesamt. (2023c). *Was meinen Sie, wo herrscht Nachholbedarf deutscher Unternehmen beim Thema Diversity?* [Where do you think German companies need to catch up when it comes to diversity?]. <https://de.statista.com/statistik/daten/studie/1272107/umfrage/nachholbedarf-diversity-in-deutschen-unternehmen/>

- Statistisches Bundesamt. (2023d). *Wie bewerten Sie die folgenden Aussagen zu Diversity bei Ihrem aktuellen Arbeitgeber?* [How would you rate the following statements about diversity at your current employer?]. <https://de.statista.com/statistik/daten/studie/1226531/umfrage/diversity-management-in-unternehmen-in-deutschland-grossbritannien-und-frankreich/>
- van Dijk, H., Kooij, D., Karanika-Murray, M., De Vos, A., & Meyer, B. (2020). Meritocracy a myth? A multilevel perspective of how social inequality accumulates through work. *Organizational Psychology Review*, 10(3-4), 240–269. <https://doi.org/10.1177/2041386620930063>
- Van Hooft, E. A. J., Born, M. P., Taris, T. W., & Van der Flier, H. (2006). Ethnic and gender differences in applicants' decision-making processes: An application of the theory of reasoned action. *International Journal of Selection and Assessment*, 14(2), 156–166. <https://doi.org/10.1111/j.1468-2389.2006.00341.x>
- Voß, E., & Würtemberger, S. (2023). *Vielfalt im Employee Lifecycle* [Diversity in the employee lifecycle]. Springer. <https://doi.org/10.1007/978-3-658-39841-5>
- Vuletich, H. A., & Payne, B. K. (2019). Stability and change in implicit bias. *Psychological Science*, 30(6), 854–862. <https://doi.org/10.1177/0956797619844270>
- Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology*, 65(6), 1093–1104. <https://doi.org/10.1037/0022-3514.65.6.1093>
- Welpel, I., & Peus, C. (2014). *Gendergerechte Personalauswahl und -beförderung. Handreichung für EntscheidungsträgerInnen in Wirtschaft und Wissenschaft* [Gender-equitable personnel selection and promotion. Handout for decision-makers in business and science]. Technische Universität München.

https://www.chancengleichheit.tum.de/fileadmin/w00blt/www/Karriere_und_Weiterbildung/Gendergerechte_Personalauswahl_und_Befoerderung_AuBeFuehr.pdf

Wiggins-Romesburg, C. A., & Githens, R. P. (2018). The psychology of diversity resistance and integration. *Human Resource Development Review*, 17(2), 179–198.

<https://doi.org/10.1177/1534484318765843>

Winkler, B. (2015). Beobachtungen zur Auswahl- und Beurteilungspraxis: Plädoyer für den Einsatz einer evidenzbasierten Expertise [Observations on selection and assessment practices: a plea for the use of evidence-based expertise]. In I. M. Welp, P. Brosi, L. Ritzenhöfer, & T. Schwarzmüller (Eds.), *Auswahl von Männern und Frauen als Führungskräfte* (pp. 219–240). Springer.

World Economic Forum. (2022). *Global Gender Gap Report 2022*. <http://reports.weforum.org/globalgender-gap-report-2022>

Yadav, S., & Lenka, U. (2020). Diversity management: A systematic review. *Equality, Diversity, and Inclusion*, 39(8), 901–929. <https://doi.org/10.1108/EDI-07-2019-0197>

Yaish, M., Mandel, H., & Kristal, T. (2021). Has the economic lockdown following the Covid-19 pandemic changed the gender division of labor in Israel? *Gender & Society*, 35(2), 256–270. <https://doi.org/10.1177/08912432211001297>

Zehnter, M. K., Olsen, J., & Kirchler, E. (2018). Obituaries of female and male leaders from 1974 to 2016 suggest change in descriptive but stability of prescriptive gender stereotypes. *Frontiers in Psychology*, 9, Article 2286.

<https://doi.org/10.3389/fpsyg.2018.02286>

Statement of Originality

Declaration in lieu of oath according to section 9 subsection 1(e) of the Regulations and Procedures Governing the Doctoral Dissertation to earn a doctoral degree in Social Sciences at the University of Mannheim of June 27, 2013, 1st amendment of June 6, 2016.

I hereby declare that the presented doctoral dissertation with the title “Promoting Gender Equality at Different Stages of the Employee Life Cycle: New and Cumulated Findings” is my own work.

- I did not seek unauthorized assistance of a third party and I have employed no other sources or means except the ones listed. I clearly marked any quotations derived from the works of others.
- I did not yet present this doctoral dissertation or parts of it at any other higher education institution in Germany or abroad.
- I hereby confirm the accuracy of the declaration above.
- I am aware of the significance of this declaration and the legal consequences in case of untrue or incomplete statements.

I affirm in lieu of oath that the statements above are to the best of my knowledge true and complete.

Mona Salwender, Mannheim

May 2nd, 2023

Statements of the Co-Authors

Statement of the Co-Authors of Manuscript 1

It is hereby confirmed that Mona Salwender primarily contributed to the following manuscript (conceptualization, data preparation and analyses, writing):

Salwender, M. & Stahlberg, D (2023). Do women only apply when they are 100% qualified, whereas men already apply when they are 60% qualified? [Manuscript in preparation].

Department of Social Psychology, University of Mannheim.

I sign this statement to the effect that Mona Salwender is credited as the primary source of ideas and the main author of this manuscript.

Statement of the Co-Authors of Manuscript 2

It is hereby confirmed that Mona Salwender primarily contributed to the following manuscript (conceptualization, data preparation and analyses, writing):

Salwender, M., Schoel, C., Bless, H., & Stahlberg, D. (2023). The politics hurdle: Joint effect of organizational culture and gender on lack of fit experiences. *Social Psychological and Personality Science*, 14(1), 84–92. <https://doi.org/10.1177/19485506221075898>

I sign this statement to the effect that Mona Salwender is credited as the primary source of ideas and the main author of this manuscript.

Statement of the Co-Authors of Manuscript 3

It is hereby confirmed that Mona Salwender primarily contributed to the following manuscript (conceptualization, data preparation and analyses, drafting methods, results, and discussion sections of the manuscript, revision of the manuscript):

Salwender, M., Maschmann, I. T., Heinzl, V. K., Sczesny, S., & Stahlberg, D (2023).

Masculine generics versus gender-inclusive language: A meta-analysis of the effects of language form on the mental representation of gender [Manuscript in preparation].

Department of Social Psychology, University of Mannheim.

I sign this statement to the effect that Mona Salwender is credited as the primary source of ideas and the main author of this manuscript.

Appendix: Manuscripts

The three manuscripts that form this dissertation are attached in the appendix in the following order:

Salwender, M., & Stahlberg, D. (2023). *Do women only apply when they are 100% qualified, whereas men already apply when they are 60% qualified?* [Manuscript in preparation]. Department of Social Psychology, University of Mannheim.

Salwender, M., Schoel, C., Bless, H., & Stahlberg, D. (2023). The politics hurdle: Joint effect of organizational culture and gender on lack of fit experiences. *Social Psychological and Personality Science*, 14(1), 84–92. <https://doi.org/10.1177/19485506221075898>

Salwender, M., Maschmann, I. T., Heinzl, V. K., Sczesny, S., & Stahlberg, D. (2023). *Masculine generics versus gender-inclusive language: A meta-analysis of the effects of language form on the mental representations of gender* [Manuscript in preparation]. Department of Social Psychology, University of Mannheim.

The online supplements can be retrieved from:

https://osf.io/ed8w7/?view_only=820b4c2e1c3f44f08b15aed028b82c05 (Manuscript 1)

https://osf.io/astpf/?view_only=380a63dde1ee4f418ef6ae740a8c8085 (Manuscript 2)

https://osf.io/ducgp/?view_only=bbe7116c8a0940a9b01bab21afc1a6e2 (Manuscript 3)

**Do Women Only Apply When They Are 100% Qualified, Whereas Men Already Apply
When They Are 60% Qualified?**

Mona Salwender and Dagmar Stahlberg

Department of Social Psychology, University of Mannheim

Author Note

Mona Salwender  <https://orcid.org/0000-0001-8431-7707>

Dagmar Stahlberg  <https://orcid.org/0000-0001-5972-0641>

Studies 2-7 were preregistered on osf.io (see links in the online supplement), data and materials are also available on https://osf.io/ed8w7/?view_only=03ca36afc527405d927bb326a18c052e. We have no known conflict of interest to disclose. This research was supported by departmental budget resources of the Chair for Social Psychology. We thank Vanessa Borytzka, Hannah Jansen, Lena Grobe, Melina Welt, Linnea Tönjes, Sophia Wolf, and Verena Hofmann for programming the surveys and collecting the data as part of their Bachelor's or Master's theses.

Correspondence concerning this article should be addressed to Mona Salwender, Social Psychology, University of Mannheim, A5 6, 68165 Mannheim, Germany, Email: mona.salwender@gmail.com

Abstract

We tested the popular claim that women only apply for jobs when they are 100% qualified, whereas men already apply with a 60% qualification fit. In Studies 1-3, we presented a job advertisement and a CV with different levels of qualification fit. Participants should imagine the CV presented as their own and were to indicate whether they would apply for the advertised job. No gender difference emerged in participants' application intentions at any level of qualification fit. In Studies 4-6 we presented a job advertisement and asked participants to indicate whether they themselves would apply for the advertised job. Afterwards, participants indicated for every qualification criterion listed in the job advertisement whether they fulfilled it or not. Again, we did not find a consistent gender difference in application intention depending on qualification fit. However, when asking how much women and men want to be prepared in application situations, women robustly indicated a higher desire for preparedness than men. In Study 7, we therefore focused on the gender difference in desire for preparedness and competitively tested 29 potential mediators. We found women's higher desire for preparedness to be explained by their higher fear of not being able to meet expectations. Overall, our results indicate that higher psychological hurdles (i.e., desire for preparedness, fear of not being able to meet expectations) are present in application situations for women than for men. However, these do not seem to translate into differential application intentions. We discuss the theoretical and practical implications of our findings.

Keywords: application intention, qualification fit, person-job fit, job advertisement, desire for preparedness, gender, gender difference

Do Women Only Apply When They Are 100% Qualified, Whereas Men Already Apply When They Are 60% Qualified?

“You’ve probably heard of the following statistic: Men apply for a job when they meet only 60% of the qualifications, but women apply only if they meet 100% of them.” (Mohr, 2014, p. 1). This statistic is mentioned in Sheryl Sandberg’s book “Lean in” (Sandberg, 2015), where it is used to explain gender differences in career paths. After the book was published, this statistic has received a lot of media coverage and has been cited extensively in online magazine and blog articles (e.g., Hannon, 2014; Rojas, 2021; Youn, 2019). However, journalists later discovered that this statistic was based on a speculative comment instead of quantitative data and that relevant quantitative data are still scarce (Hacohen & Nicks, 2019; Rice, 2014).

We address this research gap of lacking quantitative data. In the following, we report seven studies testing whether women only apply when they are 100% qualified, whereas men already apply with a 60% qualification fit. If this gender difference does exist, this would have important implications. First, organizations could use this knowledge to prevent losing potential talents and to maximize the quality of hires. They could, for example, redesign job advertisements and recruiting processes to prevent female applicants with an acceptable or even high fit from self-selecting out of the applicant pool. Such redesigns would directly support diversity goals many organizations have, as a higher rate of females in the applicant pool has been found to predict more female hires (Schmidt & Stettes, 2018). At the same time such interventions could prevent that (male) applicants with a low fit self-select into the applicant pool. By reducing the number of low-fitting applications that must be reviewed, the organization could save time (and therefore money; Böttcher, 2017). Second, female and male applicants could learn from respective findings that different application strategies exist (e.g., applying selectively, which may be a wise investment of applicants’ time, but may entail the risk of lost opportunities vs. applying extensively with a high investment of time and the chance to

meet plenty interesting opportunities). Participants, then, could reflect which strategy they are using, and make a conscious decision which strategy they want to use. Third, federal and private institutions (e.g., federal employment agencies, business consultants, application coaches) could apply this knowledge about gender differences in this field in the same line as organizations and individuals, depending on who their clients are.

In contrast, if the empirical studies reported in this line of research would not confirm a gender difference, and this null finding is replicated in future research, authors of online magazine and blog articles should refrain from citing the statistic. Otherwise, these (misleading) citations may unjustifiably create stereotype threat situations for female applicants and toxic masculinity for male applicants (e.g., pressure to behave confidently). However, either way, organizations may be well advised to check for gender imbalances in their applicant pool as many factors are known that prevent disproportionately more female than male applicants from applying for jobs (described in the following section in more detail; e.g., Born & Taris, 2010).

Application Intention and Gender

Past research on application intention and gender has found similarities as well as differences between female and male applicants. User analyses from the job search portal LinkedIn have shown that when women and men browse job postings, they both look on average at 44-46 job postings (Tockey & Ignatova, n.d.). While behaving quite similar in the job search, women applied to 20% fewer jobs than men (Tockey & Ignatova, n.d.). Barbulescu and Bidwell (2013) have shown a similar pattern. In their study women were less likely than men to apply to management jobs in finance (though women and men were similarly qualified), partly because they expected lower success of their application. Yet, women were not less likely to receive a job offer in finance than men.

Quite a few studies have meanwhile confirmed that women's application intention is more dependent on context compared to men's application intention. The type of job or industry (Barbulescu & Bidwell, 2013; Fluchtmann et al., 2022), for example, has been found to influence women's application intentions more strongly compared to men's application intention. Further, organizational culture has been identified to influence female applicants' intention to seek power positions more strongly than those of male applicants (Salwender et al., 2023). Besides such context variables, the effects of design of job advertisements have been researched intensively. Women's, but not men's, application intention increased when the required qualifications listed in the job advertisement were worded in terms of behaviors (e.g., "able to generate original ideas") instead of personal attributes (e.g., "possess high levels of creativity"; Born & Taris, 2010; Wille & Derous, 2018). Similarly, women's, but not men's, interest in a job increased when the wording used in the job advertisement carried feminine (e.g., "we are committed...") instead of masculine connotations (e.g., "we are determined..."; Gaucher et al., 2011). Such gender differences can even be found in children as girls', but not boys', interest in stereotypically male occupations was higher when gender-inclusive job titles (e.g., "businesswomen and businessmen") were used compared to masculine generics (e.g., "businessmen"; Vervecken et al., 2013).

Studying gender differences in application intention is insofar important as consequences of gender differences in application intention can accumulate in the gender pay gap (Fluchtmann et al., 2022) and the gender leadership gap (Schmidt & Stettes, 2018). Before we discuss the question whether women want to have a higher qualification fit when applying for a job than men, we will next define and shortly review research on qualification fit.

Qualification Fit Effects

Qualification fit can be classified as a special case of person-job fit, which has been researched extensively. Person-job fit is defined as “relationship between a person’s characteristics and those of the job or tasks that are performed at work” (Kristof-Brown et al., 2005, p. 284) and is divided into two subtypes. The first subtype, demands-abilities fit, describes fit between a person’s individual abilities and job demands (e.g., a person has profound knowledge in statistics and part of the job is to interpret statistics; Kristof-Brown et al., 2005). Qualification fit is a special case of this subtype, referring to a specific job advertisement. The second subtype, needs-supplies fit, describes whether a person’s individual needs are met by a job (e.g., having a challenging job, having flexible working hours; Kristof-Brown et al., 2005; Travaglianti et al., 2016).

Summarizing 62 studies and 225 effect sizes, a meta-analysis of both subtypes of person-job fit showed that higher person-job fit was associated with higher organizational attraction of potential applicants and a stronger intent of organizations to hire an applicant (both pre-entry assessments, i.e., before an individual enters an organization). Further, higher person-job fit was associated with higher job satisfaction, higher organizational commitment, lower intention to quit, higher satisfaction with coworkers and supervisors, higher organizational identification, higher performance, and lower strain (all post-entry, Kristof-Brown et al., 2005). In the pre-entry phase, recent studies additionally showed that higher person-job fit also went along with higher application intentions (Van Hooft et al., 2006) and higher intentions to accept a job offer (Carless, 2005, Chapman et al., 2005).

Therefore, in general, we expect a positive relationship of qualification fit and application intention across both genders. However, in the following section we present evidence suggesting gender differences in the qualification fit – application intention relation with a stronger relation for women than for men.

Gender Difference in Desire for Preparedness

Multiple studies have investigated to what extent men and women feel prepared to confront new tasks in education and job environments and have predominantly found men perceiving themselves to be more prepared than women (e.g., for entering graduate school; García-Aracil et al., 2018; Mohan & Sundari Ravindran, 2018; Shaw et al., 2013; Svirko et al., 2014; Thandi & Sharma, 2004). Directly relevant to our main hypothesis is Lee's (2018) examination of women's and men's *desire* for preparedness regarding qualification fit: Participants were asked to imagine that they were looking for a new job. Participants then indicated their desired qualification fit by selecting a number between 0% and 100% in response to the statement "You will apply for a job if you think you have X% of the skills and knowledge that the job requires" (Lee, 2018, p. 12). Across five studies, women strived for a higher preparedness than men, as women wanted to fulfill a higher percentage of the qualification criteria (79%) before they would apply for a job than men (72%). Further, the results showed that women held only themselves to this higher standard, as neither a gender difference occurred in the advised preparedness for a close female nor for a close male friend (Lee, 2018, Study 2).

The present research goes beyond the analysis of the percentage of qualification criteria participants desire to fulfill before applying for a job. We either manipulated qualification fit or assessed participants actual qualification fit for a concrete, realistic job advertisement for which participants indicated their application intention. We did this to reflect application processes more realistically and directly, as self-reported desired percentages are a rather simplistic and at the same time very abstract measure. It is unclear whether and how such abstract estimations of desire for preparedness are actually translated into behavioral intentions in application situations. We therefore tested whether we find a gender difference in application intention depending on manipulated or assessed qualification fit. Further, the gender difference in desire for preparedness found in Lee (2018), 72% versus 79%, was rather small compared to the 60% versus 100% hypothesis. The present studies therefore tested for significant

gender differences at different levels of qualification fit. We further tested several potential explanations for the hypothesized gender difference, which we outline next, to provide a starting point for future research and practice.

Why Women May Desire to Be More Prepared Than Men When Applying for a Job

Past research and discussions on the 60% and 100% statistic have already generated several explanations for a gender difference in application intention depending on qualification fit. At the core of the explanations why women may desire to be more prepared for jobs than men lie gendered socialization (Lawson et al., 2015; Tannen, 1990) as well as prescriptive gender stereotypes (Prentice & Carranza, 2002). We will now shortly review past research on gender differences in those explanatory variables.

Lower Self-Confidence

Oftentimes, the 60% and 100% statistic has been attributed to women having lower *self-confidence* than men (e.g., Hannon, 2014). However, in her Harvard Business Review article, Mohr (2014) reports results from a survey based on which she refutes confidence as an explanation for why women may want to be more prepared than men when applying for a job. In her survey, Mohr asked *why* exactly people did not apply for a job when they decided not to apply because they did not meet enough qualifications. A lack of confidence was the *least* common answer given (10% of women and 12% of men gave this answer). Further, gender differences in self-confidence seem to be highly dependent on context and existing results are therefore inconclusive regarding qualification fit (Matlin, 2004). However, in measures closely related to self-confidence, women have reported lower *self-esteem* (Lee, 2018), higher *perfectionism* (Lee, 2018), and stronger *impostor-feelings* (Markman, 2019) than men. These characteristics have consequently also been proposed as explanation for a gender difference in desire for preparedness (Lee, 2018; Markman, 2019). As they have not (sufficiently) been tested as explanations, they are also investigated in the present research.

Higher Risk Aversion

Research has shown that women are less willing to take risks than men (Dohmen et al., 2011; see Byrnes et al., 1999, for a meta-analysis) and this is due to women's expectation of less rewarding consequences of taking risks (Morgenroth et al., 2022). Regarding job applications, Lee (2018) showed that gender differences in "default" *risk perceptions* may pose an explanation for women wanting to be more prepared than men when applying for a job. When a job description did not contain any risk-related information, women wanted to be more prepared than men before they would apply (i.e., women indicated a higher percentage of the qualifications mentioned in the job description they desired to fulfill). In comparison, when a job was described as risky (i.e., a bad performance decreases chances for a future career), no gender difference in desire for preparedness emerged because men's desire for preparedness increased to that of women. When a job was described as not risky (i.e., a bad performance will not decrease chances for a future career), no gender difference in desire for preparedness emerged because women's desire for preparedness decreased to the level of men's desire for preparedness (Lee, 2018, Study 4). This result pattern allows the conclusion that women as a default perceive a higher risk in application situations (unless an explicit statement about low risk is present) eliciting a higher desire for preparedness. As an intervention to reduce the perceived risk Lee (2018, Study 5) examined the effects of using a *growth mindset* communication (i.e., the company values learning experiences) in a job advertisement. When the job advertisement did not include a growth mindset communication, the basic pattern was replicated as women reported a higher desire for preparedness than men. However, when the job advertisement conveyed a growth mindset, this gender difference disappeared because women's desire for preparedness was reduced (Lee 2018).

Related to women being more risk averse, women have further been found to behave less *competitive* than men (Saccardo et al., 2018). Gender differences in competition could

also be relevant for explaining gender differences in application situations due to their competitive nature and are therefore also investigated in the present research.

Higher Fear of Failing

Another set of explanatory variables that could explain why women may want to be more prepared for a job than men are fears. When Mohr (2014) asked *why* exactly people did not apply for a job when they decided not to apply because they did not meet enough qualifications, the largest descriptive difference in responses from women and men emerged on their *fear of failing*. Women (22%) indicated more fear of failing than men (13%). This is backed up by findings across different age groups that – in general – females have stronger fear of failing than males (Borgonovi & Han, 2021; Levy et al., 2012; Nelson et al., 2013). Further, Lee (2018) discusses that women might have a higher desire for preparedness than men as women have learned that they are held to a higher standard in application situations (and research shows that indeed they are, see e.g., Heilman, 2001; Gerdes & Garber, 1983). Therefore, women's *fear* that they are not able to *meet expectations* could explain their higher desire for preparedness. Other fears potentially explaining gender differences in application intention are *fear of negative evaluation/rejection* (Lee, 2018) and *fear of backlash* (Salwender et al., 2023). As these fears have not yet been tested with reference to qualification fit in application situations, we include them in the present research.

Higher Rule Abidance

Mohr (2014) found that women answered descriptively more often than men that they were following the guidelines when not applying for a job because they did not meet all qualification criteria. This is backed up by research showing that women – in general – report higher *rule abidance* than men in work contexts (Portillo & DeHart-Davis, 2009). Besides following the guidelines, a *waste of time and energy* was named as top reason to not apply when

not all qualification criteria are met (descriptively more often by men than women; Mohr, 2014).

Across the studies reported below, we investigate all these potential explanations. Notably, most discussions have focused on women so far. However, Lee (2018) noticed that women, men, or both could be the driver of differences. Men, for example, may underestimate the degree of preparedness necessary and be overly confident (Lee, 2018). It is also possible that they refrain from applying with a very high qualification fit because they perceive a lack of challenge and further qualification gains. We therefore analyze and discuss gender differences on different qualification fit levels.

Overview of the Present Research

In the following, we present seven studies investigating whether women only apply for a job when they are 100% qualified, whereas men already apply for a job when they are 60% qualified and the possible reasons for such a gender difference. Across Studies 1-6 different paradigms were used to realize conceptual replications examining gender differences in application intention depending on manipulated or assessed qualification fit. In Studies 1-3 an experimental approach was administered by using hypothetical application scenarios with a manipulated qualification fit of CVs and job advertisements. Participants were to imagine a CV as their own and to indicate how much they intend to apply for a job. Using a correlational study design, in Studies 4-6 we presented a job advertisement and asked the participants how much they intended to apply for the job. Afterwards, participants' qualification fit was assessed by having participants report which of the qualification criteria they themselves fulfilled. Across these six studies, we assessed several of the explanatory variables outlined above (e.g., fear of failing, rule abidance) as moderators or mediators. However, as the main goal of the first six studies was to examine whether the qualification fit – application intention relation is actually stronger for women than men, we focus on this main hypothesis and report

effects of the explanatory variables in the online supplement only. To foreshadow our results, we did not find strong evidence for a gender difference in application intention depending on qualification fit in those first six studies, neither when manipulating nor when assessing qualification fit. However, we consistently replicated the small gender difference in desire for preparedness reported by Lee (2018). Followingly, we examined women's and men's desire for preparedness in more detail in Study 7. Specifically, we competitively tested all explanatory variables described above as mediators to learn why women want to be more prepared than men before applying for a job. Finally, we integrate and discuss all findings in an overall discussion.

Studies 1-3

In Studies 1-3 we used hypothetical application scenarios with a manipulated qualification fit of CVs and job advertisements to test whether women only apply for a job when they are 100% qualified, whereas men already apply for a job when they are 60% qualified. As the procedure and design was very similar across Studies 1-3, we report a consolidated methods section to keep the manuscript concise (for details concerning the individual studies see online supplement). We report individual results sections for Studies 1-3 as qualification fit was manipulated between-subjects in Studies 1 and 3 and within-subjects in Study 2, requiring different sets of analyses. To keep results of the different studies comparable, analyses in the main manuscript are reported, for example, without pre-registered covariates which varied from study to study. Corresponding analyses, for example with covariates, are reported in the online supplement.

Method

Participants

In Study 1 $N = 289$ (76% female; 77% students in the field of social sciences), in Study 2 $N = 91$ (63% female; 47% students in the field of social sciences), and in Study 3 $N =$

265 (65% female; 33% students in the field of social sciences) university students participated. All three studies were conducted online. We excluded participants who identified as other gender from the analyses as their sample size was too small for gender-grouped statistical analyses.

Procedure and Design

Before starting the study participants had to agree to data protection and informed consent regulations. Next participants reported their demographics (including gender).

In all three studies, participants saw a CV (see online supplement for materials) and should imagine this was their own CV. We created different CVs for each study and aimed to create CVs our participants could widely identify with in terms of experience, skill levels, and grades. In addition to a CV, participants read one or more job advertisements (e.g., for a student assistant position) including a description of the company, the tasks, and qualifications required for the job (exemplary qualification “You have experience abroad”). Qualification fit was manipulated via a different number of qualifications listed in the job advertisement that were also present in the CV (e.g., year abroad after finishing school does meet the required qualification of having experience abroad vs. social year in Germany does not). After viewing the CV and a job advertisement, participants indicated their intention to apply for the position in the hypothetical scenario. A manipulation check measured the qualification fit perceived by the participants.

Study materials (CVs and job advertisements) and the specific experimental designs varied between the three studies to achieve conceptual replications. The design in Study 1 was chosen to test the basic hypothesis of the gender difference in application intention in a 60% versus 100% condition. It was therefore based on a 2 (participant gender: female vs. male) \times 2 (condition: 60% vs. 100% objective qualification fit) between-subjects design with random

assignment to the condition groups (i.e., participants only saw one job advertisement with either 60% or 100% qualification fit). Study 2 extended the focus on different amounts of qualification fit and uses a within-subjects design to get closer to real application processes in which different options may be available. The design in Study 2 was a 2 (participant gender: female vs. male) \times 6 (condition: 0%, 20%, 40%, 60%, 80% vs. 100% objective qualification fit) mixed design with participant gender as between-subjects and qualification fit as within-subjects factor (i.e., participants saw six job advertisements with differing qualification fit and indicated their intention to apply for each job advertisement). To maximize power to detect a gender effect in the critical 60% qualification fit condition, Study 3 was a 2 (participant gender: female vs. male) \times 1 (condition: 60% objective qualification fit) design with participant gender as between-subjects factor (i.e., all participants saw a job advertisement with 60% qualification fit and indicated their intention to apply).

In addition to the experimental test of gender differences in application intention depending on qualification fit, we asked about the participants' desire for preparedness in Studies 2 and 3 to see whether we replicate Lee's (2018) finding that women want to be more prepared for a job than men.

Measures

Participants responded to all measures on 7-point rating scales ranging from 1 = *do not agree at all* to 7 = *totally agree*, if not indicated otherwise.

Manipulation Check. In Study 1 participants ticked the criteria listed in the qualifications section in the job advertisement they fulfilled with "their" hypothetical CV resulting in scores from 0 to 10 criteria met with the hypothetical CV. In Study 2 participants estimated for one of the six job advertisements (randomly selected, to keep the experimental tasks short)

the fit of the job advertisement and their hypothetical CV [%]. In Study 3 participants estimated the percentage of qualification fit between the CV presented and the job advertisement [%].

Application Intention. In Studies 1 and 3 participants' application intention was measured with two items "I would apply to the advertised position with this CV.", and "If this were my CV, I would be very interested in applying for the advertised thesis." (adapted from Hentschel et al., 2018; Study 1: $r(287) = .69, p < .001$; Study 3: $r(263) = .56, p < .001$). Participants' scores on the two items were averaged. In Study 2 participants indicated their application intention on one item only "How likely is it that you will apply for this job with "your" CV?" (1 = *very unlikely* to 7 = *very likely*, adapted from Hentschel et al., 2018).

Desire for Preparedness. In Studies 2 and 3 we asked participants "Imagine you have just read an advertisement for a job you are interested in. How many % match do you need with the qualifications listed to apply?" [%].

Results Study 1

Manipulation Check

In the 60% qualification fit condition participants ticked on average 5.94 ($SD = 0.91$) criteria they fulfilled with their imagined CV, whereas in the 100% condition participants ticked 9.42 ($SD = 1.14$) criteria, $F(1, 285) = 621.67, p < .001, \eta^2 = .69$, indicating a successful manipulation of qualification fit. The main effect of participant gender, $F(1, 285) = 2.08, p = .151$ (female participants responded slightly higher perceived fit than male participants), and the Participant Gender \times Condition interaction were not significant, $F(1, 285) = 0.50, p = .479$.

Application Intention Depending on Participant Gender and Qualification Fit

The predicted Participant Gender \times Objective Qualification Fit interaction was not significant, $F(1, 285) = 0.01, p = .906$ (see Figure 1). Contrary to the hypothesis, female and male participants did not differ in their application intention in the 60% qualification fit condition, $M_{\text{diff}} = 0.33, SE = 0.25, p = 1.000$. Overall, in the 100% qualification fit condition, female and male participants' application intention was higher (main effect of condition: $F(1, 285) = 79.70, p < .001, \eta^2 = .22$), and again did not differ, $M_{\text{diff}} = -0.37, SE = 0.25, p = .775$. Across conditions, female participants indicated a higher application intention than male participants, $F(1, 285) = 4.11, p = .043, \eta^2 = .01$. To further test the null-effect of the interaction relevant for the main hypothesis, a Bayesian ANOVA was conducted (using the default priors in JASP software) comparing a null-model including only a random effect for the subjects and the main effects of participant gender and qualification fit condition to a model that additionally included the predicted interaction term. For the inclusion of the interaction term a Bayes Factor of $BF_{01} = 4.660$ was found. A Bayes Factor of this magnitude is conventionally described as substantial evidence for the null hypothesis (Wagenmakers et al., 2011). Thus, the Bayes Factor speaks against the presence of the predicted interaction.

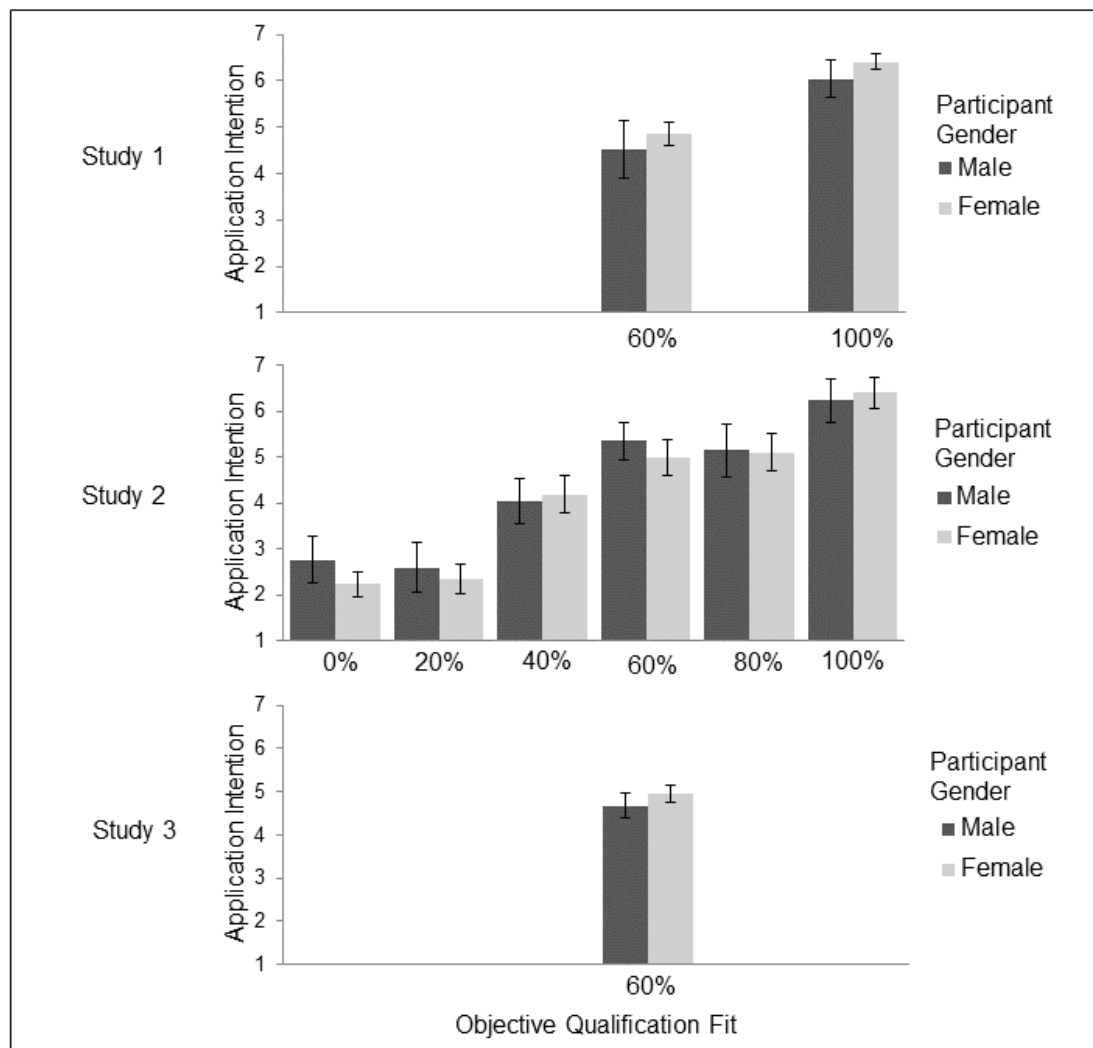
Results Study 2

Manipulation Check

Participants estimated the percentage of qualification fit between the CV presented and the job advertisement significantly different depending on qualification fit condition, $F(5, 79) = 26.86, p < .001, \eta^2 = .63$. Analyzing Helmert-contrasts showed that the manipulation check worked well (all $ps < .078$) except for the step from the 60% to the 80% qualification fit condition that was not perceived to differ ($p = .757$; see Table S1 in the online supplement for descriptives). The main effect of participant gender and the Participant Gender \times Objective Qualification Fit interaction were not significant, $F_s < 0.28, ps > .782$.

Figure 1

Application Intention Depending on Objective Qualification Fit and Participant Gender
(Studies 1-3)



Note. Error bars represent 95% confidence intervals.

Application Intention Depending on Participant Gender and Qualification Fit

The predicted Participant Gender \times Objective Qualification Fit interaction was not significant, $F(4.26, 379.49) = 1.12, p = .348$ (with Greenhouse-Geisser correction to adjust for

violations of sphericity; see Figure 1). Contrary to the hypothesis, female and male participants did not differ in their application intention at any percentage of qualification fit. For 20% to 100% qualification fit the level of significance of the participant gender contrast ranged from $p = .228$ to $.906$. For 0% qualification fit a marginal trend emerged, $p = .056$, with female participants showing a somewhat lower application intention compared to male participants. A Bayesian repeated measures ANOVA testing the interaction term supported the non-significant interaction effect, $BF_{01} = 17.750$, providing strong evidence for the null hypothesis. The main effect of condition was significant, $F(4.26, 379.49) = 133.97$, $p < .001$, $\eta^2 = .60$ (linear contrast $p < .001$) with higher application intention the higher qualification fit. Across conditions, no significant gender difference in application intentions emerged, $F(1, 89) = 0.57$, $p = .452$.

Desire for Preparedness

Female participants ($M = 77.89\%$, $SD = 9.54$) reported a higher desire for preparedness than male participants ($M = 69.59\%$, $SD = 13.75$), $t(89) = 3.40$, $p < .001$, $d = .74$.

Results Study 3

Manipulation Check

Participants estimated the percentage of the manipulated 60% qualification fit between the CV presented and the job advertisement at 66% ($SD = 16.32$). Female participants ($M = 68\%$, $SD = 14.54$) rated the fit higher than male participants ($M = 63\%$, $SD = 18.91$; $t(149.41) = 2.12$, $p = .036$, $d = .30$). We therefore controlled for this rating in the analysis of participants' application intention.

Application Intention Depending on Participant Gender at 60% Fit

Controlling for the estimated percentage of qualification fit in step 1 of a hierarchical regression analysis, application intention did not differ between female ($M = 4.96$, $SD = 1.38$)

and male participants ($M = 4.67$, $SD = 1.40$), $\beta = -.03$, $t(262) = -0.51$, $p = .608$. Dropping the control variable and running a t -test did not change the result, $t(263) = 1.59$, $p = .114$. A Bayesian independent samples t -test with male > female participants' application intention as alternative hypothesis supported this result, $BF_{01} = 17.434$, providing strong evidence for the null hypothesis.

Desire for Preparedness

Female participants ($M = 73.18\%$, $SD = 14.86$) indicated a higher desire for preparedness than male participants ($M = 67.86\%$, $SD = 18.19$), $t(263) = 2.56$, $p = .011$, $d = .33$.

Discussion

In Studies 1-3 we did not find evidence for the proposed gender difference in application intention at 60% qualification fit (not even when integrating the data from the three studies to maximize power and testing for a gender difference in the 60% qualification fit condition that was present in all three studies, see online supplement). At the same time, Studies 2 and 3 replicated a gender difference in desire for preparedness at a much narrower margin. Comparable to the findings of Lee (2018; US American adult samples) the present studies confirmed that the desire for preparedness was significantly higher for women than for men (73-78% for women vs. 67-70% for men). Conceivably, this gender difference in desire for preparedness could translate into a gender difference in application intentions depending on qualification fit. However, this difference can probably only be found at a much smaller margin than predicted in the anecdotal 60% versus 100% claim (Sandberg, 2015) and may be restricted to a narrow, specific level of qualification fit (67-78%).

Besides a much smaller actual gender difference than predicted, limitations of the design used in Studies 1-3 could explain the non-significant results of Studies 1-3. All three studies were based on hypothetical scenarios as participants had to imagine a presented CV as

their own. This can be criticized of lacking self-relevance of the situation or lacking identification with the CV (which could also be reflected in participants' low overall interest in the advertised positions). Participants may even have felt to be in the position to advise the person whose CV they read on the application decision, a situation for which Lee (2018) has already shown the absence of gender effects: Women and men gave highly similar recommendations regarding desire for preparedness to friends. We therefore developed a new method for the next set of studies to overcome these limitations.

Studies 4-6

The new method applied in Studies 4-6 comprised presenting a job advertisement and asking the participants how much they personally intended to apply. Afterwards, the participants checked the qualification criteria listed in the job advertisement they themselves fulfilled. By using this new study paradigm, no CV had to be imagined, reducing the hypotheticality of the task. All statements concerned the participants themselves. To increase the conciseness of the presentation and as the basic design was the same across Studies 4-6 (for details concerning the individual studies see online supplement), we present the results in an integrative data analysis of the merged data sets of Studies 4-6 with study number as dummy variable.

Method

Participants

The integrative data analysis comprised $N = 807$ participants (70% female). The samples in Studies 4 and 5 consisted of students, whereas Study 6 included 62% employees.

Procedure and Design

In all three studies we asked the participants to imagine they are looking for a job. Then participants saw a job advertisement (Studies 4 and 5 for two different newly created

student assistant positions, Study 6 for a consultant position). Next, participants indicated their application intention. Then participants indicated which of the ten qualification criteria listed in the job advertisement they personally fulfilled. Additional questions (see online supplement for details) followed.

Measures

Application Intention. For the integrative data analysis, we used the item “I would apply to the advertised position with this CV.”, as this was present in all three studies.

Subjective Qualification Fit. For each of the ten qualification criteria listed in the respective job advertisement (e.g., studying at a university), the participants indicated whether they fulfilled the criterion (yes/no). The resulting subjective qualification fit score ranged from 0% to 100%.

Desire for Preparedness. We asked participants “Imagine you have just read an advertisement for a job you are interested in. How many % match do you need with the qualifications listed to apply?” [%]. This measure was only included in Studies 4 and 5.

Results

Application Intention Depending on Participant Gender and Qualification Fit

We tested our hypothesis that the qualification fit – application intention relation is stronger for women than for men with a hierarchical regression analysis. Application intention was the criterion and participant gender, qualification fit, and their interactions were the predictors in the analysis. Interest in the specific job advertisement served as control variable in step 1 of the analysis, as each study had used a different job advertisement (the results were virtually identical with and without interest as control variable). Additionally, study number and the respective interactions were added to test the consistency of the results across studies.

The results are reported in Table 1 and Figure 2. Higher qualification fit predicted higher application intentions. Female participants indicated a higher application intention than male participants. The Participant Gender \times Qualification Fit interaction was significant. For female participants, the correlation between qualification fit and application intention was significantly positive, $r(564) = .30, p < .001$. For male participants the correlation was also positive, but not significant, $r(239) = .08, p = .223$. The two correlations differed significantly from each other, Fisher's z -test: $z = 2.97, p = .002$. Investigating simple slopes revealed that for 0%-60% qualification fit the gender difference in application intention was not significant ($ps > .139$), whereas for 70%-100% qualification fit female participants' application intention was higher compared to male participants' application intention ($ps < .0057$). In addition, a Bayesian linear regression was conducted comparing a null-model including participant gender, qualification fit, and interest to a model that additionally included the Participant Gender \times Qualification Fit interaction term (step 3 from Table 1). For the inclusion of the interaction term a Bayes Factor of $BF_{01} = 1.04$ was found. A Bayes Factor of this magnitude is conventionally described as anecdotal evidence for the null hypothesis (Wagenmakers et al., 2011). Thus, the Bayes Factor speaks tentatively against the presence of the predicted interaction. That the gender effect seems not very robust is backed up by the individual analyses for the three studies. Although the integrated analysis did not reveal a significant three-way interaction between gender, qualification fit, and study number, the crucial Participant Gender \times Qualification Fit interaction was not consistently found in the individual studies (β_s ranging from .05 to .83, p_s from .024 to .367, see online supplement). Overall, the pattern is in the predicted direction, but if at all the effect is very small, primarily driven by women's higher application intention (or phrased the other way around by men's lower application intention) at high levels of qualification fit, and more data is needed to come to a clear conclusion.

Table 1

Hierarchical Regression Analysis Predicting Application Intentions (Integrative Data Analysis Studies 4-6)

Variable	β	t	p	Adj. R^2	ΔR^2
Step 1				.13	
Interest	0.36	10.95	<.001		
Step 2				.16	.03
Participant Gender	0.08	2.53	.019		
Subjective Qualification Fit	0.17	5.19	<.001		
Step 3				.17	.01
Participant Gender \times Subjective Qualification Fit	0.08	2.08	.038		
Step 4				.20	.03
D2	-0.12	-2.79	.005		
D3	0.03	0.68	.494		
D2 \times Subjective Qualification Fit	0.04	0.70	.482		
D3 \times Subjective Qualification Fit	-0.07	-0.96	.336		
D2 \times Participant Gender	0.10	1.87	.062		
D3 \times Participant Gender	0.05	1.00	.318		
D2 \times Participant Gender \times Subjective Qualification Fit	-0.07	-1.20	.231		
D3 \times Participant Gender \times Subjective Qualification Fit	-0.08	-1.07	.286		

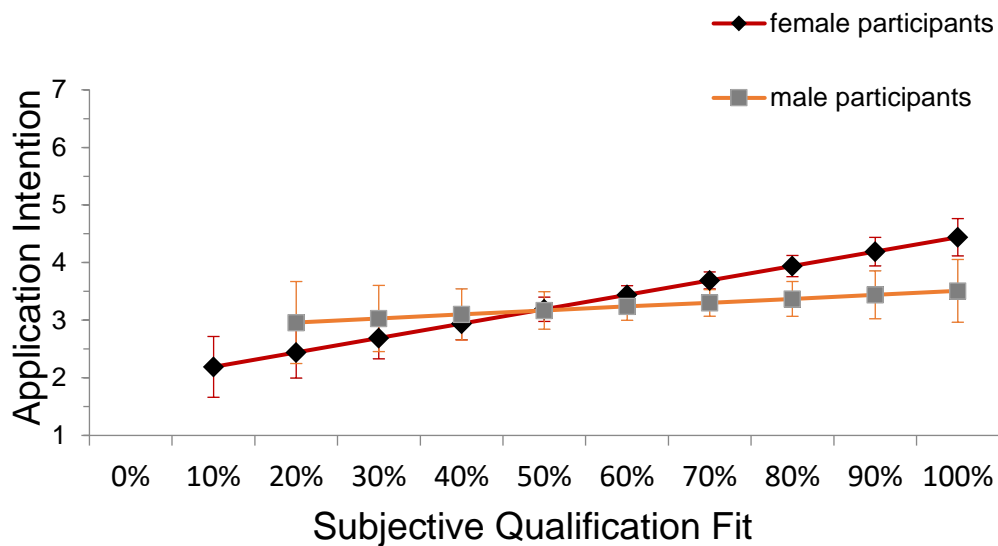
Note. Female participants = +0.5, male participants = -.05; study number: reference group is Study 4; D2 compares Study 5 to Study 4, D3 compares Study 6 to Study 4

Desire for Preparedness

Female participants ($M = 75.67\%$, $SD = 11.66$) indicated a higher desire for preparedness than male participants ($M = 73.00\%$, $SD = 11.40$), $F(1, 525) = 6.06$, $p = .014$, $\eta^2 = .011$ (study number was included as dummy in this analysis).

Figure 2

Application Intention Depending on Subjective Qualification Fit and Participant Gender (Integrative Data Analysis)



Note. Error bars represent 95% confidence intervals.

Discussion

In this integrative data analysis summarizing Studies 4-6 we found a small, but not robust gender difference in application intention depending on qualification fit. In contrast, the gender difference in desire for preparedness was robustly replicated. Before discussing these results, we present a small-scale meta-analysis on the gender difference in desire for preparedness.

Small-Scale Meta-Analysis: Women's and Men's Desire for Preparedness

We conducted a small-scale meta-analysis summarizing the gender difference in desire for preparedness across our own studies and Lee's (2018) studies ($k = 10$; from Lee we only used the control groups if an intervention was tested). The results were highly consistent

between studies. Women's desire for preparedness ranged between 73-80% and men's between 67-74%. Overall, women's desire for preparedness was significantly higher than men's desire for preparedness, $d = 0.39$, 95% CI [0.28, 0.51].

Interim Discussion

We conducted six studies to examine whether the qualification fit – application intention relation is stronger for women than for men. In Studies 1-3 we used an experimental approach presenting hypothetical CVs to the participants. In those studies, we did not find evidence for a gender difference in the strength of the qualification fit – application intention relation. However, as discussed above, these studies may have come with their limitations (e.g., hypotheticality, analogy to “giving advice to a friend”). Therefore, in Studies 4-6 we used a correlational approach eliminating those limitations by asking participants for their personal application intention and assessing participants' qualification fit. In these latter studies, we found a small indication for a stronger qualification fit – application intention relation for women than for men which was especially driven by women's higher application intention compared to men's at high levels of qualification fit. However, the results of a Bayesian analysis speak against the robustness of this gender difference and leave the conclusion that more data is needed for a definite evaluation.

For all six studies reported so far, we tried to develop application scenarios that are realistic and tailored for the participants of the respective study. In addition, we created job advertisements using gender-neutral wordings and layouts to make sure that women and men were per se equally attracted to the respective jobs. Yet, despite these efforts, interest ratings were overall low and sometimes differed between females and males. Therefore, to examine the qualification fit – application intention relationship without confounds of higher or lower interest in the advertised job, we controlled for participants' interest in the analysis, arriving at the same results (see online supplement). Future studies should invest in assessing qualification

fit and application intention in real applicants to eliminate hypotheticality altogether and maximize interest for a further critical test of the 60% versus 100% claim.

Summarizing Studies 1-6, tentative evidence hints towards qualification fit potentially being more important for women than men in their thoughts and feelings in application processes. Strongest evidence for such a conclusion is provided by a robust gender difference in desire for preparedness which consistently emerged for the general desire for preparedness measure that was independent of the specific study designs and job advertisements. Women indicated a higher desire for preparedness in application situations than men. This speaks for a higher psychological hurdle present for women compared to men as women have to overcome their higher desire for preparedness to make the decision whether to apply for a job. To learn more about the gender difference in desire for preparedness, we next present data from a survey study where we integrated the explanations discussed in the introduction (e.g., self-confidence, risk aversion) to inform further research and practice about psychological, gender-specific hurdles present in application situations.

Study 7

In this seventh study we asked participants from diverse backgrounds about their desire for preparedness by asking at what percentage of qualification fit participants would apply for a job. In addition, we included the potential explanations outlined in the introduction and competitively tested which of these explanations best predicts participants' desire for preparedness. Additional mediation analyses can inform further research as well as practice on relevant, gender-specific psychological hurdles present in application situations.

Method

Participants

In this online study $N = 196$ participants (51% female) participated. Participants were recruited via Prolific. On average participants had 8.58 ($SD = 8.78$) years' work experience, 47% percent of the sample was employed, 9% self-employed, and 18% looking for a job, 13% were in a leadership position.

Procedure and Design

Participants agreed to participate in a survey on thoughts and feelings in application situations. The survey started with filler questions (see online supplement for a full description of the survey) to introduce the topic. Afterwards, we assessed the potential explanations for women's higher desire for preparedness, including self-confidence, risk aversion, fear of failing, and rule abidance and their correlates as described in the theoretical section with reference to application situations. Further, we asked the participants at what percentage of fit they would apply for a job they were interested in (i.e., desire for preparedness). In the last section of the survey questions on demographics followed.

Measures

Desire for Preparedness. Participants answered the same item "Imagine you have just read an advertisement for a job you are interested in. How many % match do you need with the qualifications listed to apply?" [%] as used previously.

Explanatory Variables.

Self-Confidence. Participants indicated their self-confidence when applying for a job with three items, for example "I am convinced of the success of my applications." (1 = *does not apply to me at all* to 7 = *totally applies to me*; $\alpha = .73$; adapted from Dunn et al., 2021; average score). In addition to self-confidence, we assessed the highly related construct self-efficacy. Participants indicated their *self-efficacy* when applying for a job with three items, for

example “I can rely on my skills in the application process.” (1 = *does not apply to me at all* to 7 = *totally applies to me*; adapted from Beierlein et al., 2014; $\alpha = .81$).

Self-Esteem. Participants indicated their general self-esteem with four items, for example “I am happy to be the way I am.” ($\alpha = .86$; Hormuth & Lalli, 1988; average score).

Perfectionism. Participants indicated their perfectionism when applying for a job with three items, for example “I strive in the application process to appear perfect.” (1 = *does not apply to me at all* to 7 = *totally applies to me*; $\alpha = .52$; adapted from Hewitt et al., 2003; average score).

Impostor-Feelings. Impostor-feelings when applying for a job were assessed with three items, for example “I’m afraid that in the application process it will come out that I’m not as capable as expected.” (1 = *does not apply to me at all* to 7 = *totally applies to me*; $\alpha = .81$; adapted from SOEP-IS Group, 2017; average score).

Risk Aversion. Participants’ risk perceptions in application situations were measured with two items, for example “I like to take risks in the application process.” (reversed item; 1 = *does not apply to me at all* to 7 = *totally applies to me*; $r(194) = .27, p < .001$; adapted from Mandrik & Bao, 2005; average score).

Growth Mindset. Participants indicated their job-related growth mindset by responding to two items, for example, “I can learn new skills that I need for a job.” (1 = *does not apply to me at all* to 7 = *totally applies to me*; $r(194) = .31, p < .001$; self-created scale; average score).

Competition. Participants responded to five items on their perception of competition when thinking about application situations, for example “Competition makes people perform better.” ($\alpha = .73$; Kesebir et al., 2019; average score).

Fear of Failing. Participants responded to two items on their fear of failing in application situations, for example “Would you be preoccupied with the thought “The recruiter could think I’m incompetent”?” (1 = *not at all* to 7 = *very much*; $r(194) = .75, p < .001$; self-created scale; average score).

Fear of Not Being Able to Meet Expectations. Tapping into whether women think they are held to a higher standard in application situations without being too explicit about gender (as this could elicit reactance or desirability effects), we asked the participants how much they agree with the statement “I only apply when I have a high fit because I fear that I will be doubted if I don’t meet all the criteria.” (self-developed).

Fear of Negative Evaluation/Rejection. Participants indicated their fear of negative evaluation/rejection in application situations on two items, for example “Would you be nervous because of the possibility of receiving a rejection?” (1 = *not at all* to 7 = *very much*; $r(194) = .67, p < .001$; self-created scale; average score).

Fear of Backlash. Participants indicated their fear of backlash on four items, for example “Would you be concerned that the recruiter perceives you as too confident?” (1 = *not at all* to 7 = *very much*; $\alpha = .65$; see Study 5; adapted from Salvender et al., 2023; Moss-Racusin & Rudman, 2010; average score).

Rule Abidance. Participants responded to three questions on their perception of rules concerning qualification fit, for example “I think the qualifications listed are actually all expected in a job posting.” ($\alpha = .73$; self-created scale; average score).

Waste of Time and Energy. We asked participants three questions on their perception of time investment concerning qualification fit, for example “Applying for a job where I don’t meet all the criteria in the job posting would be a waste of energy and time.” ($\alpha = .79$; self-created scale; average score).

Additional Variables. For exploratory reasons, we asked participants some additional questions on their perception of qualification fit: “For the success of an application it is more important to sell yourself well compared to how many requirements you meet.” (*present oneself*), “You should apply for a job where you do not already meet all the requirements to be able to develop further.” (*room for development*), “If my skills match the job advertisement, I have a high chance of success.” (*chances for success*; self-developed).

Participants’ *success motivation* in application situations was assessed by asking to rate “When I apply for a job, I think about how nice a success would be for me.”, *failure motivation* by “In an application process, my biggest goal is to avoid rejection under any circumstances.”, and *success attributions* by two items “If I am successful in the application process, I owe it to my skills, my *efforts*, or my good preparation.” and “If I’m successful in the application process, I’m *lucky*.” (1 = *does not apply to me at all* to 7 = *totally applies to me*; self-created).

We assessed participants’ goal orientation when applying for a job with one item for *ability orientation* “When I apply for a new job, I want a position that allows me to always prove my abilities.”, one item for *outcome orientation* “... that allows me to always get the best possible evaluation of my performance.”, and one item for *learning orientation* “...that allows me to always learn and improve my skills.” (Lee, 2018).

A short 10-item version of the Big-5 inventory measured participants’ *extraversion*, *agreeableness*, *conscientiousness*, *openness*, and *neuroticism* (Rammstedt et al., 2014).

Results

Desire for Preparedness

Replicating findings of the previous studies, female participants ($M = 74.07\%$, $SD = 10.78$) indicated a higher desire for preparedness than male participants ($M = 70.84\%$, $SD = 11.53$), $t(194) = 2.02$, $p = .044$, $d = .29$.

Predictors of Participants' Desire for Preparedness

We ran a random forest model using the randomForest package in R (Breiman et al., 2022) to competitively test which variables of the many explanatory variables we assessed are most predictive of participants' desire for preparedness. Variables were z-standardized before running the machine-learning algorithm. Results are depicted in Figure 3. The most important predictor of participants' desire for preparedness was fearing of not being able to meet expectations when not fulfilling all qualification criteria. All variables with a %IncMSE larger than 1 (dots to the right of the line) are deemed relevant for participants desire for preparedness (negative values denote noise). Results for the individual predictors (gender differences in the predictors as well as prediction of desire for preparedness) as well as a table presenting the correlations between all variables assessed are reported in the online supplement.

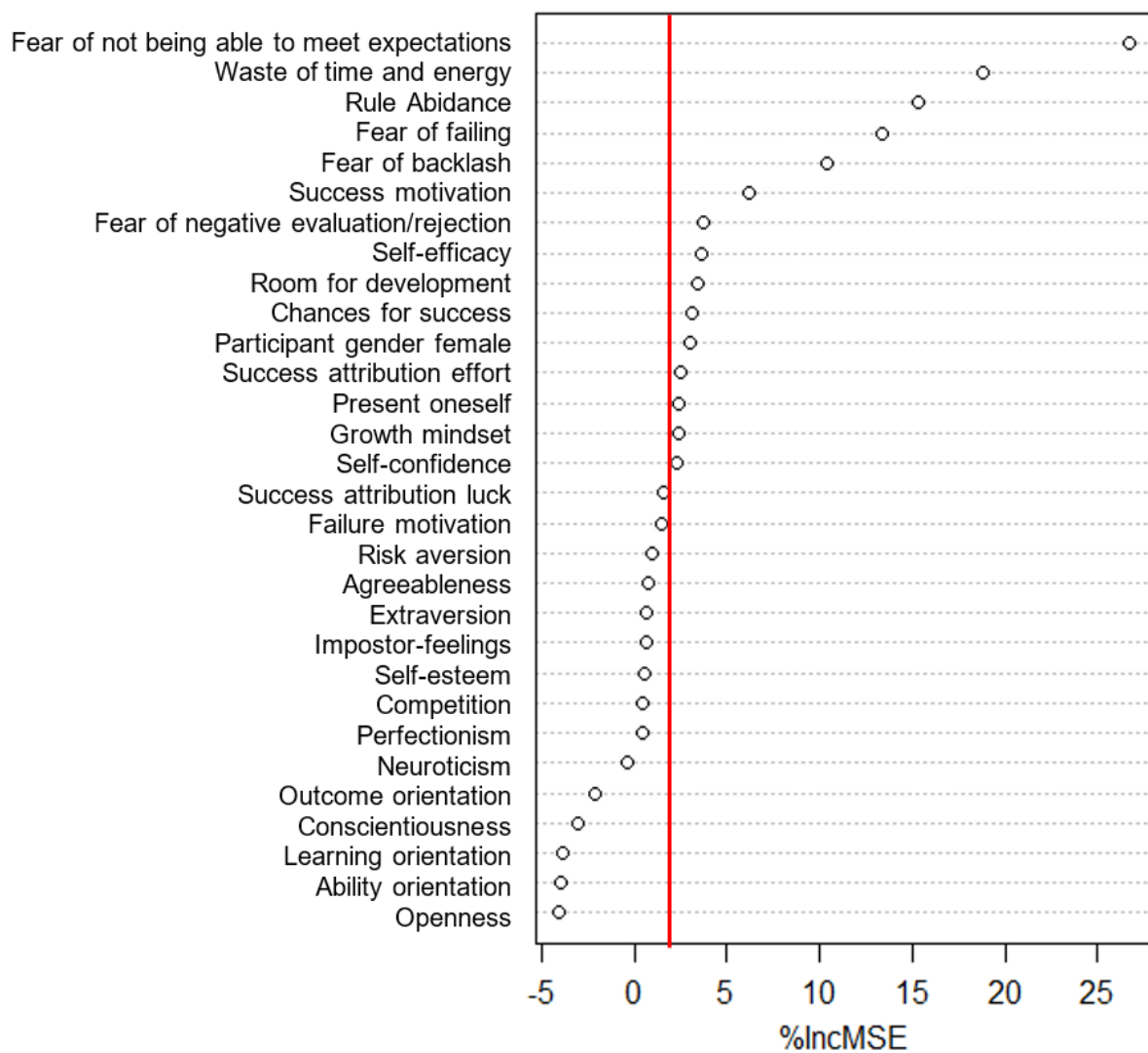
Mediation Analysis

For all the potential explanatory variables reported above we conducted a multiple mediation analysis (mma package in R, Yu & Li, 2017). Specifically, we ran a multiple mediator model with participant sex as predictor, desire for preparedness as outcome and 29 potential mediators. In the first step, mediators were identified by selecting those variables that are correlated with the predictor as well as the outcome. In the second step indirect effects were calculated (generalized linear model). Replicating the t -test presented before, overall, women were more likely to indicate a higher desire for preparedness than men. In the third step, statistical inferences were drawn using bootstrapping procedures. Using the quantile confidence interval, the results indicated at a significance level of 0.2 that fearing not being able to meet expectations, that is “fearing that doubts are cast on oneself when not fulfilling all criteria”,

had a significant indirect effect on the gender difference in desire for preparedness, which explained about 61% of the gender difference (average mediation effect from bootstrap samples: 0.61 [0.32; 2.07]). To a much smaller degree fear of backlash (7%) and impostor-feelings (5%) were significant mediators. No other mediators were significant.

Figure 3

Predictors of Desire for Preparedness (Study 7, Random Forest Model)



Discussion

Using a comprehensive set of survey questions, we replicated the gender difference in desire for preparedness. We tested multiple explanations and found the fear of not being able to meet expectations was especially relevant. Next, we discuss these results together with the previous studies.

General Discussion

We addressed the research gap on the question whether women only apply for a job if they are 100% qualified, whereas men already apply when they are 60% qualified. In six studies with different paradigms and methods we did not find convincing evidence for the 60% and 100% hypothesis. In Studies 1-3 no gender difference emerged in participants' application intentions at any level of manipulated qualification fit. In Studies 4-6 we found a stronger positive qualification fit (assessed) – application intention relationship for women than for men. Yet, this result did not emerge consistently, and the results of Bayesian analyses spoke against this Participant Gender \times Qualification Fit interaction, indicating that more data are needed to come to an informed conclusion. Overall, the results of the first six studies cast doubt on whether the 60% and 100% hypothesis, which has not yet been empirically proven, holds at least as a general gender difference. Beyond that, gender differences may be in much narrower ranges, as the data on women's and men's desire for preparedness suggest. The ranges in the preparedness data were between 73-80% of qualifications women desired to fulfill before applying for a job and between 67-74 % for men. In Study 7 we tested 29 potential mediators explaining this gender difference in desire for preparedness. Female participants consistently indicated a higher desire for preparedness than male participants and fearing not being able to meet expectations—and to a lesser extent the fear of backlash and having impostor-feelings—could explain this gender difference.

Theoretical Contributions

The results of Studies 1-6 raise the question: Does a gender difference at 60% qualification fit exist at all? Further, the results from Studies 4-6 suggest that it may be more worthwhile to focus on gender differences at high levels of qualification fit as gender differences in application intention in our data mainly emerged at high levels of qualification fit. This could be an important focus for future research, especially for an examination of mediators. While past narratives focused on women and deficits (e.g., lacking self-confidence; in the application area in general a large amount of past research has focused on female applicants, from both actor and observer perspective, e.g., Moss-Racusin & Rudman, 2010; Salwender et al., 2023; Williams & Tiedens, 2016), a gender difference at high levels of qualification fit might be better explained by men striving more (and women striving less) for challenges and opportunities for development and therefore intending to apply less (or more) when their fit is very high.

Additional mechanisms may play a role as well, such as men over-estimating their overall fit (Nicks et al., 2022). In this regard, gender differences may appear in consecutive steps in the application process and may add up to final differences in application intention and success (see also the discussion of small, cumulative gender effects in Szillis & Stahlberg, 2007). Female and male applicants may for example differ in their subjective decision how much knowledge or skill they need to fulfill a qualification criterion as well as in their overall fit perception. At first glance, our Studies 4-6 (see online supplement) seem to contradict this possibility: female and male applicants similarly estimated the degree to which they fulfill a single qualification criterion depending on their self-perception of performance (e.g., whether an applicant says they fulfill the criterion speaking English fluently or not depends on their self-rated language skills but not on gender). However, these findings rely on self-reports (and men's self-perception of skill may be inflated or women's may be underestimated) and should be complemented with other-reports as well as more objective measures (e.g., English test). In

addition, qualification fit does not seem to play a major role for men with regards to their application intention (see non-significant correlation in Studies 4-6). It would be informative for research and practice to investigate which predictors are relevant for men's application intentions.

Our findings of no gender difference in application intentions depending on qualification fit in experimental and correlational designs (Studies 1-6) and at the same time a robust gender difference in desire for preparedness (small-scale meta-analysis) may appear contradictory at the first glance. One potential explanation could be that for women higher psychological hurdles are present than for men when deciding whether to apply for a job or not, but women are willing and capable to go the extra mile and consequently no gender difference in application intentions based on qualification fit emerges. This might be part of a larger phenomenon as, for example, Salwender et al. (2023) found very small differences in women's and men's intention to seek power position and at the same time large gender differences for psychological hurdles such as fear of backlash.

Despite the psychological hurdle explanation other explanations are conceivable explaining at first glance contradictory findings, for example context. We sought to test the 100% versus 60% claim with university students in an as neutral context as possible to test its' validity in a neutral, common, and generalizable setting. For this purpose, we used gender-inclusive language and neutral adjectives in the job advertisements and gender balanced jobs (see detailed methods sections of Studies 1-6 in the online supplement) and did not find strong support for the claim that women only apply when they are 100% qualified, whereas men already apply when they are 60% qualified using correlational and experimental designs. However, context may be an important moderator in this regard. The neutral context as well as the student assistant positions targeted in the reported studies (that were chosen to achieve

relevance and closeness to reality for the student samples) contrast with the career advancement/leadership context for which the statistic has been proposed. The proposed effect may only be found in a high-stakes, male-dominated area, as such a context could induce lack-of-fit mechanisms such as fears of not being able to meet expectations (Heilman, 1983). Therefore, future research should investigate context as moderator.

Further, even though much can be learned from the studies presented, several important questions remain unanswered. It would be very interesting to learn which application strategy (e.g., applying selectively, which may be a wise investment of applicants' time, but may entail the risk of lost opportunities) is more advantageous and whether female applicants' fears to fall short of expectations are valid, even when women are equally qualified as men. In general, applicants with relatively high qualifications are perceived as more hireable than applicants with relatively low qualifications (Hardy et al., 2022), showing that taking qualification fit into account when making application decisions is in general a good strategy. An analysis of a recruiting company, however, finds no increase in chances of being invited to an interview from 50% qualification fit to 100% qualification fit, for women even from 40% onwards, and therefore recommend applying once a match of 50% is given (Jaja, 2018). First evidence hints even towards drawbacks of very high qualification fits. In one study, for example, being highly capable led to less favorable impressions by hiring managers compared to being adequately capable as, among other things, highly capable candidates were perceived to have more outside job options and potentially quit sooner (Galperin et al., 2020). Men might be more aware of such drawbacks and/or integrate this awareness more strongly into their application strategy.

These findings from prior research would speak for a recommendation to apply at lower levels of fit than 100%, but this advice might not work for women: Research shows that

female applicants' larger fears of not being able to meet expectations compared to male applicants' fears (identified as explanation for a gender difference in desire for preparedness in Study 7) seem valid. The same performance has been found to be perceived differently depending on the gender of the actor in the direction that women's performance was devalued (Heilman, 2001). This is even more intriguing as past research has found that in hiring decisions for women competence plays a central role, whereas for evaluations of men potential plays a larger role (Gerdes & Garber, 1983; Player et al., 2019). Further, Niessen-Ruenzi and Zimmerer (2023) showed that especially when the evaluation of leadership potential of candidates proved difficult (e.g., outside hires), signaling skills was more important for female than male candidates.

Limitations and Future Directions

Rejecting the 60% and 100% hypothesis altogether based on the data presented might be premature because of several shortcomings in our studies. Future research could improve the research design by incorporating real application situations for actual job seekers, use job advertisements from high-stakes, male-dominated areas, or use company application data instead of hypothetical situations (see also interim discussion on low interest in the respective jobs in our studies). Further, with a dichotomous yes/no decision that has to be taken in real life situations gender differences may be aggravated. It therefore seems especially relevant to learn more about the discrepancy of women indicating overall a higher application intention than men in our studies, whereas Tockey and Ignatova (n.d.) reported that women applied to fewer jobs than men on the actual job market. This discrepancy could have arisen due to response tendencies, social desirability, different types of jobs examined, etcetera. Future research should investigate with dichotomous application decisions for how many jobs women and men apply and whether qualification fit can predict which jobs they apply to.

Further, future research should investigate whether the null findings in experimental and correlational designs measuring application intentions in Studies 1-6 replicate using different methods. For example, in a longitudinal study participants could first indicate for a large pool of qualification criteria which they fulfill. Then participants could be randomly assigned to different fit conditions and the job advertisements could be created based on the qualification criteria information provided by the participants. Also, in cooperation with organizations, applications could be rated regarding their qualification fit and this could be regressed on applicants' gender. For internal hires, available employee data could be used for a more objective assessment of qualification fit (e.g., age as proxy for experience).

A clear limitation of Study 7 was the explorative nature of including multiple mediators, often with self-created, translated items adapted for the application context. For some scales problematic scale metrics appeared (e.g., low reliability for perfectionism). Here, methodical improvements should be aspired for future research. Besides methodological improvements, a close look at the content scope of the items may be worthwhile. For impostor-feelings, for example, we only asked about impostor-feelings in application situations (i.e., others might notice in the application process that I am not as capable as expected). Perhaps impostor-feelings when thinking about the future job itself (i.e., others might notice after I am hired that I am not capable of doing the job) are more relevant and should be assessed. In this regard, qualitative interviews could be informative to learn about applicants' thoughts and relevant psychological hurdles.

Practical Implications

Based on our research findings, authors of online magazine and blog articles should refrain from citing anecdotal findings or mere hypotheses on male and female application behavior as the matter seems more complex (e.g., psychological hurdles) and does not seem evidence-based at the moment. Organizations are advised to check for gender imbalances in their

applicant pool as many factors are known that prevent disproportionately more female than male applicants from applying for jobs. Especially gender imbalances in psychological hurdles in application processes (e.g., fearing not being able to meet expectations) should be considered when designing application processes. Further, hiring biases can entail negative monetary effects for organizations (utility loss; Hardy et al., 2022). This calls for thorough investigations of biases in recruiting processes leading to different standards and outcomes for female and male applicants and ways to reduce those. Studies that research recruiters' perceptions of men and women who apply with different amounts of qualification fit will be informative in this regard. Such studies can complement previous work to develop interventions that have been found to be highly effective for both female and male applicants, especially when teaching job search skills, improving self-presentation, boosting self-efficacy, and encouraging proactivity (Liu et al., 2014). These aspects should be considered by providers of such interventions and by individuals when selecting those.

Conclusion

We tested the popular claim that women only apply for jobs when they are 100% qualified, whereas men already apply when they are 60% qualified. Across three experimental and three correlational studies we did not find robust gender differences when measuring application intentions depending on qualification fit. However, we consistently found higher psychological hurdles present for women, specifically a higher desire for preparedness based on the fear of not being able to meet expectations, present in application situations.

References

- Barbulescu, R., & Bidwell, M. (2013). Do women choose different jobs from men? Mechanisms of application segregation in the market for managerial workers. *Organization Science*, 24(3), 737–756. <https://doi.org/10.1287/orsc.1120.0757>
- Böttcher, G. (2017, September 27). Teures Recruiting im Vertrieb [Expensive recruiting in sales]. *Springer Professional*. <https://www.springerprofessional.de/recruiting/marketingstrategie/teures-recruiting-im-vertrieb/15080096>
- Borgonovi, F., & Han, S. W. (2021). Gender disparities in fear of failure among 15-year-old students: The role of gender inequality, the organisation of schooling and economic conditions. *Journal of Adolescence*, 86(1), 28–39. <https://doi.org/10.1016/j.adolescence.2020.11.009>
- Born, M. P., & Taris, T. W. (2010) The Impact of the wording of employment advertisements on students' inclination to apply for a job. *The Journal of Social Psychology*, 150(5), 485–502. <https://doi.org/10.1080/00224540903365422>
- Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2022). *Package ‘randomForest’*. <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>
- Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3), 367–383. <https://doi.org/10.1037/0033-2909.125.3.367>
- Carless, S. A. (2005). Person-job fit versus person-organization fit as predictors of organizational attraction and job acceptance intentions: A longitudinal study. *Journal of Occupational and Organizational Psychology*, 78(3), 411–429. <https://doi.org/10.1348/096317905X25995>

- Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., & Jones, D. A. (2005). Applicant attraction to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology*, 90(5), 928–944. <https://doi.org/10.1037/0021-9010.90.5.928>
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European economic association*, 9(3), 522–550. <https://doi.org/10.1111/j.1542-4774.2011.01015.x>
- Dunn, E. W., Chen, L., Proulx, J. D. E., Ehrlinger, J., & Savalei, V. (2021). Can researchers' personal characteristics shape their statistical inferences? *Personality and Social Psychology Bulletin*, 47(6), 969–984. <https://doi.org/10.1177/0146167220950522>
- Fluchtmann, J., Glenney, A., Harmon, N. A., & Maibom, J. (2022). The gender application gap: Do men and women apply for the same jobs?. *IZA Discussion Paper*, Article 14906. <http://dx.doi.org/10.2139/ssrn.4114410>
- Galperin, R. V., Hahl, O., Sterling, A. D., & Guo, J. (2020). Too good to hire? Capability and inferences about commitment in labor markets. *Administrative Science Quarterly*, 65(2), 275–313. <https://doi.org/10.1177/0001839219840022>
- García-Aracil, A., Monteiro, S., & Almeida, L. S. (2021). Students' perceptions of their preparedness for transition to work after graduation. *Active learning in higher education*, 22(1), 49–62. <https://doi.org/10.1177/1469787418791026>
- Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1), 109–128. <https://doi.org/10.1037/a0022530>

- Gerdes, E. P., & Garber, D. M. (1983). Sex bias in hiring: Effects of job demands and applicant competence. *Sex Roles: A Journal of Research*, 9(3), 307–319.
<https://doi.org/10.1007/BF00289666>
- Hacohen, R., & Nicks, L. (2019, November 20). Women only apply for jobs when 100% qualified. Fact or fake news?. *The Behavioral Insights Team*.
<https://www.bi.team/blogs/women-only-apply-for-jobs-when-100-qualified-fact-or-fake-news/>
- Hannon, K. (2014, September 11). Are women too timid when they job search?. *Forbes*.
<https://www.forbes.com/sites/nextavenue/2014/09/11/are-women-too-timid-when-they-job-search/?sh=73d9b451411d>
- Hardy III, J. H., Tey, K. S., Cyrus-Lai, W., Martell, R. F., Olstad, A., & Uhlmann, E. L. (2022). Bias in context: Small biases in hiring evaluations have big consequences. *Journal of Management*, 48(3), 657–692. <https://doi.org/10.1177/0149206320982654>
- Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269–298.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657–674.
<https://doi.org/10.1111/0022-4537.00234>
- Hentschel, T., Horvath, L. K., Peus, C., & Sczesny, S. (2018). Kick-starting female careers: Attracting women to entrepreneurship programs. *Journal of Personnel Psychology*, 17(4), 193–203. <https://doi.org/10.1027/1866-5888/a000209>
- Hewitt, P. L., Flett, G. L., Sherry, S. B., Habke, M., Parkin, M., Lam, R. W., McMurtry, B., Ediger, E., Fairlie, P., & Stein, M. B. (2003). The interpersonal expression of perfec-

- tion: Perfectionistic self-presentation and psychological distress. *Journal of Personality and Social Psychology*, 84(6), 1303–1325. <https://doi.org/10.1037/0022-3514.84.6.1303>
- Hormuth, S., & Lalli, M. (1988). Eine Skala zur Erfassung der bereichsspezifischen Selbstzufriedenheit [A scale to measure domain-specific self-esteem]. *Diagnostica*, 34(2), 148–166.
- Jaja, C. (2018, November 27). The science of the job search, part VII: You only need 50% of job “requirements”. *Talent Works*. <https://talent.works/2018/11/27/the-science-of-the-job-search-part-vii-you-only-need-50-of-job-requirements/>
- JASP Team (2023). JASP (Version 0.17.1)[Computer software].
- Kesebir, S., Lee, S. Y., Elliot, A. J., & Pillutla, M. M. (2019). Lay beliefs about competition: Scale development and gender differences. *Motivation and Emotion*, 43(5), 719–739. <https://doi.org/10.1007/s11031-019-09779-5>
- Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individual's fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology*, 58(2), 281–342. <https://doi.org/10.1111/j.1744-6570.2005.00672.x>
- Lawson, K. M., Crouter, A. C., & McHale, S. M. (2015). Links between family gender socialization experiences in childhood and gendered occupational attainment in young adulthood. *Journal of Vocational Behavior*, 90, 26–35. <https://doi.org/10.1016/j.jvb.2015.07.003>
- Lee, F. (2018). *How prepared do I have to be for a job? Gender differences in the desire for preparedness* [Unpublished doctoral dissertation]. Stanford University.

- Levy, L. M., Gandy, M., Solomon, R., McLaughlin, A. C., Allaire, J. C., & Whitlock, L. A. (2012, May 29–June 1). *Fear of failure: Gender differences in older adult gamers* [Poster presentation]. International Conference on the Foundations of Digital Games, Raleigh, NC, USA. <https://doi.org/10.1145/2282338.2282405>
- Liu, S., Huang, J. L., & Wang, M. (2014). Effectiveness of job search interventions: A meta-analytic review. *Psychological Bulletin*, 140(4), 1009–1041. <https://doi.org/10.1037/a0035923>
- Mandrik, C. A., & Bao Y. (2005). Exploring the concept and measurement of general risk aversion. *Advances in Consumer Research*, 32, 531–539.
- Markman, A. (2019, May 20). You don't need to meet every qualification to apply for a job. *Harvard Business Review*. <https://hbr.org/2019/05/you-dont-need-to-meet-every-qualification-to-apply-for-a-job>
- Matlin, M. W. (2004). *The psychology of women* (5th edition). Wadsworth/Thomson Learning.
- Mohan, M., & Sundari Ravindran, T. K. (2018). Conceptual framework explaining “preparedness for practice” of dental graduates: A systematic review. *Journal of Dental Education*, 82(11), 1194–1202. <https://doi.org/10.21815/JDE.018.124>
- Mohr, T. (2014, August 25). Why women don't apply for jobs unless they're 100% qualified. *Harvard Business Review*. <https://hbr.org/2014/08/why-women-dont-apply-for-jobs-unless-theyre-100-qualified>
- Morgenroth, T., Ryan, M. K., & Fine, C. (2022). The gendered consequences of risk-taking at work: Are women averse to risk or to poor consequences?. *Psychology of Women Quarterly*, 46(3), 257–277. <https://doi.org/10.1177/03616843221084048>

- Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in women's self-promotion: The backlash avoidance model. *Psychology of Women Quarterly*, 34(2), 186–202.
<https://doi.org/10.1111/j.1471-6402.2010.01561.x>
- Nelson, K. L., Newman, D. N., McDaniel, J. R., & Buboltz, W. C. (2013). Gender differences in fear of failure amongst engineering students. *International Journal of Humanities and Social Science*, 3(16), 10–16. https://www.researchgate.net/publication/335703198_Gender_Differences_in_Fear_of_Failure_amongst_Engineering_Students/link/5d769d6f4585151ee4ab02dd/download
- Nicks, L., Gesiarz, F., Valencia, L., Hardy, T. & Lohmann, J. (2022, March). Gender differences in response to requirements in job adverts. *Research Report*.
<https://www.bi.team/wp-content/uploads/2022/03/Gender-differences-in-response-to-requirements-in-job-adverts-March-2022.pdf>
- Niessen-Ruenzi, A., & Zimmerer, L. (2023). The value of skill signals for women's careers. *European Corporate Governance Institute – Finance Working Paper No. 888/2023*.
<https://ssrn.com/abstract=3987238>
- Player, A., Randsley de Moura, G., Leite, A. C., Abrams, D., & Tresh, F. (2019). Overlooked leadership potential: The preference for leadership potential in job candidates who are men vs. Women. *Frontiers in Psychology*, 10, Article 755.
<https://doi.org/10.3389/fpsyg.2019.00755>
- Portillo, S., & DeHart-Davis, L. (2009). Gender and organizational rule abidance. *Public Administration Review*, 69(2), 339–347. <https://doi.org/10.1111/j.1540-6210.2008.01978.x>

- Prentice, D. A., & Carranza, E. (2002). What women should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26(4), 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Rice, C. (2014, April 24). How McKinsey's story became Sheryl Sanberg's statistic – and why it didn't deserve to. *Huffington Post*. https://www.huffingtonpost.co.uk/curt-rice/how-mckinseys-story-became-sheryl-sandbergs-statistic-and-why-it-didnt-deserve-to_b_5198744.html?guccounter=1
- Rojas, M. (2021, August 3). Dear female job seeker: Apply for the job, ignore the 'qualifications'. *FastCompany*. <https://www.fastcompany.com/90661349/dear-female-jobseeker-apply-for-the-job-ignore-the-qualifications>
- Saccardo, S., Pietrasz, A., & Gneezy, U. (2018). On the size of the gender difference in competitiveness. *Management Science*, 64(4), 1541–1554. <https://doi.org/10.1287/mnsc.2016.2673>
- Salwender, M., Schoel, C., Bless, H., & Stahlberg, D. (2023). The politics hurdle: Joint effect of organizational culture and gender on lack of fit experiences. *Social Psychological and Personality Science*, 14(1), 84–92. <https://doi.org/10.1177/19485506221075898>
- Sandberg, S. (2015). *Lean in: Women, work, and the will to lead*. WH Allen.
- Schmidt, J., & Stettes, O. (2018, April 30). Frauen in Führungspositionen [Women in leadership positions]. *IW-Report*. <https://www.iwkoeln.de/studien/joerg-schmidt-oliver-stettes-frauen-in-fuehrungspositionen.html>
- Shaw, K., Holbrook, A., & Bourke, S. (2013). Student experience of final-year undergraduate research projects: an exploration of 'research preparedness'. *Studies in Higher Education*, 38(5), 711–727. <https://doi.org/10.1080/03075079.2011.592937>

SOEP-IS Group. (2017). *SOEP-IS 2015 – Fragebogen für die SOEP-Innovations-Stichprobe* [Questionnaire for the SOEP-Innovation-Sample]. DIW Berlin/SOEP.

https://www.diw.de/documents/publikationen/73/diw_01.c.571107.de/diw_ssp0447.pdf

Statistisches Bundesamt. (2021, November 11). Studierende an Hochschulen: Fächersystematik [Students at universities: Subject classification]. https://www.destatis.de/DE/Methoden/Klassifikationen/Bildung/studenten-pruefungsstatistik.pdf?__blob=publication-File

Svirko, E., Lambert, T., & Goldacre, M. J. (2014). Gender, ethnicity and graduate status, and junior doctors' self-reported preparedness for clinical practice: National questionnaire surveys. *Journal of the Royal Society of Medicine*, 107(2), 66–74.

<https://doi.org/10.1177/0141076813502956>

Szillis, U., & Stahlberg, D. (2007). The face-ism effect in the internet differences in facial prominence of women and men. *International Journal of Internet Science*, 2(1), 3–11.

Tannen, D. (1990). *You just don't understand! Women and men in conversation*. Morrow.

Thandi, H., & Sharma, R. (2004). MBA Students' preparedness for entrepreneurial efforts. *Tertiary Education & Management*, 10(3), 209–226.

<https://doi.org/10.1080/13583883.2004.9967128>

Tockey, D., & Ignatova, M. (n.d.). Gender insights report. *LinkedIn*. <https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions-lodestone/body/pdf/Gender-Insights-Report.pdf>

- Travaglianti, F., Babic, A., & Hansez, I. (2016). The role of work-related needs in the relationship between job crafting, burnout and engagement. *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, 42(1), a1308.
<http://dx.doi.org/10.4102/sajip.v42i1.1308>
- Van Hooft, E. A. J., Born, M. P., Taris, T. W., & Van der Flier, H. (2006). Ethnic and gender differences in applicants' decision-making processes: An application of the theory of reasoned action. *International Journal of Selection and Assessment*, 14(2), 156–166.
<https://doi.org/10.1111/j.1468-2389.2006.00341.x>
- Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (s)expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior*, 82(3), 208–220.
<https://doi.org/10.1016/j.jvb.2013.01.008>
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432.
<https://doi.org/10.1037/a0022790>
- Wille, L., & Deros, E. (2018). When job ads turn you down: How requirements in job ads may stop instead of attract highly qualified women. *Sex Roles: A Journal of Research*, 79(7-8), 464–475. <https://doi.org/10.1007/s11199-017-0877-1>
- Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, 142(2), 165–197. <https://doi.org/10.1037/bul0000039>

- Youn, S. (2019, March 8). Women are less aggressive than men when applying for jobs, despite getting hired more frequently: LinkedIn. *abc News*. <https://abcnews.go.com/Business/women-aggressive-men-applying-jobs-hired-frequently-linkedin/story?id=61531741>
- Yu, Q., & Li, B. (2017). mma: An R package for mediation analysis with multiple mediators. *Journal of Open Research Software*, 5(11). <http://doi.org/10.5334/jors.160>

The Politics Hurdle: Joint Effect of Organizational Culture and Gender on Lack of Fit Experiences

Mona Salwender¹ , Christiane Schoel¹, Herbert Bless¹,
and Dagmar Stahlberg¹

Social Psychological and
Personality Science
2023, Vol. 14(1) 84–92
© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/19485506221075898

journals.sagepub.com/home/spp



Abstract

We propose that an organizational culture where playing politics is important for advancement, compared with an organizational culture where showing competencies is important, elicits stronger lack of fit experiences for women than for men. In a pre-study, playing politics was perceived as dominant, typically male work behaviors, whereas showing competencies was perceived as competent, typically female work behaviors. We then tested in two experiments (689 individuals, integrated in a small-scale meta-analysis) the joint effect of organizational culture and gender on four lack of fit indicators (self-concept conflict, fear of backlash, intention to seek power positions, concerns about one's skills). As expected, women indicated more lack of fit experiences than men in politics cultures, but not in competencies cultures. Our findings suggest that perceived organizational culture may play an important role in understanding the dynamics of career advancement of women and men.

Keywords

organizational culture, gender, leadership, lack of fit, stereotypes

What is required to advance to higher positions in an organization? Not surprisingly, research on antecedents and lay theories of power identified competence, structural position (e.g., access to information), demographics, and personality as predictors for advancement (Anderson & Brion, 2014; Belmi & Laurin, 2016; ten Brinke & Keltner, 2020). In addition, and of prime interest in the present research, political behaviors such as deciding strategically, using social connections instrumentally, and forming alliances have been discussed as necessary and useful to attain higher positions (Anderson & Brion, 2014; Belmi & Laurin, 2016). Playing politics is different from single instrumental work behaviors (e.g., instrumental networking) as it includes multiple behaviors focusing on advancement to power positions using a “pragmatic and Machiavellian approach to impression management and social relationships to get ahead” (Belmi & Laurin, 2016, p. 505). Although strategic networks and assertive self-promotion drive success (Belmi & Laurin, 2016), playing politics is often viewed negatively (Doldor et al., 2013; Pfeffer, 2010). After all, playing politics behaviors are manipulative and self-interested behaviors (Belmi & Laurin, 2016).

In this article, we propose that, at least for some individuals, playing politics may violate self-definitions and societal expectations regarding adequate behavior and will, therefore, lead to lack of fit experiences. This should be especially true for women for whom society prescribes

communal behavior and proscribes dominant behavior (Prentice & Carranza, 2002).

This assumption is in line with previous findings: Even though women and men both rate playing politics as effective and necessary behaviors to acquire power (Belmi & Laurin, 2016, Online Supplement), and do not differ in their political skill levels or effective use of those skills (Harris et al., 2007; Judge & Bretz, 1994; Todd et al., 2009; Treadway et al., 2005), qualitative interview studies suggest that politics cultures constitute a comparatively high hurdle for women's advancement. Women described playing politics as dominant, stereotypical male behaviors, reported to dislike playing politics, perceived playing politics to be at odds with female identity, and preferred success built on achievements (Davey, 2008; Doldor et al., 2013).

We further argue that the gender difference regarding lack of fit experiences expected in a politics culture should be less pronounced or even disappear in gender-neutral competencies cultures that require doing an excellent job, being highly motivated, and being a team player to succeed.

¹University of Mannheim, Germany

Corresponding Author:

Mona Salwender, Chair of Social Psychology, University of Mannheim, A5.6, Mannheim, 68159, Germany.

Email: mona.salwender@gmail.com

The competence facet of agency is nowadays perceived as gender-neutral, whereas the dominance facet of agency is still stereotypically associated with masculinity, and communion is still stereotypically associated with femininity (Abele et al., 2016; Eagly et al., 2019). For men, lack of fit experiences in competencies cultures might depend on how much stereotypical female communal skills are emphasized in a specific competencies culture.

In the present research, we examine four indicators of a lack of fit experience. First, self-concept conflict reflects the experienced mismatch between one's self-concept and an organization's culture (Belmi & Laurin, 2016). Due to gender stereotypes and socialization, women's (compared with men's) self-concepts are more strongly characterized by communion and less strongly characterized by the dominance component of agency (Guimond et al., 2006; Hentschel et al., 2019). These self-concepts can be hypothesized to contrast with the dominant, typically male behaviors required in politics cultures. If so, politics cultures should elicit more self-concept conflict for women than for men, whereas in competencies cultures differences in self-concept should be less pronounced.

Second, fear of backlash is defined as being afraid of others' reactions when violating expectations based on stereotypes (Moss-Racusin & Rudman, 2010). Gender stereotypes prescribe women to show communal and proscribe to show dominant behaviors (Prentice & Carranza, 2002). According to prior research, women are evaluated negatively for showing dominance (i.e., backlash), but not for showing competence or communion (Rudman & Glick, 2001; Williams & Tiedens, 2016). Moreover, women anticipate backlash (i.e., fear of backlash) and consequently self-limit their behavior (Amanatullah & Morris, 2010; Moss-Racusin & Rudman, 2010). So far, fear of backlash has been studied in specific situations (e.g., self-promotion, negotiation, Amanatullah & Morris, 2010; Moss-Racusin & Rudman, 2010). Going beyond specific situations, we examine whether organizational culture can induce general feelings of fear of backlash as one indicator of lack of fit experiences.

Third, lack of fit experiences (e.g., self-concept conflict) may lead to self-limiting behaviors (Heilman, 1983). These self-limiting behaviors can be reflected in lowered intentions to seek power which constitute a more indirect measure of lack of fit. This assumption can be built on research by Belmi and Laurin (2016) who found that individuals with low (vs. high) social class strive less for power when playing politics (vs. showing competencies¹) is important for advancement. Interestingly, with respect to gender differences in power motivation, prior research has revealed mixed findings. Some research suggest similar levels of power motivation for women and men (Anderson & Brion, 2014), whereas other research reports a lower implicit (Denzinger et al., 2016) and explicit power motive in women than in men (Gino et al., 2015). Although women

view advancement as equally attainable, they perceive power as less desirable (Gino et al., 2015). These mixed findings point toward moderating variables. We propose that organizational cultures can function as moderators and hypothesize increased gender differences in the intention to seek power positions in politics (compared with competencies) cultures.

Fourth, limited experience and self-stereotyping may raise one's concerns about one's skills in playing politics (Belmi & Laurin, 2016). If so, women may believe that they lack the necessary skills to play politics (vs. showing competencies).

This present research is—to the best of our knowledge—the first quantitative, experimental investigation of the joint effect of organizational culture and gender on lack of fit experiences (in particular, including critical comparisons between different organizational cultures) showcasing the importance to disentangle the competence and dominance facets of agency. This approach, thus, also echoes the call to address structural hurdles in gender research (England et al., 2020) by linking structural aspects (organizational culture) with individuals' resulting lack of fit experiences.

Pre-Study

The assumption that playing politics is perceived as typically male behavior and showing competencies as gender-neutral (or slightly female behavior) has so far been only addressed in qualitative research (based on interviews with women). To our knowledge, quantitative research with female and male participants has not yet been reported. Therefore, in a first step, we tested whether women and men both perceive playing politics as typically male behavior and showing competencies as gender-neutral (or even slightly female behavior). Moreover, we assessed dominance and competence as facets of agency as well as warmth and morality as facets of communion. We expected that playing politics is perceived by both women and men as more dominant, less competent, less warm, and less moral behavior compared with showing competencies.

Method

Participants. We recruited 100 participants via a participant pool from University of Mannheim and via posts on social media (75% female, $M_{\text{age}} = 33.90$, $SD_{\text{age}} = 11.90$, 97% with higher educational level, 30% students, 72% working, average work experience $M = 9.40$ years, $SD = 10.80$, 16% held a leadership position). For all studies reported participant treatment was in accordance with ethical standards from APA and University of Mannheim, participants could sign up for a lottery for Amazon vouchers, and materials, data and code for all studies are available through OSF (https://osf.io/astpf/?view_only=5774fd3a384b426aaff0e2d4c8ba506b).

Procedure and Design. After providing their demographics, participants were presented with two sets of work behaviors. Seven behaviors described the examples for playing politics and seven examples described showing competencies reflecting a one-factor (playing politics vs. showing competencies) within-subjects design. An example for playing politics is “Make decisions strategically, that is, decide on a purely business basis, and not pay attention to the sensitivities of others (e.g., colleagues).” An example for showing competencies is “Do excellent work and be competent at what you do” (based on Belmi & Laurin, 2016). Each behavior was evaluated on several dimensions (see below). The order of the two sets and the order of the behaviors within one set were randomized. Finally, participants indicated their self-ascribed masculinity-femininity (we expected comparable results for biological sex and self-ascribed gender identity; see Online Supplement).

Measures

Gender Typicality Rating. Participants rated how typically feminine or masculine they perceive described work behaviors to be ($-3 = \text{typically female}$ to $+3 = \text{typically male}$). The reliabilities of these ratings were $\alpha = .67$ for playing politics and $\alpha = .62$ for showing competencies.

Agency and Communion Ratings. Participants rated each behavior in terms of agency and communion. For each behavior, participants indicated how much they associated a person showing the behavior in question with 12 different adjectives ($1 = \text{not at all}$ to $7 = \text{very much}$). Three adjectives each represented the four sub-facets dominance, competence, warmth, and morality (Abele et al., 2016). For example, participants indicated how much they associated “Do excellent work and be competent at what you do” with a person who is assertive (one of three indicators for dominance; $\alpha_{\text{politics}} = .88$; $\alpha_{\text{competencies}} = .87$), efficient (competence; $\alpha_{\text{politics}} = .91$; $\alpha_{\text{competencies}} = .87$), caring (warmth; $\alpha_{\text{politics}} = .91$; $\alpha_{\text{competencies}} = .86$), and trustworthy (morality; $\alpha_{\text{politics}} = .88$; $\alpha_{\text{competencies}} = .85$).

Results and Discussion

As predicted, playing politics was perceived as a typically male behavior ($M = 0.61$, $SD = 0.67$), $t(99) = 9.17$, $p < .001$, $d = 0.92$ (significant deviation from scale midpoint zero), whereas showing competencies was perceived as a typically female behavior ($M = -0.43$, $SD = 0.52$), $t(99) = -8.22$, $p < .001$, $d = 0.82$.

Furthermore, a repeated measure ANOVA revealed the expected Work Behavior \times Agency-Communion-Rating interaction, $F(1.84, 159.05) = 420.73$, $p < .001$, $\eta^2 = .81$ (with Greenhouse–Geisser correction). Participants rated playing politics compared with showing competencies as higher in dominance (mean difference of 0.93, $SE = 0.07$, $p < .001$), lower in competence (mean difference of 1.07,

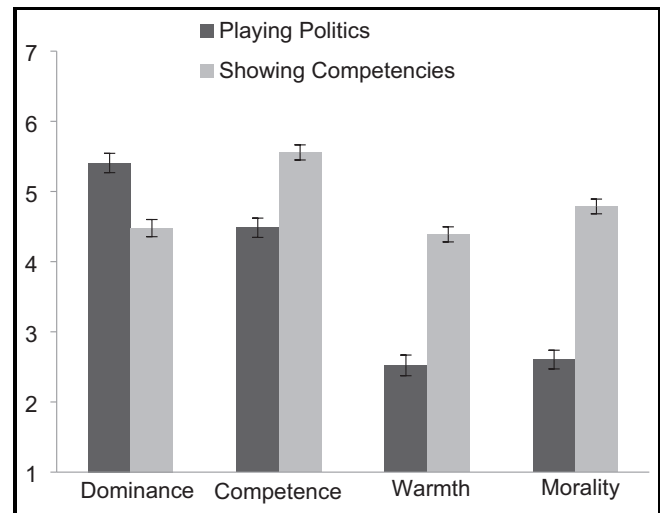


Figure 1. Participants' Dominance, Competence, Warmth, and Morality Ratings of Playing Politics and Showing Competencies
Note. Participants' ratings of how much ($1 = \text{not at all}$ to $7 = \text{very much}$) they associate showing competencies and playing politics behaviors with sub-facets of agency and communion. Error bars represent 95% confidence intervals.

$SE = 0.08$, $p < .001$), lower in warmth (mean difference of 1.87, $SE = 0.08$, $p < .001$), and lower in morality (mean difference of 2.18, $SE = 0.08$, $p < .001$; see Figure 1).

These results confirm our starting assumption that playing politics is perceived as typically male, dominant behavior. Showing competencies is perceived as typically female, competent, communal behavior. Women and men seem to share this perception. These findings underline the importance to disentangle the competence and dominance facets of agency.

Studies 1a and 1b

The pre-study clearly indicated that playing politics (compared with showing competencies) was perceived as stereotypic male and dominant behavior. Given these perceptions, organizational cultures emphasizing playing politics are more likely to elicit lack of fit experiences in women than in men. In Study 1a, participants imagined working for a company either with a playing politics culture or a showing competencies culture. Study 1b was a direct replication of Study 1a using the same design and materials to test the stability and reliability of the results and increase test power. We hypothesized that women would report more lack of fit experiences (i.e., self-concept conflict, fear of backlash, intention to seek power positions, and concerns about one's skills) in an organizational culture of playing politics, whereas this gender difference should be less pronounced in an organizational culture of showing competencies.

Method

We present an integrative data analysis based on the pooled raw data of the two independent samples of Studies 1a and 1b including a dummy indicating the respective sample, as well as a small-scale meta-analysis (Curran & Hussong, 2009). Separate analyses per sample are reported in the Online Supplement. Although the descriptive pattern of results was highly consistent, differences in significances between the pooled data and the individual studies are reported in footnotes.

Participants. We recruited university students via www.surveycircle.com and by spreading the study link via e-mail at multiple universities across Germany. We excluded three participants in Study 1a and eight participants in Study 1b who indicated at the end of the study that they did not answer the questions conscientiously. Furthermore, we excluded two participants in Study 1b with gender category “other” as this sample was too small for statistical analyses and 10 participants in Study 1b who indicated they had participated in this study before. For Study 1a, a power analysis for an ANOVA with interactions revealed a minimum sample size of $N=128$ to detect a medium-sized effect, $f=0.25$, with a power of $(1-\beta)=.80$ at $\alpha=.05$ for the proposed univariate Participant Sex \times Culture interactions (G*Power, Faul et al., 2007). After reaching the suggested sample size, we continued recruitment until each condition contained at least 50 respondents (see suggestion by Simmons et al., 2013) resulting in 317 participants. With those 317 observations in four groups the analysis had 80% power to detect an interaction effect of $f \geq 0.16$ at $\alpha=.05$. For replication purposes, for Study 1b we aimed to reach a sample comparable to Study 1a. The resulting sample size was 372 for Study 1b. The integrative data analysis based on the pooled raw data of Studies 1a and 1b ($N=689$, 68% women) had 80% power to detect univariate Participant Sex \times Culture interaction effects of $f \geq 0.11$.³ See Table S11 in the Online Supplement for detailed demographics of the participants in Studies 1a and 1b.

Procedure and Design. After providing their demographics, participants had to imagine starting their career after university in a well-regarded (fictitious) German consulting company Swathmore International. All participants read that they would enter, as every new employee does, as Associate Consultant with the option to advance to higher positions. Then half of the participants learned, allegedly from sources within the company, that it would take playing politics to move up the ranks, whereas the other half learned that it would take showing one's competencies. This reflects a 2 (participant sex: male vs. female) \times 2 (condition: politics culture vs. competencies culture) between-subjects design with random assignment to conditions. We presented the same examples for showing competencies or playing politics as in

the pre-study to describe the respective company culture.⁴ Afterward, participants answered a manipulation check, indicated their self-concept conflict and fear of backlash, as well as their intention to seek power and concerns about their skills when using the strategies that promise advancement in the respective organization. Last, we added some exploratory questions (e.g., self-ascribed masculinity-femininity). Procedure and materials were adopted from Study 2 from Belmi and Laurin (2016).

Measures. Participants responded to all measures on 7-point rating scales ranging from 1 = *strongly disagree* to 7 = *strongly agree* if not indicated otherwise.

Manipulation Check. We asked participants to rate their agreement with two statements: “What people at Swathmore International primarily need to advance to a higher position is strategic skills” and “What people at Swathmore International primarily need to advance to a higher position is hard work and motivation.”, $r(687) = -.34$, $p < .001$ (Belmi & Laurin, 2016). We also asked participants at the end of the study how well they could imagine the described situation (*not at all* to *very well*).

Self-Concept Conflict. Participants indicated their agreement with six statements ($\alpha=.96$) about self-concept conflicts regarding the behaviors needed to advance to a higher position, for example, “I feel that these strategies conflict with who I am as a person” (Belmi & Laurin, 2016).

Fear of Backlash. We assessed fear of backlash with an adapted selection of six items ($\alpha=.92$) from Rudman and Fairchild's (2004) fear of backlash scale. Participants were asked to imagine performing the behaviors needed to advance to a higher position (in the respective experimental condition) and to indicate to what extent (1 = *not at all* to 7 = *very much*) they would fear their colleagues' reactions, for example, “Would you be concerned that your colleagues might dislike you?”⁵

Intention to Seek Power. We asked for the participants' intention to seek power when thinking about the respective behaviors needed to advance to a high position with four questions, for example, “I would aspire to be in a high-ranking position in this organization” (Belmi & Laurin, 2016). In addition, we presented stairs depicting the seven different positions available in the company from lowest [*Assistant Consultant*] to highest [*Partner*] and asked “Which position would you like to occupy in this organization in the future?” (Belmi & Laurin, 2016). We combined both measures by standardizing and then averaging the items ($\alpha=.96$).

Concerns About One's Skills. We assessed participants' concerns about their skills with two items, $r(687) = .52$,

$p < .001$ (Belmi & Laurin, 2016), for example, “I feel that I do not have the skills to perform these behaviors.”

Additional Measures. Identical to the pre-study, we assessed participants’ social gender with the German version of the Traditional Masculinity-Femininity Scale (Kachel et al., 2016; see Online Supplement). Furthermore, as prior research found that social class predicts the intention to seek power positions (Belmi & Laurin, 2016), we included a measure of the participants’ subjective social class (Euteneuer et al., 2015). Controlling for social class led to consistent results (see Online Supplement).

Results

Manipulation Check. Participants could overall imagine the situation well ($M = 5.11$, $SD = 1.56$), with no significant differences due to participant sex, condition, study, and the interactions (all $ps > .076$). As intended, participants in the politics culture condition rated playing politics ($M = 6.51$, $SD = 0.98$) as more helpful than showing competencies ($M = 4.10$, $SD = 1.97$), whereas the reverse pattern emerged in the competencies culture, $M_{\text{competencies}} = 6.51$, $SD = 0.94$ versus $M_{\text{politics}} = 4.58$, $SD = 1.83$; Manipulation Check Items \times Condition interaction, $F(1, 681) = 591.98$, $p < .001$, $\eta^2 = .47$. Importantly, the three-way interaction Manipulation Check Items \times Participant Sex \times Study

Dummy as well as the other two-way interactions were not significant, $F_s < 1.82$, $ps > .178$.

Lack of Fit—Multivariate Results. A MANOVA (see Table S12 in the Online Supplement for intercorrelations of the lack of fit indicators) revealed a significant main effect for participant sex, $F(4, 678) = 9.84$, $p < .001$, $\eta^2 = .06$, a significant main effect for condition, $F(4, 678) = 206.64$, $p < .001$, $\eta^2 = .55$, and the hypothesized Participant Sex \times Condition interaction, $F(4, 678) = 5.87$, $p < .001$, $\eta^2 = .03$. The main effect of study, the two-way interactions with study as well as the three-way interaction were non-significant (see Online Supplement, also for the following results). Planned univariate 2 (participant sex: male vs. female) \times 2 (condition: politics culture vs. competencies culture) contrasts are reported below (see Figure 2).

Self-Concept Conflict. Participants reported more self-concept conflict in the politics compared with the competencies culture condition, $F(1, 681) = 760.35$, $p < .001$, $\eta^2 = .53$. Crucially, this effect was more pronounced for women than for men, $F(1, 681) = 19.06$, $p < .001$, $\eta^2 = .03$ (Participant Sex \times Condition interaction). As predicted, women reported more self-concept conflict in the politics culture condition than men, $t(183.01) = -3.90$, $p < .001$. Unexpectedly, in the competencies culture condition, men

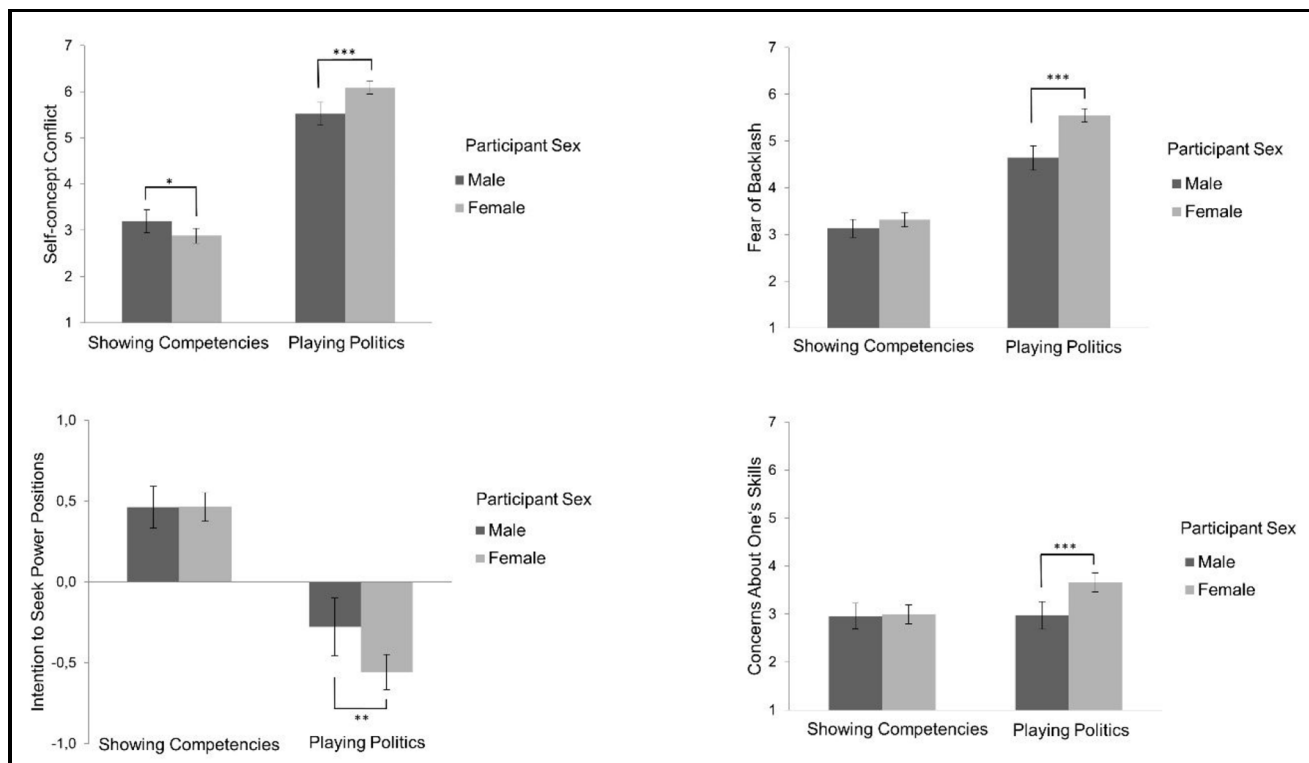


Figure 2. Lack of Fit Experiences Depending on Organizational Culture and Participant Sex

Note. Error bars represent 95 % confidence intervals.

* Indicates a significant contrast at $p < .050$. ** Indicates a significant contrast at $p < .010$. *** Indicates a significant contrast at $p < .001$.

reported more self-concept conflict than women, $t(203.83) = 2.14, p = .034$.⁶

Fear of Backlash. Women reported more fear of backlash than men, $F(1, 681) = 30.58, p < .001, \eta^2 = .04$. Furthermore, participants in the politics culture condition reported more fear of backlash than participants in the competencies culture condition, $F(1, 681) = 372.48, p < .001, \eta^2 = .35$. Most importantly, the latter effect was more pronounced for women than for men, $F(1, 681) = 13.73, p < .001, \eta^2 = .02$ (Participant Sex \times Condition interaction). In line with our hypothesis, women reported more fear of backlash in the politics culture condition than men, $t(178.24) = -5.97, p < .001$, whereas no gender difference emerged in the competencies culture condition, $t(246.37) = -1.43, p = .153$.

Intention to Seek Power Positions. Women reported less intention to seek power positions than men, $F(1, 681) = 4.68, p = .031, \eta^2 = .01$.⁷ Furthermore, participants indicated less intention to seek power positions in the politics culture condition compared with the competencies culture condition, $F(1, 681) = 180.10, p < .001, \eta^2 = .21$. With respect to our hypothesis, women reported a lower intention to seek power positions than men in the politics culture, $t(197.77) = 2.64, p = .009$,⁸ whereas this difference did not emerge in the competencies culture condition, $t(219.90) = -0.01, p = .99$; $F(1, 681) = 4.37, p = .037, \eta^2 = .01$ (Participant Sex \times Condition interaction).

Concerns About One's Skills. Women reported more concerns about their skills than men, $F(1, 681) = 7.96, p = .005, \eta^2 = .01$.¹⁰ Participants in the politics culture condition reported more concerns about their skills than those in the competencies culture condition, $F(1, 681) = 7.71, p = .006, \eta^2 = .01$. A significant Participant Sex \times Condition interaction emerged, $F(1, 681) = 6.91, p = .009, \eta^2 = .01$.¹¹ As expected, women reported more competence-based concerns in the politics culture condition than men, $t(685) = -3.90, p < .001$, whereas no gender difference emerged in the competencies culture condition, $t(685) = -0.19, p = .847$.

Robustness Check With Small-Scale Meta-Analysis. Beyond the integrated data analysis with the pooled data sets, we conducted a small-scale meta-analysis (fixed-effect model) across the individual studies. We transformed the η^2 of the interactions into d -effect sizes and weighted them by their sample's inverse variance weight (i.e., one over the square of their standard errors, Lipsey & Wilson, 2001). Results are summarized in Table 1. The small-scale meta-analysis demonstrates significant Participant Sex \times Condition interactions for all four lack of fit indicators across Studies 1a and 1b. The effect sizes indicate small effects according

Table 1. Results of the Small-Scale Meta-Analysis With the Participant Sex (Female vs. Male) \times Condition (Politics Culture vs. Competencies Culture) Interaction Effects of Studies 1a and 1b

Dependent variable	d	95% CI
Self-concept conflict	0.34	[0.23, 0.45]
Fear of backlash	0.28	[0.18, 0.39]
Intention to seek power	0.16	[0.05, 0.27]
Concerns about one's skills	0.21	[0.10, 0.32]

Note. 95% CI = 95% confidence interval for d .

to Cohen's (1988) classification, with self-concept conflict showing the strongest gender difference in lack of fit depending on organizational culture.

Discussion

The obtained findings suggest that women experience a stronger lack of fit than men when faced with situations in which playing politics (relative to showing competencies) is the key driver for career advancement. Beyond this significant interaction effect, it is noteworthy that lack of fit experiences was lower for all participants in the politics compared with the competencies condition. Importantly, the interaction pattern holds after controlling for subjective social class (see Online Supplement). Supporting the reliability of these findings, consistent results were obtained across two studies. The small-scale meta-analysis suggests that the hypothesized effect is presumably smaller for the intention to seek power than for the other three indicators of lack of fit experiences.

General Discussion

In summary, the present study clearly demonstrated the joint effects of gender and organizational cultures on lack of fit expectations of potential applicants for higher organizational positions. The results of a pre-study document that playing politics (vs. showing competencies) are perceived as dominant, typically male behavior, thus supporting prior qualitative research (Davey, 2008; Doldor et al., 2013). Subsequently, across two studies, women indicated more lack of fit experiences than men in organizational cultures where playing politics (vs. showing competencies) drives success. Effects were consistent across four different indicators of lack of fit experiences: Women reported more self-concept conflict, more fear of backlash, more concerns regarding the skills allegedly responsible for success in the respective organizations, and less intentions to seek power positions than men.

Presumably, the stronger experienced lack of fit results from women's internalized or perceived prescriptive stereotypes to be communal and prescriptive stereotypes to refrain from being dominant (e.g., forming alliances for

one's own career advancement, see Belmi & Laurin, 2016). However, future studies will have to explicitly test these assumptions (and the relative importance of communal prescriptions and dominance proscriptions). Moreover, it would be particularly interesting to examine whether women's fear of being negatively viewed for behaving politically has a valid basis. Even if the evaluation of dominant men and women does not differ, as reported by Steffens and colleagues (2009, see also Williams & Tiedens, 2016, for reduced backlash effects outside North America), women's lack of fit experiences are still highly consequential due to the self-limiting behaviors.

Effect sizes for the different indicators varied in our studies but proved reliable in a small-scale meta-analysis. Interestingly, the weaker effects regarding women's reduced intention to seek power in political culture organizations resemble inconsistent findings regarding gender differences in the intention to seek power in prior research (e.g., Anderson & Brion, 2014; Denzinger et al., 2016; Gino et al., 2015). One may speculate that, compared with the other indicators of lack of fit, the intention to seek power is a more distal outcome that is not only influenced by joint effects of organizational culture and gender, but also, for example, by prestige or salary.

The obtained findings mirror prior research on the interplay of social class and organizational culture—with the results for women reflecting the same pattern as those for low socio-economic status individuals (Belmi & Laurin, 2016). Importantly, however, women's lack of fit experiences were observed independently of subjective social class.

The present research was based on students' reactions to descriptions of different organizational cultures. Although entering vocational activities for the first time is a crucial decision, it would be interesting to investigate lack of fit of employees already working in organizations with more or less playing politics cultures. One may speculate that the actual confrontation with the respective culture enhances lack of fit.¹²

From a practical point of view, it is also worth stressing that beyond the gender differences that were our major research topic, the present results also show that politics cultures overall are perceived negatively, both by women and men. Therefore, organizations may well be advised to develop an organizational culture where competencies (not politics) are perceived important for advancement. Such a focus will also attract women for leadership positions. In this respect, it is interesting to note that men indicated more self-concept conflict for showing competencies compared with women (for the other three lack of fit measures no gender difference emerged). This might be due to the specific measure of showing competencies (e.g., having good communication skills) used in the present study (based on

the work of Belmi & Laurin, 2016). Organizations may therefore be well-advised to focus on competencies that are not perceived as gender-specific.

Evidence about a stagnating progress in gender equality in the past years has been combined with the call to attend more to structural hurdles (England et al., 2020). Organizational cultures may constitute such structural hurdles as they potentially create boundary conditions for the effectiveness of diversity interventions (Baron et al., 2021). The present research, thus, echoes the call addressing structural aspects by focusing on the interplay of organizational culture and gender on individuals' lack of fit experiences. Although readily acknowledging the necessity of future research, we strongly believe that the present results constitute an important step in understanding that organizational culture affects gender equality at work.

Acknowledgments

The authors thank Kerry L. Marsh and Felicia Pratto for their valuable feedback regarding the design of Studies 1a and 1b.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by departmental budget resources of the Chair for Social Psychology and the Chair for Microsociology and Social Psychology.

Open Practices Statement

Pre-study and Study 1a were not preregistered. We preregistered Study 1b on OSF (https://osf.io/astpf/?view_only=5774fd3a384-b426aaff0e2d4c8ba506b). In an exploratory data analysis in Study 1a we found the interaction term concerning intention to seek power to be significant for participants being highly interested in consulting and therefore aimed to investigate this further. However, in Study 1b interest in consulting significantly differed between the politics and the competencies condition. Therefore, we did not analyze the results as preregistered. Details can be obtained from the first author (msalwend@mail.uni-mannheim.de).

ORCID iD

Mona Salvender  <https://orcid.org/0000-0001-8431-7707>

Supplemental Material

The supplemental material is available in the online version of the article.

Notes

1. We replaced the term “prosocial” used in Belmi and Laurin (2016) with “competencies” culture although we relied on the same materials for the culture manipulation. This was done because five out of the seven work examples items address competence, and only two prosocial behaviors. The results of our pre-study support this change as competency turned out as the most defining characteristic (i.e., highest ratings) across the seven used examples.
2. The two main effects of work behavior, $F(1, 99) = 405.77$, $p < .001$, $\eta^2 = .80$, and of agency-communion-rating, $F(1.68, 166.08) = 404.38$, $p < .001$, $\eta^2 = .80$ (with Greenhouse–Geisser correction), were also significant, but are not interpreted due to the interaction. In some instances, a marginal interaction with participant sex emerged (in both rating tasks) indicating that women’s compared with men’s ratings were more extreme (see Online Supplement).
3. The same f -value applies for the Participant Sex \times Culture \times Study Dummy interaction.
4. The stimulus material was thankfully provided by Peter Belmi.
5. Items were thankfully provided by Lauri Rudman.
6. Study 1a: $p = .071$; Study 1b $p = .248$.
7. Study 1a: $p = .075$; Study 1b: $p = .203$.
8. Study 1a: $p = .071$; Study 1b: $p = .070$.
9. Note that the interaction did not reach significance in Study 1a ($p = .265$) and in Study 1b ($p = .228$), which might be due to low power. The descriptive pattern is highly consistent (see Online Supplement). We nevertheless test the robustness of this finding in a small-scale meta-analysis.
10. Study 1a: $p = .515$; Study 1b: $p = .001$. This is reflected in the marginal participant sex \times study interaction ($p = .055$).
11. The interaction effect was $p = .011$ in Study 1a and $p = .228$ in Study 1b. The descriptive pattern is again highly consistent (see Online Supplement and meta-analysis).
12. We are also grateful to one of our reviewers who pointed out that future research may gain from using a real control group (e.g., no advancement strategies presented) that will allow to test whether playing politics enhances lack of fit experiences in general and gender differences in particular and/or whether showing competencies eliminates those.

References

- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness—Communion with warmth and morality. *Frontiers in Psychology*, 7, Article 1810. <https://doi.org/10.3389/fpsyg.2016.01810>
- Amanatullah, E. T., & Morris, M. W. (2010). Negotiating gender roles: Gender differences in assertive negotiating are mediated by women’s fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology*, 98(2), 256–267. <https://doi.org/10.1037/a0017094>
- Anderson, C., & Brion, S. (2014). Perspectives on power in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 1, 67–97. <https://doi.org/10.1146/annurev-orgpsych-031413-091259>
- Baron, J. A., Ganglmair, B., Persico, N., Simcoe, T., & Tarantino, E. (2021). *Representation is not sufficient for selecting gender diversity* (No. w28649). National Bureau of Economic Research. <https://www.nber.org/papers/w28649>
- Belmi, P., & Laurin, K. (2016). Who wants to get to the top? Class and lay theories about power. *Journal of Personality and Social Psychology*, 111(4), 505–529. <https://doi.org/10.1037/pspi0000060>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, 14(2), 81–100. <https://doi.org/10.1037/a0015914>
- Davey, K. M. (2008). Women’s accounts of organizational politics as a gendering process. *Gender, Work and Organization*, 15(6), 650–671. <https://doi.org/10.1111/j.1468-0432.2008.00420.x>
- Denzinger, F., Backes, S., Job, V., & Brandstätter, V. (2016). Age and gender differences in implicit motives. *Journal of Research in Personality*, 65, 52–61. <https://doi.org/10.1016/j.jrp.2016.09.003>
- Doldor, E., Anderson, D., & Vinnicombe, S. (2013). Refining the concept of political will: A gender perspective. *British Journal of Management*, 24(3), 414–427. <https://doi.org/10.1111/j.1467-8551.2012.00812.x>
- Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2019). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, 75(3), 301–315. <https://doi.org/10.1037/amp0000494>
- England, P., Levine, A., & Mishel, E. (2020). Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences*, 117(13), 6990–6997. <https://doi.org/10.1073/pnas.1918891117>
- Euteneuer, F., Süssbach, P., Schäfer, S. J., & Rief, W. (2015). Subjektiver sozialer Status. MacArthur-Skalen zur Erfassung des wahrgenommenen sozialen Status im sozialen Umfeld (SSS-U) und in Deutschland (SSS-D) [Subjective social status. MacArthur scales for the assessment of perceived social status in the social environment (SSS-U) and in Germany (SSS-D)]. *Verhaltenstherapie*, 25(3), 229–232. <https://doi.org/10.1159/000371558>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Gino, F., Wilmuth, C. A., & Brooks, A. W. (2015). Compared to men, women view professional advancement as equally attainable, but less desirable. *Proceedings of the National Academy of Sciences*, 112, 12354–12359. <https://doi.org/10.1073/pnas.1502567112>
- Guimond, S., Chatard, A., Martinot, D., Crisp, R. J., & Redersdorff, S. (2006). Social comparison, self-stereotyping, and gender differences in self-construals. *Journal of Personality and Social Psychology*, 90(2), 221–242. <https://doi.org/10.1037/0022-3514.90.2.221>
- Harris, K. J., Zivnuska, S., Kacmar, K. M., & Shaw, J. D. (2007). The impact of political skill on impression management effectiveness. *Journal of Applied Psychology*, 92(1), 278–285. <https://doi.org/10.1037/0021-9010.92.1.278>
- Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269–298.

- Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at Men's and women's characterizations of others and themselves. *Frontiers in Psychology, 10*, Article 11. <https://doi.org/10.3389/fpsyg.2019.00011>
- Judge, T. A., & Bretz, R. D. (1994). Political influence behavior and career success. *Journal of Management, 20*(1), 43–65. <https://doi.org/10.1177/014920639402000103>
- Kachel, S., Steffens, M. C., & Niedlich, C. (2016). Traditional masculinity and femininity: Validation of a new scale assessing gender roles. *Frontiers in Psychology, 7*, Article 956. <https://doi.org/10.3389/fpsyg.2016.00956>
- Lipsey, M. W., & Wilson, D. B. (2001). *Applied social research methods series; Vol. 49. Practical meta-analysis*. SAGE.
- Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in women's self-promotion: The backlash avoidance model. *Psychology of Women Quarterly, 34*(2), 186–202. <https://doi.org/10.1111/j.1471-6402.2010.01561.x>
- Pfeffer, J. (2010). Power play. *Harvard Business Review, 88*(7/8), 84–92. <https://hbr.org/2010/07/power-play>
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*(4), 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology, 87*(2), 157–176. <https://doi.org/10.1037/0022-3514.87.2.157>
- Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743–762. <https://doi.org/10.1111/0022-4537.00239>
- Simmons, J., Nelson, L., & Simonsohn, U. (2013). Life after p-hacking [Paper presentation]. Presentation at the 2013 Society for Personality and Social Psychology Meeting, New Orleans, LA, United States.
- Steffens, M. C., Schult, J. C., & Ebert, I. D. (2009). Feminization of management leads to backlash against agentic applicants: Lack of social skills, not gender, determines low hireability judgments in a student sample. *Psychology Science Quarterly, 51*(1), 16–46.
- ten Brinke, L., & Keltner, D. (2020). Theories of power: Perceived strategies for gaining and maintaining power. *Journal of Personality and Social Psychology, 122*(1), 53–72. <https://doi.org/10.1037/pspi0000345>
- Todd, S. Y., Harris, K. J., Harris, R. B., & Wheeler, A. R. (2009). Career success implications of political skill. *The Journal of Social Psychology, 149*(3), 279–304. <https://doi.org/10.3200/SOCP.149.3.279-304>
- Treadway, D. C., Hochwarter, W. A., Kacmar, C. J., & Ferris, G. R. (2005). Political will, political skill, and political behavior. *Journal of Organizational Behavior, 26*(3), 229–245. <https://doi.org/10.1002/job.310>
- Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin, 142*(2), 165–197. <https://doi.org/10.1037/bul0000039>

Author Biographies

Mona Salwender is a PhD student in Social Psychology at University of Mannheim, Germany. She studies gender differences and similarities in work contexts.

Christiane Schoel is a private lecturer at University of Mannheim, Germany. Her research topics cover, among others, ostracism, narcissism, leadership, and language.

Herbert Bless is a professor of Microsociology and Social Psychology at University of Mannheim, Germany. He has published extensively on a wide range of topics, focusing recently on social inequalities.

Dagmar Stahlberg is a professor of Social Psychology at University of Mannheim, Germany. She has published extensively on a wide range of topics, focusing recently on backlash effects, self-compassion, and gender-fair language.

Handling Editor: Robyn Mallett

Masculine Generics Versus Gender-Inclusive Language: A Meta-Analysis of the Effects of Language Form on the Mental Representations of Gender

Mona Salwender¹, Ira T. Maschmann¹, Viktoria K. Heinzl¹, Sabine Sczesny², and Dagmar Stahlberg¹

¹ Department of Social Psychology, University of Mannheim

² Department of Social Neuroscience and Social Psychology, University of Bern

Author Note

Mona Salwender  <https://orcid.org/0000-0001-8431-7707>

Ira T. Maschmann  <https://orcid.org/0000-0001-5392-6390>

Sabine Sczesny  <https://orcid.org/0000-0002-1666-1263>

Dagmar Stahlberg  <https://orcid.org/0000-0001-5972-0641>

This meta-analysis was pre-registered (https://osf.io/93np4/?view_only=344b97f397b848699c83626349bc4e90), data and materials are available on https://osf.io/ducgp/?view_only=54d7129820454a9488fedceca40d3b72. We have no known conflict of interest to disclose. We thank Carolin Geckeler and Charlotte Grefenstien for helping to code the studies, Luise Metzger and Janin Roessel for their insightful ideas and comments on this research project. We thank all authors who shared their data with us or provided additional information about their studies.

Correspondence concerning this article should be addressed to Mona Salvender, Social Psychology, University of Mannheim, A5 6, 68165 Mannheim, Germany, Email: mona.salwender@gmail.com

Abstract

Since the 1970s, the effects of masculine generics in language—namely the reduced mental representations of women—in comparison to gender-inclusive alternatives, have been discussed in science, politics, media, and society at large. The present research provides the first systematic pre-registered meta-analysis assessing how women are mentally represented and feel represented themselves, in reaction to masculine generics compared to gender-inclusive alternatives. The multi-level random-effects model comprises 357 effect sizes from 95 independent samples involving 19,582 individuals with nine native languages (German, English, French, Italian, Spanish, Hebrew, Dutch, Swedish, and Norwegian). The calculated overall effect size is based on studies that describe the difference in mental representations male participants have of women, and female participants have of other women, depending on the use of masculine generics versus gender-inclusive alternatives (other-perspective) and effect sizes that describe the difference in mental representations female participants have of themselves depending on the language form (self-perspective). The result, $d = 0.46$ (95% CI [0.36, 0.55]), provides evidence for a significantly higher mental representation of women when gender-inclusive alternatives are used compared to masculine generics. This holds for the other- and the self-perspective, female-typed, male-typed, and gender-neutral contexts, grammatical and natural gender languages, singular and plural forms, and between- and within-participant designs. Effects were larger for multi-gendering strategies (e.g., Swedish new gender-encompassing pronoun “hen”) compared to feminization (e.g., he/she) and neutralization strategies (e.g., chairperson). The results inform theory, politics, media, and society at large about the impact of gendered language on the mental representation of women.

Keywords: masculine generics, gender-inclusive language, mental representation, gender, meta-analysis

Public Significance Statement: This meta-analytic review reveals that when gender-inclusive language is used, women are mentally more strongly represented from an observer perspective and feel more addressed themselves than when masculine generics are used, documenting the failure of masculine forms to serve as truly generic.

Masculine Generics Versus Gender-Inclusive Language: A Meta-Analysis of the Effects of Language Form on the Mental Representations of Gender

When asked to draw a *chairperson*, do people draw women more often as exemplars than when being asked to draw a *chairman*? Do women feel more strongly addressed by a job advertisement written in a gender-inclusive form (e.g., “firefighter”) in comparison to a job advertisement that uses masculine generics (“fireman”)? As early as the mid-20th century, linguists Edward Sapir and Benjamin Lee Whorf hypothesized that thoughts are influenced by language (nowadays known as Sapir-Whorf Hypothesis; Lucy, 2001). Since the 1970s, the Sapir-Whorf idea has also been applied to language addressing gender and numerous researchers have examined questions, such as the two outlined above. In the present research we quantify the effect of gender-inclusive language forms (e.g., “chairperson”) versus masculine generics (e.g., “chairman”) on the mental representation of women, both from the other- and the self-perspective. We also test for variations of effect sizes due to relevant potential moderators, such as language family (natural vs. grammatical gender languages), study specifics (e.g., within- or between-participants designs), and attributes of the participants (e.g., age).

Is What is Supposed to be Generic Really Generic? A Theoretical Controversy

Gender is encoded in nearly all language systems (for more detailed overviews including genderless languages see Hellinger & Bußmann, 2001; Stahlberg et al., 2007), but the degree of gender encoding differs between specific languages and language families. In *grammatical gender languages* (e.g., German, Italian, or Hebrew) every noun and pronoun has a grammatical gender, which can be masculine as for instance the German noun “der Löffel” [the spoon], feminine as for instance “die Gabel” [the fork], or neuter as for instance “das Messer” [the knife]. In *natural gender languages* (e.g., English, Swedish, or Norwegian) only personal pronouns (e.g., “she” and “he”) reflect the gender of the person they refer to. Biological/social gender is not grammatically marked in nouns, not even in personal nouns (Stahlberg et al., 2007). For instance, the personal noun “teacher” does not entail any information

about the person's sex/gender and thus can refer to a female and/or male person. In contrast, in grammatical gender languages personal nouns differ depending on a person's sex/gender (e.g., in Spanish "la profesora" refers to a female teacher and "el profesor" refers to a male teacher).

In both, grammatical and natural gender languages, masculine generics exist. Masculine generics represent a language rule, whereby the grammatically masculine form (e.g., "el profesor", "he") refers to either a male person, a group of male persons, but also a group of people of mixed gender. Further, masculine generics can also be used when a person's gender is unknown or irrelevant (Gygax & Gabriel, 2010; Gygax et al., 2009; Stahlberg et al., 2007). From a classical linguistic position, masculine generics include female and diverse genders not only in grammar but also in interpretation. According to this view, masculine generics do not convey the gender of a person. Two examples from grammatical gender languages support this view: inanimate nouns also have a grammatical gender without having a biological or social gender (see the "spoon" and "knife" examples above), and words exist where grammatical gender and biological/social gender mismatch (e.g., "das Mädchen" [the girl] in German, which is grammatically neuter, but describes a female person; Braun et al., 2005; Stahlberg et al., 2007).

Critique raised by feminist linguists calls the generic application of masculine generics into doubt, suggesting that women are not and do not feel represented as intended (Guentherodt et al., 1980). The proponents of this position argue that while for inanimate nouns the grammatical gender does obviously not express biological/social gender, a correspondence between grammatical and biological/social gender exists for most personal nouns (Guentherodt et al., 1980; Stahlberg et al., 2007). Feminist language critique emphasizes the ambiguity of masculine forms: sometimes 'men' and other masculine generics may mean men only and sometimes it may encompass men, women, and other genders. Or in other words, it is always correct to interpret a masculine generic as referring to male persons, whereas a reference

to women, groups of women and men, or people of other genders is only sometimes correct. Therefore, the association of masculine generics with male interpretations should be the strongest, resulting in a male bias of masculine generics (Braun et al., 2005).

Gygax and Gabriel (2010) as well as Irmen and Linner (2005) have further outlined the cognitive processes that could lead to women not being and feeling represented when masculine generics are used: When processing a linguistic input (e.g., "...the average student will feel isolated in his introductory courses" from Moulton et al., 1978, p. 1034), recipients build a mental model (Carreiras et al., 1996). These mental models contain explicitly stated information (e.g., studying person) and implicitly made inferences (e.g., person probably young as most students are young) based on the accessibility of the information. As the masculine generic can be interpreted in several ways (e.g., male person or gender unknown), the receiver of the message must solve the ambiguity of masculine generics by making inferences (Gygax & Gabriel, 2010; Irmen & Linner, 2005). These inferences contain mostly a concrete representation of a person (or group) with a specific gender (composition) instead of an abstractly represented person (or group) without a representation of gender (Hyde, 1984; Irmen & Linner, 2005). For making inferences, bottom-up (e.g., grammatical gender) as well as top-down (e.g., context) information is used. According to cognitive processing models (Gygax & Gabriel, 2010; Irmen & Linner, 2005), in the present example, without context information, the representation would likely be "male person" based on the bottom-up information of masculine grammatical gender.

Alternative Language Forms to Enhance the Mental Representation of Women

Based on the feminist language critique, different strategies have been proposed to enhance women's visibility in language or the gender inclusiveness of language (Lindqvist et al., 2019). The first strategy uses feminization to make the feminine equally as salient as the masculine. Feminization can be implemented for instance via word pairs ("he/she"; "las alum-

nas y los alumnos” [Spanish female and male students]; Lindqvist et al., 2019). Many languages have found ways to achieve feminization by moving beyond their own alphabet and implementing symbols to create new word forms. For instance, the middle dot is specific for French language (“musicien·ne·s” [French gender-inclusive form for musicians combining grammatical feminine and masculine form]; Xiao et al., 2022).

The second strategy uses neutralization to create gender neutrality in language by reducing the salience of gender (de-gendering). It is implemented via the usage of gender-neutral expressions (e.g., “chairperson” instead of “chairman”) and pronouns (e.g., singular “they” instead of “he”; Lindqvist et al., 2019). In German, for instance, gender neutrality can be implemented using neutral forms that neither contain male nor female suffixes (e.g., the use of participle constructions such as “Studierende” [German gender-inclusive form for students] instead of “Studenten” [German masculine generic form for students]; Stahlberg et al., 2007).

The third strategy, multi-gendering, highlights a non-binary understanding of gender (Lindqvist et al., 2019), for instance, by introducing new pronouns (e.g., “hen” [Swedish singular pronoun to refer to a non-binary person, but also as generic singular pronoun form]; Bäck et al., 2013; the singular “they” is nowadays used in an analogous manner in English; Lindqvist et al., 2019) or by using special characters such as the asterisk (“Professor*innen” [German gender-inclusive form for professors including grammatically male and female form with the star representing gender diversity]; Zacharski & Ferstl, 2022).

The Impact of Gendered Language on Perceivers and Actors: A Short Summary of Empirical Evidence

In the 1970s, scientists started to run tests on the controversy of whether masculine generics are understood generically or elicit a male bias in mental representations (e.g., Harrison, 1975; Schneider & Hacker, 1973). As the mental representation of women has been the major focus of research on gender-inclusive language so far, other variables that have been

studied as effects of masculine generic versus gender-inclusive language forms (e.g., readability, see online supplement for a short overview of variables) are not included in the present meta-analysis. Similarly, the mental representation of non-binary people or those of other genders is not included in the present meta-analysis, as only very recently researchers have started to investigate effects of language forms that are not only supposed to enhance women's mental representation but also strive to include gender identities outside the binary (Zacharski & Ferstl, 2022). Included in the present meta-analysis is research on the impact of gendered language on the mental representation of women, which comprises two perspectives: the perspective of perceivers (other-perspective) and the perspective of actors (self-perspective).

Regarding the *other*-perspective, numerous studies have covered perceivers' mental construal and its outcomes, specifically the mental representations men have of women, and women have of other women depending on the use of masculine generics versus gender-inclusive alternatives. In their seminal experimental study, Schneider and Hacker (1973) presented fictional headlines for textbook chapters to their participants. The headlines were either depicted in a generic male form (e.g., "Societal man") or a gender-inclusive form (e.g., "Society"). Participants' task was to choose a picture that fits the headline. Women were more represented in pictures chosen by men and women when a gender-inclusive form was presented than when a generic male form was presented (Schneider & Hacker, 1973). Many of the ensuing research followed this classic paradigm and replicated a higher mental representation of women in reaction to gender-inclusive language versus gender-inclusive language while using different manipulations of the independent variables and different measures of the dependent variables. For example, Stahlberg and Szesny (2001) presented in a study conducted in German profession categories (e.g., singers, athletes, politicians) either in the masculine generic form or via female-male word pairs and asked the participants to name three famous people for each profession. More women were named when profession categories were presented via

word pairs than via masculine generics (Stahlberg & Sczesny, 2001). Measures in this line of research range from naming people and/or writing stories (e.g., Keith et al., 2022), drawing or selecting pictures of women or men (e.g., Snizek & Jazwinski, 1986), estimating the gender of target individuals/groups (e.g., Horvath et al., 2016), responding whether a sentence can refer to a woman (e.g., Körner et al., 2022), success expectations of the person described (e.g., Vervecken et al., 2013), and making associations (e.g., Scheele & Gauler, 1993).

In addition to those explicit measures of perceivers' mental representations of women, research has employed implicit measures such as reaction times. Kim et al. (2022), for example, presented alternately male and female names in combination with professions in French and measured how long participants needed to respond whether an individual with the presented name could be a member of the respective profession. When professions were presented in a generic male form, participants responded faster to male compared to female names. When profession categories were presented in a gender-neutral form participants responded slightly faster to female compared to male names (Kim et al., 2022). Most of the studies using implicit measures used a study design comparing reaction times for female and male stimuli (e.g., names or pictures) testing whether masculine generics lead to slower reaction times compared to gender-inclusive alternatives when female (vs. male) stimuli were presented. As we investigate mental representations of women in reaction to masculine generics versus gender-inclusive language, only reaction times to female stimuli are included in the main analysis reported below. Additional analyses for the implicit measures are reported in the online supplement (i.e., comparing reaction times to female and male stimuli depending on language form as well as a separate meta-analysis on the Stimulus Gender \times Language Form interaction effect).

Regarding the *self*-perspective, several studies have investigated the difference between masculine generics and gender-inclusive language on whether women feel or think themselves to be addressed by a text (e.g., a job advertisement). In one of the early studies,

MacKay presented a text either containing the masculine generic “he” or containing the gender-inclusive “they”. Women evaluated the text written in masculine generics as less relevant for themselves than the text written in gender-inclusive “they”. However, in a direct replication, this effect was reversed (MacKay, 1980). More recently, Stout and Dasgupta (2011) found that women’s sense of belonging during job interviews was higher when recruiters used gender-inclusive language (“he or she”, “one”) compared to masculine generics (“he”). Further research replicated the language form effect again studying diverse dependent variables, such as sense of belonging, job interest (e.g., Vervecken et al., 2013), or success expectations (e.g., Hentschel et al., 2018b). To test whether women’s mental representations of themselves is higher when gender-inclusive language is used compared to masculine generics, the main analysis reported below includes only the self-perspective of *female participants*. However, a critical comparison from the self-perspective is to examine whether the degree of mental representation differs for women and men. If masculine generics do not adequately serve as a true generic form and rather promote mental representations of men, the effect of language form should be stronger for female than male participants from the self-perspective. In other words, male participants should feel addressed by both masculine generics and gender-inclusive alternatives as both include mental representations of men, whereas female participants should feel more addressed by gender-inclusive alternatives compared to masculine generics as the first promote the mental accessibility of women (Vervecken et al., 2013). If the language form effect is comparable for female and male participants (or even stronger for male participants), this would not speak against a generic interpretation of masculine generics. We address this in two additional analyses reported below (i.e., comparing the language form effect from the self-perspective for female and male participants as well as a separate meta-analysis on the Participant Gender \times Language Form interaction effect).

An increasing number of empirical studies from both self- and other-perspective, using explicit as well as implicit measures, have meanwhile reported on the mental representation of

gender in reaction to masculine generics versus gender-inclusive language. Diverse research designs have been applied to examine the impact of masculine generics versus gender-inclusive language, indicating that although masculine generics were intended to be generic, they oftentimes result in a lower mental representation of women compared to gender-inclusive language forms (= *language form effect*). This was corroborated in literature reviews (e.g., Irmen & Linner, 2005; Stahlberg et al., 2007), yet a systematic analysis of the available research is still lacking. We close this gap by providing the first meta-analysis of the available gender-related language form effects on the mental representation of women. Based on the existing literature, we expect a higher mental representation of women in reaction to gender-inclusive language forms compared to masculine generics from both other- and self-perspectives. Beyond investigating the quantity of the language form effect, we also examine its boundary conditions, such as gender typicality of context and publication year.

Potential Moderators of the Strength of the Language Form Effect

This meta-analysis examines several potential moderators of the language form effect, namely characteristics of the studies and of the participants. Moderator analyses are exploratory as the availability effect sizes greater than ten required to conduct the analyses (Fernández-Castilla et al., 2020) was not predictable at the beginning of this research project. Despite the exploratory nature of the analyses, assumptions regarding the effects of the respective moderators are discussed in the following.

Study Characteristics

Gender Typicality of Context. Gygax and Gabriel (2010) and Irmen and Linner (2005) outlined in their theoretical model that the bottom-up inference “male person” from masculine generics can be overridden by top-down influences. Chances are high that the “average *psychology* student feeling isolated in his introductory courses” is interpreted as mixed genders being addressed or even as being female due to knowledge about the large numbers of female psychology students (base-rate; Olos & Hoff, 2006). Correspondingly, Stahlberg et

al. (2007) hypothesize that in female-typed contexts the language form effect would be smaller compared to gender-neutral and male-typed contexts, as in female-typed contexts knowledge that women have a high base-rate overrides the bottom-up masculine interpretation of masculine generics. However, in recent years gender-inclusive forms have become more frequent (Adler & Hansen, 2020; Gustaffson Sendén et al., 2015) as well as counter-stereotypical role models (e.g., male nurses; Hussein et al., 2016). Thus, masculine generics in a female context could nowadays be interpreted as exclusively referring to males. If this mechanism was at work, the language form effect would have become larger in female-typed contexts compared to gender-neutral and male-typed contexts over time. We therefore not only tested the moderating role of context but also whether the size of the language form effect in the different contexts has changed over time.

Language. As outlined above, gender is encoded in all language systems (Hellinger & Bußmann 2001; Stahlberg et al., 2007), but to varying degrees. Here we focus solely on the distinction between grammatical gender languages and natural gender languages, as in both language systems masculine generics exist.

Again, contrasting assumptions regarding the effect of language can be formulated. On the one hand, one may assume that the language form effect should be larger for grammatical gender languages as in these languages gender is expressed not only via pronouns but also via the gender of nouns. The gender prime should therefore be more influential in comparison to natural gender languages. However, the same argument can also lead to the opposite assumption: Since inanimate objects (fork, spoon) have a grammatical gender in grammatical gender languages, people may be more used to disentangle grammatical and biological gender, and extend this to the case of social stimuli. First empirical results by Gygax et al. (2008) speak for the first assumption. For natural gender languages, they found that the specific male interpretation of masculine generics was overruled by a stereotypical interpretation, whereas for

grammatical gender languages this was not the case. This allows the conclusion that the language form effect is more robust in grammatical gender languages than in natural gender languages.

Language Form. As described above, the gender-inclusive alternatives comprise neutralization (e.g., they, human beings), feminization (e.g., pair form, capital-I, middle dot), and – only recently – multi-gendering (e.g., hen, gender asterisk *) strategies. Regarding their efficiency in enhancing the mental representation of women, past research hints towards larger language form effects for feminization compared to neutralization strategies (Stahlberg et al., 2007). This can be explained by the explicit reference to the female gender in feminine language forms. However, it remains an open question how large language form effects are for the multi-gendering strategy with respect to female representation, which is rather new.

For masculine generics, different forms have been used, for example expressions containing “man” (e.g., Sniezek & Jazwinski, 1986), as well as role nouns and pronouns in the grammatically male form (e.g., Stout & Dasgupta, 2011). Presumably, the language form effect may be larger for man-expressions as the most explicit reference to the mental representation of a male person.

Singular Versus Plural. Singular masculine generics are less likely to be interpreted as generic expressions in comparison to plural forms, as for plural forms increase the likelihood of including women into the mental representation of a less personalized and more abstract group. Therefore, the effects of language form may be stronger in singular than plural forms (Irmén & Linner, 2005).

Publication Year. Due to the lively public debate on masculine generics versus gender-inclusive language, masculine generics have been used less frequently in recent years, at least in public language contributions (Adler & Hansen, 2020). Consequently, this language form may be further losing its generic properties over time, and may be interpreted more in the sense of men only (Adler & Hansen, 2020; Gustaffson Sendén et al., 2015). However, the

discussion of gender-inclusive language may also have increased the sensitivity to the risk of not including women mentally in reaction to masculine generics. This process would lead to the opposite prediction of a weaker language form effect in recent research.

Nation-Level Gender Equality. Grammatical gender has been found to shape perception, judgement, decision making, and behavior (Prewitt-Freilino et al., 2012; Stahlberg et al., 2007). It has been theorized that a higher degree of gendering in languages should be associated with lower societal gender equality as a distinction between genders is very present in those languages (Prewitt-Freilino et al., 2012). Indeed, past studies have found that the more gendered languages were, the lower women's participation in the labor force, the more traditional views of gender roles (Jakiela & Ozier, 2018), and the lower the nations' level of achieved gender equality (Prewitt-Freilino et al., 2012).

Between- Versus Within-Participants Design. Between-participants and within-participants designs may differ in the size of language form effects. On the one hand, between-participants designs may elicit stronger language form effects than within-participants designs as the comparison group is unknown by the participants and therefore socially desirable responding is less likely. On the other hand, within-participants designs may elicit stronger effects. By having both language forms available, masculine generics could be interpreted as especially masculine (if not in the beginning, then over the course of the experiment, resulting in smaller differences in the aggregate; Gygax & Gabriel, 2008).

Participant Characteristics

Participant Gender in the Other-Perspective. In the other-perspective, all genders are expected to mentally represent women less often than men when reading about other people in masculine generic language forms (e.g., reading about the work of firemen vs. firefighters). We exploratorily test whether an interaction with participant gender emerges beyond this main effect of language form.

Participant Gender in the Self-Perspective. As described above, a critical comparison from the self-perspective is to examine whether the language form effect differs for women and men. Women should feel less mentally represented by masculine generics compared to gender-inclusive language, whereas men should feel addressed by both language forms similarly. We therefore tested not only the main effect of language form for women only (that was part of the main analysis) but also report two additional analyses. First, we test in a separate data set that includes only the self-perspective, but in addition to female participants also male participants whether the language form effect from the self-perspective is stronger for female than male participants. Second, we report the results of a meta-analysis testing whether the Participants Gender \times Language Form interaction effect emerges robustly across studies.

Participant Age. The language form effect has been investigated in children as well as adults and has been discussed to vary in its strength with age (Switzer, 1990). With development from childhood to adulthood, thinking becomes more abstract (Marini & Case, 1994). This would make it easier to interpret masculine generics as generics potentially leading to a smaller language form effect with increased age. However, mental representations in general could become more inclusive (also in reaction to gender-inclusive language forms) leading to higher female interpretations with increasing age in both language forms (Switzer, 1990). We therefore tested exploratorily whether, and if so how, the language form effect varies with participants' age.

Artifact Characteristics

The public debate on the research on the language form effect has been highly controversial in the past. For example, many studies on gender-inclusive language have been criticized for only using a single sentence while more context is needed to interpret masculine generics correctly as generics (Kurfer, 2022). To test the robustness of the language form effect, we additionally tested relevant moderators, such as length of the stimulus material. Detailed

descriptions and analyses can be found in the online supplement. We also coded publication status of the studies as one of several tests for publication bias (see sensitivity analyses in the Results section below).

Relevance of the Present Meta-Analysis

The mental representation of gender in masculine generics has been discussed controversially on a theoretical as well as empirical level. This question has gained high interest not only in the scientific community, but also in politics, media, and the society at large. For example, in a New York Times article from 1981, McFadden calls for more elegant gender-inclusive forms than using a slash (as in he/she) since the slash can complicate language and may be distracting (McFadden, 1981). Recent media reports from Germany cover political, legal, and societal discussions, for example, whether it is lawful to take points in a university exam for not using gender-inclusive language (Schmoll, 2022; Zoske, 2021). Also recently, in France the introduction of a gender-inclusive pronoun “iel” has been discussed controversially. Proponents have advocated for the inclusive mental representations promoted by “iel”, opponents have criticized such language reforms as unnecessary with reference to the generic function of masculine generics (Driyef, 2021).

This meta-analysis bears highly relevant implications for these ongoing debates. First, it will inform politics, media, and broader society about accumulated scientific knowledge of the effect of language on people’s mental representations of gender. Second, the moderation analyses will reveal, for example, whether the strength of effects differ for languages (see the discussion of languages with grammatical or natural gender above) or whether different gender-inclusive forms (e.g., neutralization or feminization) show difference in the size of their effects. Learning about the boundary conditions will allow theoretical accounts to be refined and guidelines for gender-inclusive language to be updated. For example, companies can use the accumulated knowledge in their hiring and communication strategies (e.g., language in job advertisements). Third, as gender-related language forms predict gender prejudices (DeFranza

et al., 2020) and nation-level gender inequality (Jakiela & Ozier, 2018; Prewitt-Freilino et al., 2012), the present investigation of language is meaningful beyond simply studying mental representations, as it bears weight for these highly relevant, ongoing challenges for society.

Method

Literature Search

The goal of the present meta-analysis was to compare gender-inclusive language with masculine generics on women's mental representation. We therefore restricted our search to those terms: (language OR linguist*) AND (gender* OR sex* OR masculin* OR femini* OR woman OR men) AND (fair OR bias* OR neutral OR generic* OR grammatical gender OR inclusi* OR stereotyp* OR equal*) AND (cognit* OR representati* OR visibil* OR evaluat* OR belief OR perception OR occupation OR availab* OR attitud*). We used five databases to locate relevant studies: PsychInfo, PSYINDEX, PsychArticles, Web of Science Core Collection, and ProQuest Dissertation and Theses Global (last search date December 23rd, 2020). Those databases were chosen based on their common inclusion in meta-analyses in Social Psychology. In addition, we published a call for (unpublished) data on September 10th, 2020, to minimize the file drawer problem (Rosenthal, 1979). Further, we cross-checked references in book chapters, reviews, and screened articles, and used the "cited by" function in Google Scholar to complement our comprehensive literature search.

After removing duplicates using Mendeley (www.mendeley.com), we used Abstrackr (Wallace et al., 2012) for screening 6535 titles and abstracts (see Prisma flow chart in Figure 1 for detailed numbers; Moher et al., 2009). The first 100 abstracts were screened by all four screeners (the two first authors and two research assistants) as pilots. After discussing disagreements and refining the inclusion and exclusion criteria, we switched to double-screening mode (i.e., each abstract was screened by two randomly selected screeners). Abstrackr uses a machine learning algorithm that sorts the abstracts according to their relevance based on previous screening decisions. When Abstrackr indicated that 0% of the remaining articles (about

2600) were likely to be relevant, we switched to single-screening mode (i.e., each abstract was screened by one screener).

Study Selection

After all abstracts were screened, the first two authors examined all 201 records that at least one screener indicated to be relevant for inclusion and decided together upon their eligibility for the meta-analysis. Disagreement about the inclusion or exclusion of studies was resolved through discussion with the last author. With respect to the independent variables, we included all studies containing an experimentally manipulated comparison of masculine generics and one or more gender-inclusive language form(s). Since the criterion for the independent variable was an experimental approach, we excluded corpus and language analyses. We also excluded studies analyzing (grammatical) congruency or match/mismatch effects (e.g., reaction times in response to stereotypic matching pairs “doctor-he” and mismatching pairs “doctor-she”; Banaji & Hardin, 1996) or comparing grammatical male versus female forms (e.g., Formanowicz & Sczesny, 2016). Further, we excluded studies that did not contain a masculine generic condition or contained language manipulations that were confounded with other manipulations (e.g., masculine generic + sexist wording vs. gender-inclusive language + non-sexist wording; Bem & Bem, 1973). With respect to the dependent variables, we included all studies measuring female representation as the dependent variable (see also the TOPICS+M criteria presented in Table 1 for detailed information on inclusion criteria; Johnson & Hennessy, 2019).

We only included articles published in German and English language (according to the authors’ language abilities). Bachelor’s or Master’s theses were only included when an intention to publish was declared to ensure that supervisors had checked methods and results for accurateness.

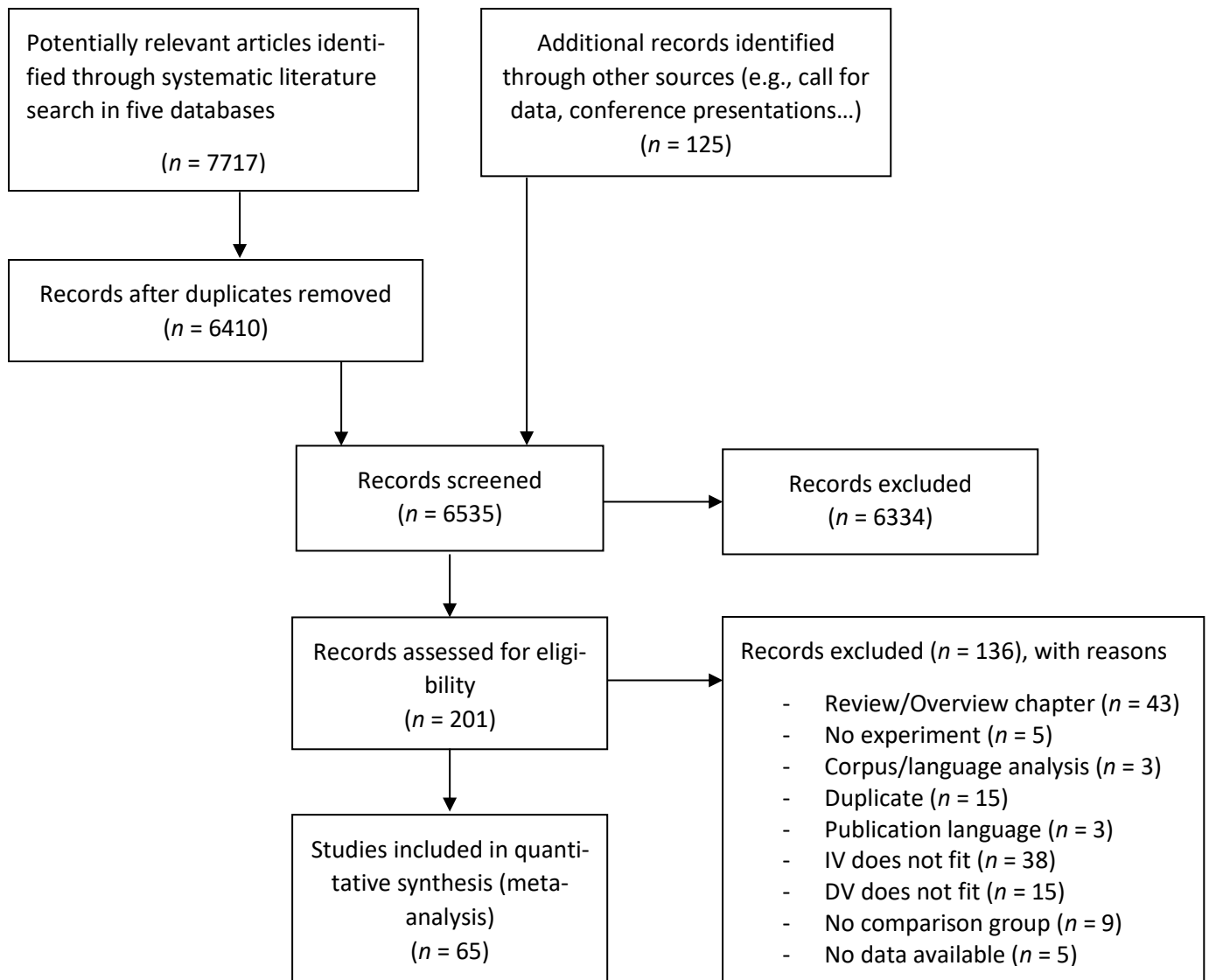
Figure 1*PRISMA Flow Chart of Literature Search*

Table 1

Inclusion Criteria for the Present Meta-Analysis according to the TOPICS+M Criteria

Criterion	Explanation
Time	Studies of all publication years are of interest; short-term effects (i.e., manipulation of independent variables and measurement of dependent variables within the same experimental session)
Outcome	(a) other-perspective (mental representations men have of women and women have of other women): - naming persons and/or writing stories - drawing or selecting pictures - estimating the gender of target persons/groups - responding whether a sentence can refer to a woman - success expectations - drawing associations - reaction times (b) self-perspective (mental representations female participants have of themselves): - sense of belonging - job interest - success expectations
Population	humans
Intervention	Gender-inclusive language form(s), e.g. - Feminization (e.g., pair form) - Neutralization (e.g., neutral expressions) - Multi-gendering (e.g., Gender *, neologism)
Comparison	Masculine generic language form(s), e.g., -man, masculine generic nouns and pronouns
Study design	experimental between- and within-subjects designs comparing masculine generics with gender-inclusive language forms on female representation
Exploratory Moderators	Self-other-perspective, gender typicality of context, language, language form, singular versus plural, publication year, nation-level gender equality, between- versus within-participants design, participant gender in the other-perspective, participant gender in the self-perspective, participant age, artifact characteristics

Coding

All studies were coded by one of four trained raters (second and third author and two research assistants) and double-checked by the first author. Disagreements were solved via discussion between raters. If data were missing, we contacted the first author of the respective article up to two times. The final set of data included in the main analysis of the meta-analysis

was 65, containing $m = 95$ independent samples, and $k = 357$ effect sizes, from $N = 19,582$ participants. For most samples we retrieved multiple effect sizes as we coded separate effect sizes, if available, for female and male participants from the other-perspective, gender typicality of contexts, multiple dependent variables, and multiple masculine generics and/or gender-inclusive language form comparisons.

Coding of the Moderators

Gender Typicality of Context. Context was coded either as female-typed, male-typed, gender-neutral context or as not available.

Language. For each effect size, we coded whether the language used in the experiment was a grammatical or natural gender language. To determine the language category, we used the classification of Prewitt-Freilino et al. (2012). We also examined whether the language form effect in grammatical and natural gender languages changes over time to rule out that change over time drives a potential difference between grammatical and natural gender languages as all studies conducted before 1992 (19 articles) were in English.

Language Form. We coded gender-inclusive language forms as neutralization (e.g., they, human beings), feminization (e.g., pair form, capital-I, middle dot), neutralization plus feminization used, multi-gendering (e.g., hen, gender asterisk *), and other (e.g., pair form alternated with masculine generics or a footnote explaining that all genders are meant to be included in masculine generics). Masculine generics were coded as generic man, generic nouns, generic pronouns, and combinations of those.

Singular Versus Plural. For each effect size, we coded whether it contained singular, plural, or mixed language forms.

Publication Year. To include publication year in the meta-analysis we assigned the oldest study from 1973 a value of 0 ascending to the newest studies from 2022 a value of 49.

Nation-Level Gender Equality. For each sample, we coded the nation-level gender equality using the Global Gender Gap Index (World Economic Forum, 2022).

Between- Versus Within-Participants Designs. We coded whether the manipulation of language form was between- or within-participants for each effect size.

Participant Gender in the Other-Perspective. For each effect size from the other-perspective, we coded whether the effect size contained data from female participants only, male participants only, or all participants from the study combined (if separate data were not available).

Participant Gender in the Self-Perspective. In the main analyses reported below (summarizing across other- and self-perspective), we only include the self-perspective of *female participants* to test whether overall (i.e., from both perspectives) women's mental representation is higher when gender-inclusive language is used compared to masculine generics. For the additional analyses on participant gender in the self-perspective we created two separate data sets. In one data set we coded for each effect size from the self-perspective whether the effect size contained data from female participants only or male participants only. In the other data set we coded for each study reporting data from the self-perspective the *d* effect size for the Participant Gender \times Language Form interaction effect itself.

Participant Age. We coded age of participants as school children, students, adults, mixed age groups, and age information not available. For those studies for which means or medians of participants' age were reported, we also coded age as continuous variable.

Publication Status. For publication status we coded whether the study was published with peer review, published without peer review, unpublished, or publication information was not available.

Effect Size Calculation

Language form was the dichotomous independent variable and female mental representation was the continuous dependent variable. For each comparison between masculine generics and gender-inclusive language form we retrieved the effect size Cohen's *d*. Positive values of *d* reflect more and negative values reflect less female representation in the gender-

inclusive compared to the masculine generic condition. If not reported in the original study, we computed d from descriptive statistics (means and information on variability, proportions), from test statistics (F tests, t tests, regression coefficients), or other effect sizes (η_p^2 , OR). For each effect size, we coded the degree of estimation necessary to calculate the effect ranging from 0 = *no estimation* (e.g., d given, or direct calculation of d from means, standard deviations, and cell sizes) to 4 = *highly estimated* (e.g., estimation of d using the reported means and an estimated pooled standard deviation stemming from an F -statistic with two numerator degrees of freedom). The degree of estimation in effect size calculation did not moderate the language form effect, providing a robustness check (see online supplement).

Transparency and Openness

We adhered to the MARS guidelines for meta-analytic reporting (Appelbaum et al., 2018). All meta-analytic data, analysis code, and research materials (including our coding scheme) are available at https://osf.io/ducgp/?view_only=54d7129820454a9488fedceca40d3b72. Data were analyzed following the guide from Harrer et al. (2021) using R, version 4.0.3 (R Core Team, 2020) and the packages meta (version 4.19.1; Balduzzi et al., 2019), Matrix (version 1.2.18; Bates & Maechler, 2019), effectsize (version 0.6.0.1; Ben-Shachar et al., 2020), pwr (version 1.3.0; Champely, 2020), powerAnalysis (version 0.2.1; Fan, 2017), robumeta (version 2.0; Fisher et al., 2017), dmetar (version 0.0.9000; Harrer et al., 2019), purrr (version 0.3.4; Henry & Wickham, 2020), esc (version 0.5.1; Lüdtke, 2019), report (version 0.5.6; Makowski et al., 2023), tibble (version 3.1.4; Müller & Wickham, 2021), clubSandwich (version 0.5.8; Pustejovsky, 2022), compute.es (version 0.2.5; Del Re, 2013), netmeta (version 1.5.0; Rücker et al., 2021), metasens (version 0.6.0; Schwarzer et al., 2021), metafor (version 3.0.2; Viechtbauer, 2010), ggplot2 (version 3.3.5; Wickham, 2016), stringr (version 1.4.0; Wickham, 2019), forcats (version 0.5.1; Wickham, 2021), tidyr (version 1.1.3; Wickham, 2021), tidyverse (version 1.3.1; Wickham et al., 2019), readxl (version 1.3.1; Wickham & Bryan, 2019), dplyr (version 1.0.5;

Wickham et al., 2021), readr (version 1.4.0; Wickham & Hester, 2020), zoo (version 1.8.8; Zeileis & Grothendieck, 2005) and lme4 (version 0.9.38; Zeileis & Hothorn, 2002). The hypothesis, literature search string, inclusion criteria, and moderators of the meta-analysis were preregistered at https://osf.io/93np4/?view_only=344b97f397b848699c83626349bc4e90.

Statistical Analysis

Multilevel random-effects models with robust variance estimation were used to address the multilevel nature of our data (three levels: participants nested in effect sizes nested in study samples) in our analyses (Harrer et al., 2021). Effect sizes were nested in study samples as, for example, in several articles more than one gender-inclusive condition was compared to the same masculine generics condition or as multiple dependent variables were assessed in the same sample. Robust variance estimation accounts for correlated sampling errors within studies by adjusting standard errors and confidence intervals (Tipton, 2013). Random-effects models enabled the testing of our proposed moderators (Borenstein et al., 2009).

Results

The main goal of the current meta-analysis was to quantify the effect of gender-inclusive language compared to masculine generics on the mental representation of women. The 65 articles and data sets included in the meta-analysis were published between 1973 and 2022. Across the $m = 95$ samples the mean percent of female participants was 60%. Approximately 46% of the samples were student samples, 17% adults outside university, 17% children, 18% mixed samples, and for 2% of studies the sample makeup was not reported. Of the $k = 357$ effect sizes 273 captured the language form effect from the other-perspective. 60% of the effect sizes contained a feminization strategy as gender-inclusive form, 30% a neutralization strategy, 6% a multi-gendering strategy, 4% neutralization and feminization mixed, and 1% other strategies. Most effects sizes came from a between participants design (94%), 70% from grammatical gender languages. The following languages were included in the meta-analysis:

German (50% of effect sizes), *English* (27%), French (6%), Italian (5%), Spanish (5%), Hebrew (4%), Dutch (3%), *Swedish* (2%), *Norwegian* (0.3%, natural gender languages in italics).

Overall Effect of Language Form on Mental Representation of Women

Overall (i.e., across self- and other-perspective), the average language form effect was $d = 0.46$, $SE = 0.05$, 95% CI [0.36, 0.55], $t(90.1) = 9.54$, $p < .001$, $k = 357$, $m = 95$, indicating a higher mental inclusion of women for gender-inclusive forms compared to masculine generics. The effect was reliable for both studies using the other-perspective ($d = 0.49$, 95% CI [0.40, 0.59], $k = 273$) and those using the self-perspective ($d = 0.34$, 95% CI [0.18, 0.50], $k = 84$) although it was marginally stronger for studies using the other-perspective ($p < .10$). Across all included studies, effect sizes varied widely, ranging from -1.11 to 3.14 , most effects were positive (81%). These results are depicted in Figure 2 and Figure 3.

The observed effects were heterogeneous $Q(356) = 1742.73$, $p < .001$. The I^2 statistic indicated that 80% of the observed variation was due to between-study heterogeneity, which is conventionally interpreted as substantial heterogeneity (Harrer et al., 2021). This underlines the adequacy of choosing a random effects model and enables testing for moderation, which are reported in the following.

Exploratory Moderator Analyses

Gender Typicality of Context

The mean effect size of language form on the mental representation of women did not differ between contexts, $F(3, 353) = 1.77$, $p = .152$. Despite the non-significant moderation, we investigated the language form effect within the different contexts to examine descriptive trends. The language form effect was marginally stronger ($p < .10$) for male-typed contexts ($d = 0.62$, 95% CI [0.43, 0.80], $k = 47$) than female-typed ($d = 0.48$, 95% CI [0.29, 0.67], $k = 50$) and significantly stronger ($p < .05$) for male-typed contexts than gender-neutral contexts ($d = 0.46$, 95% CI [0.28, 0.63], $k = 65$). The language form effects in the female-typed and gender-

neutral context did not differ from each other ($p > .10$) and were comparable to the language form effect when context information was not available ($d = 0.45$, 95% CI [0.33, 0.56], $k = 195$).

In addition, we analyzed whether the size of the language form effect in the different contexts has changed over time. In male-typed contexts the language form effect has become smaller over time ($b = -0.05$, 95% CI [-0.09, 0.006]), whereas in female-typed ($b = 0.02$, 95% CI [-0.01, 0.06]), gender-neutral contexts ($b = -0.02$, 95% CI [-0.06, 0.01]), and when no context information was available ($b = -0.03$, 95% CI [-0.06, 0.006]) effect sizes did not differ depending on publication year (see plot with effect sizes by year and context in the online supplement).

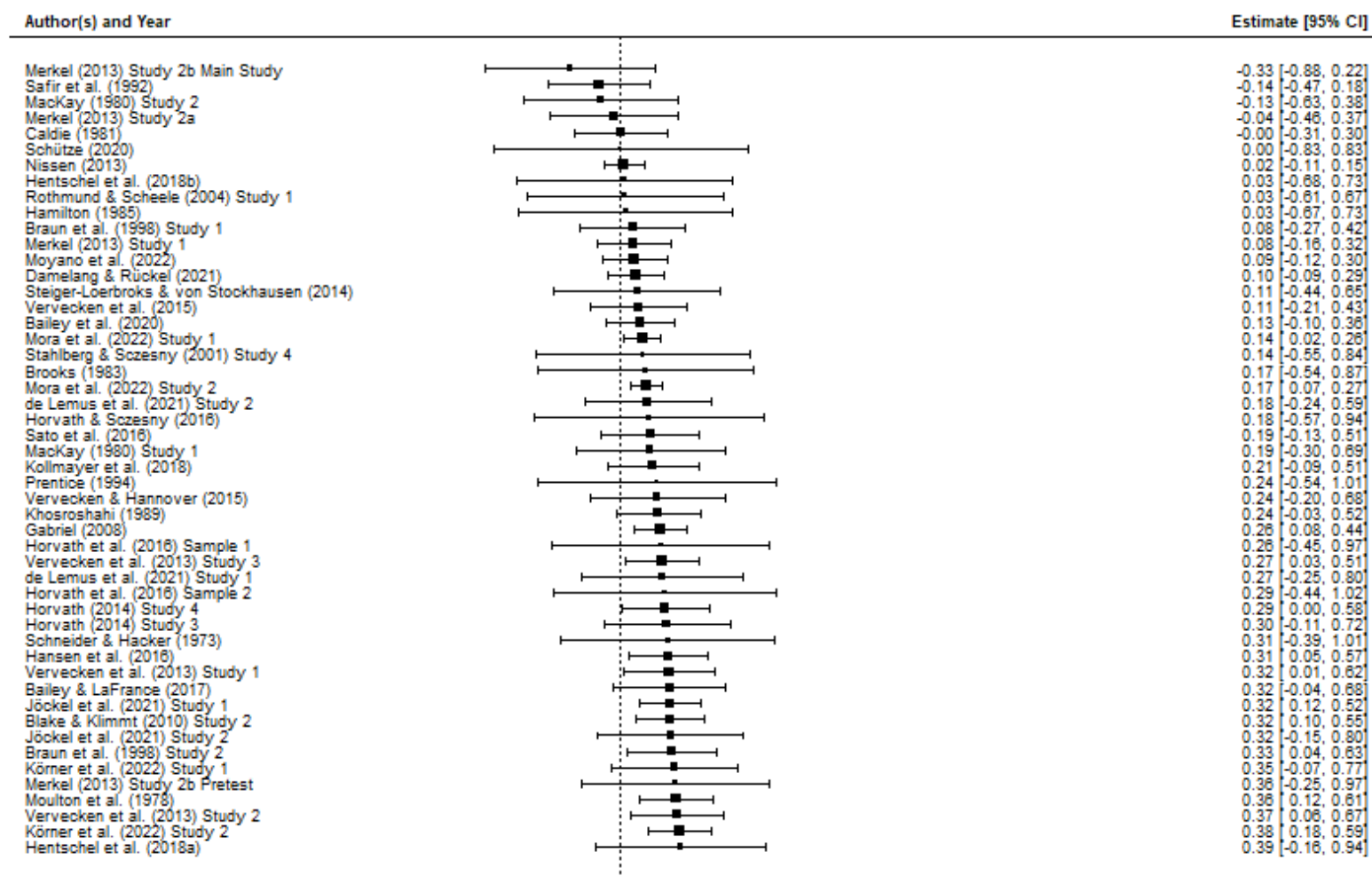
Language

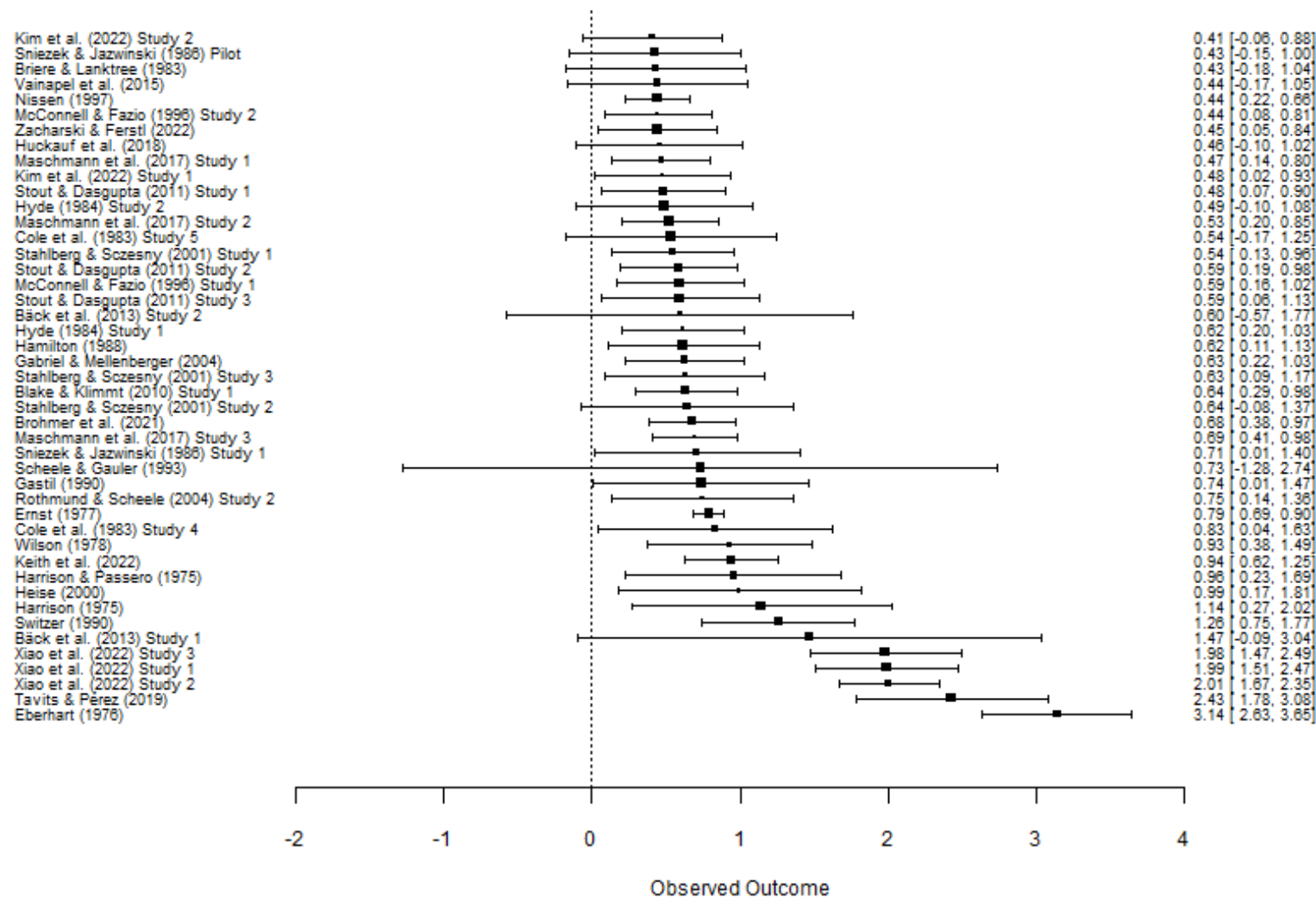
The mean effect size of language form on the mental representation of women did marginally differ between grammatical and natural gender languages, $F(1, 355) = 3.71$, $p = .055$. The difference in the mental representation of women between gender-inclusive and masculine generic forms was marginally stronger ($p < .10$) in natural gender languages ($d = 0.59$, 95% CI [0.43, 0.75], $k = 106$) than in grammatical gender languages ($d = 0.40$, 95% CI [0.28, 0.51], $k = 251$).

We also analyzed whether the language form effect in grammatical and natural gender languages has changed over years to rule out that changes over time drive a potential difference between grammatical and natural gender languages, as all studies conducted before 1992 (19 articles) were in English. Publication year did not moderate the language form effect, neither for natural ($b = -0.008$, 95% CI [-0.03, 0.01]) nor grammatical gender languages ($b = 0.007$, 95% CI [-0.007, 0.02], see plot with effect sizes by year and language in the online supplement).

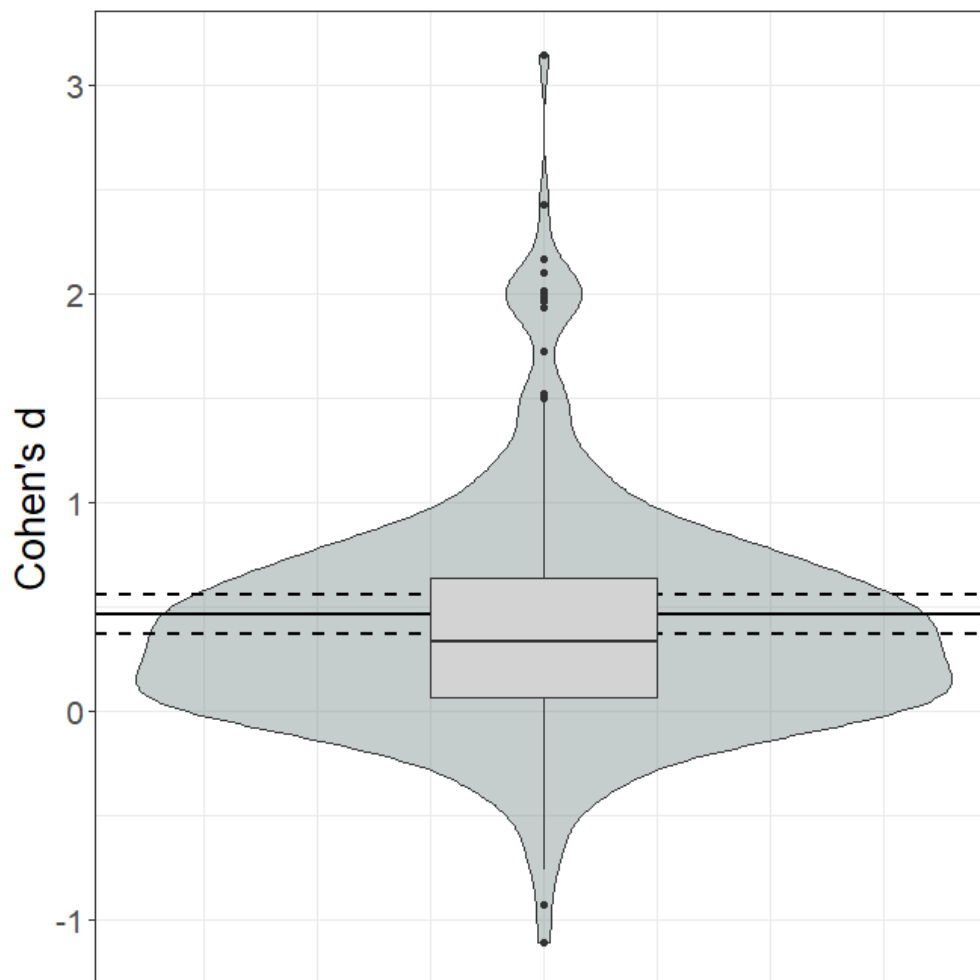
Figure 2

Forest Plot





Note. The forest plot shows the $k = 357$ effect sizes aggregated across the $m = 95$ samples.

Figure 3*Violin Plot*

Note. A higher mental representation of women in reaction to gender-inclusive language compared to masculine generics is reflected in more positive values of d . The solid line represents the mean effect size, and the dashed lines depict the corresponding 95% confidence interval. The width of the violin shows the distribution of the $k = 357$ effect sizes. In the center of the violin plot a boxplot of the effect sizes is depicted (the black points are outliers > 1.5 interquartile range).

Language Form

The mean effect size of language form on the mental representation of women differed between gender-inclusive language forms, $F(4, 352) = 2.71, p = .030$. The effect was largest for multi-gendering strategies ($d = 0.69$, 95% CI [0.49, 0.89], $k = 21$), followed by feminization ($d = 0.49$, 95% CI [0.38, 0.60], $k = 213$), feminization plus neutralization ($d = 0.41$, 95% CI [0.07, 0.74], $k = 15$), and neutralization ($d = 0.40$, 95% CI [0.27, 0.52], $k = 105$). Other strategies (e.g., footnotes) were not significant ($d = 0.06$, 95% CI [-0.54, 0.66], $k = 3$). The language form effect was significantly stronger in multi-gendering strategies compared to feminization ($p < .05$), neutralization ($p < .01$) and other strategies ($p < .05$). Comparing the language form effect between feminization and neutralization strategies revealed that the language form effect was marginally stronger ($p < .10$) in feminization compared to neutralization strategies.

For different masculine generics (such as “man” or “he”) no moderation of the language form effect appeared, $F(3, 353) = 1.53, p = .207$. Descriptively, the mean effect size was larger for generic man wordings ($d = 0.76$, 95% CI [0.45, 1.08], $k = 31$) compared to generic pronouns ($d = 0.52$, 95% CI [0.30, 0.73], $k = 58$), generic nouns ($d = 0.42$, 95% CI [0.30, 0.53], $k = 245$), and combinations of those ($d = 0.39$, 95% CI [-0.01, 0.79], $k = 23$). The language form effect was significantly stronger for generic man wordings than for generic nouns ($p < .05$; no other differences emerged in comparing the different masculine generics).

Singular Versus Plural

The language form effect was also not moderated by singular versus plural forms, $F(2, 354) = 1.32, p = .268$. The mean effect sizes for plural forms ($d = 0.53$, 95% CI [0.39, 0.67], $k = 150$), singular forms ($d = 0.45$, 95% CI [0.30, 0.59], $k = 162$), and mixed forms ($d = 0.30$, 95% CI [0.05, 0.55], $k = 45$) did not differ from each other ($ps > .10$).

Publication Year

Year of publication did not moderate the language form effect, $b = -0.003$, 95% CI $[-0.009, 0.003]$.

Nation-Level Gender Equality

To test whether the language form effect differs depending on nation-level gender equality, we included the global gender gap index as moderator in our meta-analytic model. As multiple samples were nested within nations, this analysis added a fourth level in the random effects model. No significant moderation emerged, $F(1, 355) = 2.30$, $p = .130$. Correspondingly, the effect of the global gender gap index on the language form effect was not significant, $b = 5.22$, 95% CI $[-1.55, 11.99]$.

Between- Versus Within-Participants Design

No moderation of the language form effect by design appeared, $F(1, 355) = 1.91$, $p = .167$, and the mean effect sizes of between-participants designs ($d = 0.49$, 95% CI $[0.38, 0.58]$, $k = 334$) and within-participants designs ($d = 0.26$, 95% CI $[-0.01, 0.56]$, $k = 23$) did not differ from each other ($p > .10$).

Participant Gender in the Other-Perspective

In studies using the other-perspective, the average language form effect was $d = 0.49$, $SE = 0.05$, 95% CI $[0.38, 0.60]$, $t(77.7) = 9.00$, $p < .001$, $k = 273$, $m = 82$. The moderation of participant gender was not significant, $F(2, 270) = 2.01$, $p = .136$. Descriptively, the mean effect size was largest for gender not differentiated ($d = 0.54$, 95% CI $[0.41, 0.65]$, $k = 164$), then male participants ($d = 0.48$, 95% CI $[0.26, 0.69]$, $k = 48$), followed by female participants ($d = 0.34$, 95% CI $[0.14, 0.54]$, $k = 61$). The language form effect did not differ between female and male participants ($p > .10$).

Participant Gender in the Self-Perspective

To analyze gender differences in the self-perspective, we used a separate data set that included the self-perspective of female and male participants (in contrast to the other analyses reported in the results section that included female participants only in the self-perspective).

With this separate data set, we tested whether the mental representation differed for female and male participants from the self-perspective, which should not be the case if masculine generics were understood generically. The crucial moderation of participant gender was significant, $F(1, 117) = 25.59, p < .001$, showing a significant language form effect for female participants ($d = 0.24$, 95% CI [0.15, 0.32], $k = 84$), whereas the language form effect was not significant for male participants ($d = -0.02$, 95% CI [-0.13, 0.09], $k = 35$). The language form effect for female participants was significantly larger than the language form effect for male participants ($p < .001$). In addition, we calculated a separate meta-analysis of the Participant Gender \times Language Form interaction effect on mental representations from the self-perspective. The results of a random effects model revealed a significant Participant Gender \times Language Form interaction effect, $d = 0.17$, 95% CI [0.06, 0.27], $k = 28$.

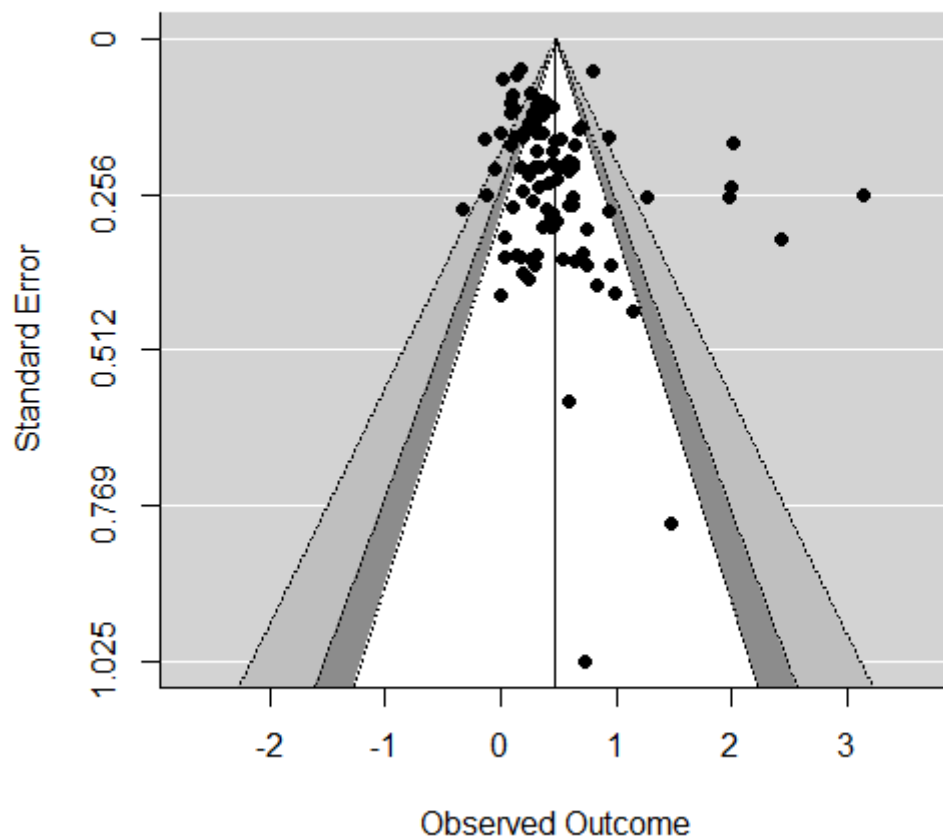
Participant Age

We tested whether the language form effect differs depending on the age group of participants in the respective studies. A significant difference between age categories emerged, $F(4, 352) = 2.48, p = .044$. The language form effect was strongest for adult samples ($d = 0.77$, 95% CI [0.54, 0.99], $k = 69$), followed by children ($d = 0.52$, 95% CI [0.30, 0.74], $k = 61$), students ($d = 0.39$, 95% CI [0.26, 0.53], $k = 180$), mixed age samples ($d = 0.34$, 95% CI [0.12, 0.56], $k = 45$), and not significant when age information was unavailable ($d = 0.23$, 95% CI [-0.47, 0.93], $k = 2$). The language form effect was significantly stronger ($p < .01$) for adult samples compared to students and mixed age samples. For those studies for which means or medians of participants' age were reported, we also tested age as continuous predictor. Participant age had a significant, positive effect on the language form effect, $b = 0.01$, 95% CI [0.0006, 0.03]. This means that with increasing mean age of a sample the difference in mental representation of women increases between masculine generics and gender inclusive forms.

Sensitivity Analyses

We conducted outlier analyses and identified two effect sizes (Eberhart, 1976; Tavits & Pérez, 2019) with studentized residuals > 1.96 as outliers. In the next step, we calculated Cook's distance to examine whether these two outliers were influential. As both outliers were also among the top ten influential cases, we repeated our main analysis excluding those two cases as robustness check. When excluding the outliers results were virtually identical, $d = 0.42$, $SE = 0.04$, 95% CI [0.33, 0.50], $t(90.1) = 10.30$, $p < .001$, $k = 355$, $m = 93$.

Besides outlier analyses, we tested for publication bias. 71% of the effect sizes stemmed from data published with peer review, 11% from data published without peer review, 14% from unpublished data, and for 4% of the effect sizes publication information was not available. We first tested publication status as moderator. This analysis revealed no significant moderation, $F(3, 353) = 1.05$, $p = .372$, and the language form effects of all publication statuses were significant ($ps < .03$). Next, we visually examined a funnel plot (Figure 4; for this and the following sensitivity analyses we discarded the multi-level structure by aggregating per sample). The visual inspection showed some asymmetry indicating that our results could be affected by publication bias. We used Egger's regression test to quantify funnel plot asymmetry. The result, $z = 2.06$, $p = .039$, indicated that the data were indeed asymmetric. However, in addition to publication bias, an asymmetric funnel plot can also be the result of extreme heterogeneity of effect sizes due to diverse study designs (Harrer et al., 2021). We therefore applied the three-parameter selection model (3PSM; McShane et al., 2016) to detect publication bias. This selection model is a generalization of other methods to detect publication bias (e.g., p -curve analysis). The non-significant test statistic, $\chi^2 = 0.36$, $p = .551$, indicates that our result was not substantially influenced by publication bias. Additional sensitivity analyses (e.g., p -curve analysis) can be found in the online supplement.

Figure 4*Funnel Plot*

Note. The white area denotes a significance area of $p < .01$, the dark grey area $p < .05$, and the outer light gray area $p < .10$. The three effect sizes with very large standard errors stem from odds ratios with treatment-arm correction due to zero cells (Scheele & Gauler, 1993; Bäck et al., 2013). Exclusion of these three samples did not change the main results.

Discussion

When exposed to masculine generics as opposed to gender-inclusive forms, how are women mentally represented by themselves and others? The present meta-analysis summarized research conducted between 1973 and 2022 that investigated the mental representation

of women in masculine generics compared to gender-inclusive language across diverse study designs from a self- as well as other-perspective. The key finding is that women are more mentally represented (other-perspective) and feel more represented (self-perspective) when gender-inclusive language is used compared to masculine generics (medium-sized effect, robust in sensitivity analyses).

This first meta-analytic evidence of the language form effect is in line with prior literature reviews describing gender-inclusive language as an effective way to increase the mental representations of women (e.g., Stahlberg et al., 2007): (1) Female and male participants show a higher mental representation of other women when being asked to draw a chairperson than when being asked to draw a chairman; (2) Female participants themselves show stronger job interest, success expectations, and feelings of belonging when gender-inclusive language forms are used compared to masculine generics, whereas male participants showed no such difference depending on language form. The results also imply that not only do others overlook, for example, female candidates for a position, but also female candidates themselves feel less addressed in reaction to masculine generics compared to gender-inclusive language forms. As such, masculine generics are not simply a grammatical device within languages, but carry a statistically significant male bias compared to more gender-inclusive language forms. The results of this meta-analysis therefore clearly support the proposition of feminist linguistics that masculine generics obscure female visibility. The validity of these findings seems not limited by publication bias or other methodological or artifact characteristics. Beyond this basic finding, we tested for moderating effects (e.g., context, language) to inform theory and practice about boundary conditions for the language form effect, described next.

Boundary Conditions for the Language Form Effect

The language form effect was significant for all three contexts (female-typed, male-typed, gender-neutral) and slightly larger for male-typed contexts. This descriptive trend is in line with the cognitive process model (Gygax & Gabriel, 2010; Irmen & Linner, 2005) as

male interpretations based on masculine generics are postulated to be more easily overruled by female-typed context information leading to smaller differences between gender-inclusive forms and masculine generics in female-typed contexts compared to male-typed contexts. As the language form effect was not weaker in female-typed compared to gender-neutral contexts, it seems that especially male-typed contexts elicit non-generic interpretations compared to gender-neutral and female-typed contexts. This allows the conclusion that male-typed, not female-typed contexts, drive context effects. Still, even in female-typed and gender-neutral contexts the mental representation of women was significantly and robustly higher when gender-inclusive language was used compared to masculine generics.

Moreover, the language form effect was stronger in natural gender languages compared to grammatical gender languages (which cannot be attributed to peculiarities of old studies). A potential explanation for a stronger language form effect in natural gender languages could be a consequence of the ease of avoiding gendered language in natural gender languages. As such, in natural gender contexts masculine generics may be interpreted less generically than in grammatical gender languages where the grammatical feature of masculine forms is more prominent (in social stimuli as well as in inanimate objects).

With regard to whether different forms of gender-inclusive language are equally effective in enhancing female representation, the meta-analysis revealed that multi-gendering inclusive forms were more successful at increasing female mental representations of women than feminization and neutralization strategies. In line with past literature reviews (Stahlberg et al., 2007) the language form effect was stronger when feminization compared to neutralization strategies were used, which can be explained by the explicit reference to the female gender in feminine language forms. The number of effect sizes investigating other gender-inclusive forms (e.g., a footnote explaining that alle genders are meant to be included in masculine generics) was very small, as those few initial studies found that such forms are ineffective in changing mental representations (Rothmund & Scheele, 2004). Regarding masculine generics

no moderation emerged, but especially expressions containing “man” seemed to trigger male biased interpretations compared to generic nouns.

Beyond those study characteristics the language form effect was robustly found for singular and plural forms, as well as between- and within-participant designs (which did not differ from each other). The language form effect neither varied by publication year nor by degree of achieved gender equality of nations.

Concerning the age of participants, the language form effect was found for adults, students, children, and for samples with mixed age groups. We expected that the language form effect would become smaller as abstract thinking increases and masculine generics might therefore be interpreted as more generic. Contrary to our expectation the language form effect became larger, not smaller, from childhood to adulthood. A post-hoc explanation might be that besides abstract thinking also attention to grammatical features of language, knowledge about stereotypes, and “people = men” biases could increase with higher age and more strongly influence gendered interpretations of language than abstract thinking.

Theoretical Implications

Overall, the present findings corroborate the feminist linguistic view that masculine generics reduce the mental representation of women compared to more gender-inclusive language forms. This is likely due to the ambivalent meaning of masculine generics: men only versus men and women alike. Beyond the large body of research on mental representations summarized in this meta-analysis, recent research has shown this using physiological measures (specifically event-related potentials; Misersky et al., 2019). Moderator analyses further supported the cognitive process model on the interpretation of masculine generics (Gygax & Gabriel, 2010; Irmen & Linner, 2005). The language form effect was descriptively stronger in male-typed contexts compared to female-typed contexts, which can be explained

by the model's proposed overwriting with context information (i.e., female-typed context information promotes mental representations of women when masculine generics are used, resulting in a smaller language form effect).

Further research should inspect divergences in the language form effect between grammatical and natural gender languages in more detail to learn more about the particularities of gendered interpretations in natural and grammatical gender languages. This would also allow predictions of how changes in language (e.g., through the increased usage of gender-inclusive language) translate into changes in mental representations of gender. In this regard, the Norwegian language could play a crucial role for future research, as Norwegian was formerly a grammatical gender language and by discarding grammatical female forms an attempt was made towards more diverse mental representations of gender in the present natural gender Norwegian language (Gabriel, 2008).

In addition, future research should also investigate the finding that the language form effect increases with age to better understand what mechanism drives this effect and followingly which approaches and features contribute to inclusiveness in language (e.g., attention to grammatical features of language, knowledge about stereotypes).

Another important, unsolved question concerns the mental representation of gender depending on language in terms of absolute values. In the present meta-analysis, we investigated relative values and compared whether the mental representation of women was higher or lower when masculine generics versus gender-inclusive language forms were used. A future meta-analysis should investigate absolute values of mental representation and examine whether gender-inclusive forms are powerful enough to increase the mental representation of women to a theoretically or empirically derived level (e.g., 50%, or more or less depending, for example, on the base rate) or even elicit a female bias. In combination with examining absolute values, future research should investigate base rates as a relevant moderator (Braun et al., 2005; Brohmer et al., 2022).

Practical Implications

Current research shows that everyday speech is biased towards men (Caliskan et al., 2022; MacArthur et al., 2020). If a goal is to increase the mental representation of women, based on the results of this meta-analysis, guidelines for gender-inclusive language should recommend multi-gendering language forms (e.g., hen, gender asterisk *) and discourage the use of wordings containing “man”. This can be relevant, for example, for companies in their pursuit to attract diverse talents and communicate inclusively, for political parties to reach diverse groups of voters, for news and marketing companies to speak to a diverse audience, and so forth. Prior research has shown that gender-inclusive language sometimes seems to reduce the perceived aesthetics of texts but is nevertheless very comprehensible content wise (Braun et al., 2007). Over time, habituation effects for aesthetics may occur parallel to higher acceptance rates of the changes the introduction of gender-inclusive language entails (as have been observed for “hen” in Swedish language; Gustafsson Sendén et al., 2021).

Limitations

A clear limitation of this meta-analysis is that only studies investigating the mental representation of women could be included as only very recently research has started to examine the mental representation of diverse genders. First results suggest that using multi-gendering language forms leads to inclusive mental representations including people beyond the gender binary (Merkel & Roessel, 2023; Zacharski & Ferstl, 2022). Further, only very few studies so far looked at spoken language, and only very recently on multi-gendering spoken forms (see Jöckel et al., 2021; Körner et al., 2022). While already being used in daily life, as discussions about German public news show (Payr, 2022), more research is needed to understand how such forms are perceived in spoken language and test their effectiveness in increasing female and diverse mental representations of gender.

Further, although this meta-analysis included research on nine languages (all research reports that were available in English and German), future research should endeavor to examine other languages, particularly those which use scripts other than Latin, or include sign languages, as their absence from the literature on this topic undermines a holistic understanding of gendered mental representations.

Conclusions

In this meta-analysis 357 effect sizes from 95 independent samples provided evidence for a significantly higher mental representation of women in gender-inclusive alternatives compared to masculine generics. Effects were very robust considering publication bias, outlier, and other sensitivity analyses. The difference between gender-inclusive forms and masculine generics in mental representations of women was largest for multi-gendering strategies (e.g., hen, gender asterisk *). Future research should endeavor to investigate the mental representation of people beyond the gender binary.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Adler, A., & Hansen, K. (2020). Studenten, StudentInnen, Studierende? Aktuelle Verwendungspräferenzen bei Personenbezeichnungen [Students (masculine generic form), students (combination of grammatical female and male form), students (neutral form)? Current usage preferences for personal names]. In J. A. Bär, R. Freudenberg-Findeisen, T. Menzel, D. Nübling, & P. Schlobinski (Eds.), *Muttersprache* (pp. 47–63). Gesellschaft für deutsche Sprache.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, 73(1), 3–25. <https://doi.org/10.1037/amp0000191>
- *Bäck, E. A., Lindqvist, A., & Gustafsson Sendén, M. (2013, October). *Hen can do it!: Effects of using a gender-neutral pronoun in recruitment* [Conference paper]. 8th Nordic Conference on Language and Gender, Iceland.
- *Bailey, A. H., & LaFrance, M. (2017). Who counts as human? Antecedents to androcentric behavior. *Sex Roles*, 76(11-12), 682–693. <https://doi.org/10.1007/s11199-016-0648-4>
- *Bailey, A. H., LaFrance, M., & Dovidio, J. F. (2020). Implicit androcentrism: Men are human, women are gendered. *Journal of Experimental Social Psychology*, 89, Article 103980. <https://doi.org/10.1016/j.jesp.2020.103980>
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evidence-based mental health*, 22(4), 153–160. <https://doi.org/10.1136/ebmental-2019-300117>

- Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, 7(3), 136-141. <https://doi.org/10.1111/j.1467-9280.1996.tb00346.x>
- Bates, D., Maechler, M. (2019). *Matrix: Sparse and dense matrix classes and methods*. R package version 1.2-18, <https://CRAN.R-project.org/package=Matrix>
- Bem, S. L., & Bem, D. J. (1973). Does sex-biased job advertising "aid and abet" sex discrimination? *Journal of Applied Social Psychology*, 3(1), 6–18. <https://doi.org/10.1111/j.1559-1816.1973.tb01290.x>
- Ben-Shachar, M. S., Lüdtke, D., Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, 5(56), 2815. <https://doi.org/10.21105/joss.02815>
- *Blake, C., & Klimmt, C. (2010). Geschlechtergerechte Formulierungen in Nachrichtentexten [Gender-fair wording in news texts]. *Publizistik*, 55(3), 289–304. <https://doi.org/10.1007/s11616-010-0093-2>
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.
- *Braun, F., Gottburgsen, A., Sczesny, S., & Stahlberg, D. (1998). Können Geophysiker Frauen sein? Generische Personenbezeichnungen im Deutschen [Can geophysicists be women? Generic personal names in German]. *Zeitschrift Für Germanistische Linguistik*, 26(3), 265–283. <https://doi.org/10.1515/zfgl.1998.26.3.265>
- Braun, F., Oelkers, S., Rogalski, K., Bosak, J., & Sczesny, S. (2007). "Aus Gründen der Verständlichkeit ...": Der Einfluss generisch maskuliner und alternativer Personenbezeichnungen auf die kognitive Verarbeitung von Texten ["For reasons of intelligibility ...": How masculine generics and alternative forms affect the cognitive processing of a

text]. *Psychologische Rundschau*, 58(3), 183–189. <https://doi.org/10.1026/0033-3042.58.3.183>

Braun, F., Szesny, S. & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, 30(1), 1–21. <https://doi.org/10.1515/comm.2005.30.1.1>

*Briere, J., & Lanktree, C. (1983). Sex-role related effects of sex bias in language. *Sex Roles*, 9(5), 625–632. <https://doi.org/10.1007/BF00290069>

*Brohmer, H., Hofer, G., Bauch, S. A., Beitner, J., Berkessel, J., Corcoran, K., Dinslaken, I. Garcia, D., Gruber, F. M., Jauk, E., Malkoc, S., Metzler, H., Mües, H. M., Otto, K., Rahal, R.-M., Salwender, M., Szesny, S., Stahlberg, D., Wehrt, W., Giuliani, F., Athenstaedt, U. (2022). *Effects of the generic masculine and its alternatives: A multi-lab replication and extension of Stahlberg, Szesny, and Braun, 2011* [Registered Report]. Institute of Psychology, University of Graz, Austria.

*Brooks, L. (1983). Sexist language in occupational information: Does it make a difference? *Journal of Vocational Behavior*, 23(2), 227–232. [https://doi.org/10.1016/0001-8791\(83\)90036-2](https://doi.org/10.1016/0001-8791(83)90036-2)

*Caldie, R. W. (1981). *Dominance and language: A new perspective on sexism*. University Press of America.

Caliskan, A., Ajay, P. P., Charlesworth, T., Wolfe, R., & Banaji, M. R. (2022, July). Gender bias in word embeddings: A comprehensive analysis of frequency, syntax, and semantics. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 156-170). <https://doi.org/10.1145/3514094.3534162>

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The*

- Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 49A(3), 639–663. <https://doi.org/10.1080/027249896392531>
- Champely, S. (2020). *pwr: Basic functions for power analysis*. R package version 1.3-0, <https://CRAN.R-project.org/package=pwr>
- *Cole, C. M., Hill, F. A., & Dayley, L. J. (1983). Do masculine pronouns used generically lead to thoughts of men? *Sex Roles*, 9(6), 737–750. <https://doi.org/10.1007/BF00289802>
- *Damelang, A., & Rückel, A.-K. (2021). Was hält Frauen von beruflichen Positionen fern? Ein faktorieller Survey zum Einfluss der Gestaltung einer Stellenausschreibung auf deren Attraktivitätseinschätzung [What keeps women away from professional positions? A factorial survey on the influence of the design of a job advertisement on its attractiveness assessment]. *KZfSS Kölner Zeitschrift Für Soziologie und Sozialpsychologie*, 73(1), 109–127. <https://doi.org/10.1007/s11577-021-00729-z>
- DeFranza, D., Mishra, H., & Mishra, A. (2020). How language shapes prejudice against women: An examination across 45 world languages. *Journal of Personality and Social Psychology*, 119(1), 7–22. <https://doi.org/10.1037/pspa0000188>
- Del Re, A. C. (2013). *compute.es: Compute effect sizes*. <https://cran.r-project.org/package=compute.es>
- *de Lemus, S., & Estevan-Reina, L. (2021). Influence of sexist language on motivation and feelings of ostracism (La influencia del lenguaje sexista en la motivación y el sentimiento de ostracismo). *International Journal of Social Psychology*, 36(1), 61–97. <https://doi.org/10.1080/02134748.2020.1840230>
- Driyef, Z. (2021, November, 26). De l'écriture inclusive à la transidentité, « iel » dans « Le Monde » [From inclusive writing to transidentity, "iel" in "Le Monde"]. *Le Monde*.

https://www.lemonde.fr/m-le-mag/article/2021/11/26/de-l-ecriture-inclusive-a-la-transidentite-iel-dans-le-monde_6103681_4500055.html

*Eberhart, O. M. Y. (1976). *Elementary students' understanding of certain masculine and neutral generic nouns* [Unpublished doctoral dissertation]. Kansas State University.

*Ernst, S. B. (1977). *An investigation of students' interpretation of inclusionary and exclusionary gender generic language* [Unpublished doctoral dissertation]. Washington State University.

Fan, F. Y. (2017). *powerAnalysis: Power analysis in experimental design*. R package version 0.2.1, <https://CRAN.R-project.org/package=powerAnalysis>

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*, 52(5), 2031–2052. <https://doi.org/10.3758/s13428-020-01373-9>

Fisher, Z., Tipton, E., Zhipeng, H. (2017). *robumeta: Robust variance meta-regression*. R package version 2.0, <https://CRAN.R-project.org/package=robumeta>

Formanowicz, M., & Szczesny, S. (2016). Gender-fair language and professional self-reference: the case of female psychologists in Polish. *Journal of Mixed Methods Research*, 10(1), 64–81. <https://doi.org/10.1177/1558689814550877>

*Gabriel, U. (2008). Language policies and in-group favoritism. *Social Psychology*, 39(2), 103–107. <https://doi.org/10.1027/1864-9335.39.2.103>

*Gabriel, U., & Mellenberger, F. (2004). Exchanging the generic masculine for gender-balanced forms – The impact of context valence. *Swiss Journal of Psychology*, 63(4), 273–278. <https://doi.org/10.1024/1421-0185.63.4.273>

- *Gastil, J. (1990). Generic pronouns and sexist language: The oxymoronic character of masculine generics. *Sex Roles*, 23(11-12), 629–643. <https://doi.org/10.1007/BF00289252>
- Guentherodt, I., Hellinger, M., Pusch, L. F., & Tromel-Plotz, S. (1980). Richtlinien zur Vermeidung sexistischen Sprachgebrauchs [Guidelines for avoiding sexist language use]. *Linguistische Berichte Braunschweig*, 69, 15–21.
- Gustafsson Sendén, M., Bäck, E. A., & Lindqvist, A. (2015). Introducing a gender-neutral pronoun in a natural gender language: The influence of time on attitudes and behavior. *Frontiers in Psychology*, 6, Article 893. <https://doi.org/10.3389/fpsyg.2015.00893>
- Gustafsson Sendén, M., Renström, E., & Lindqvist, A. (2021). Pronouns beyond the binary: The change of attitudes and use over time. *Gender & Society*, 35(4), 588–615. <https://doi.org/10.1177/08912432211029226>
- Gygax, P., & Gabriel, U. (2008). Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie*, 67(3), 143–151. <https://doi.org/10.1024/1421-0185.67.3.143>
- Gygax, P., & Gabriel, U. (2010). Gender representation in language: More than meets the eye. In R. M. Mishra & N. Srinivasan (Eds). *Language-cognition interface: State of the art*. Lincom Europa.
- Gygax, P., Gabriel, U., Sarrasin, O., Garnham, A. & Oakhill, J. (2008). There is no generic masculine in French and German: When beauticians, musicians and mechanics are all men. *Language and Cognitive Processes*, 23, 464–485.

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2009). Some grammatical rules are more difficult than others: The case of the generic interpretation of the masculine. *European journal of psychology of education*, 24(2), 235–246.

<https://doi.org/10.1007/BF03173014>

*Hamilton, M. C. (1985). *Linguistic relativity and sex bias in language: Effects of masculine "generics" on the imagery of the writer and the perceptual discrimination of the reader* [Unpublished doctoral dissertation]. University of California.

*Hamilton, M. C. (1988). Using masculine generics: Does generic he increase male bias in the user's imagery? *Sex Roles*, 19(11-12), 785–799.

<https://doi.org/10.1007/BF00288993>

*Hansen, K., Littwitz, C., & Sczesny, S. (2016). The social perception of heroes and murderers: Effects of gender-inclusive language in media reports. *Frontiers in Psychology*, 7, Article 369. <https://doi.org/10.3389/fpsyg.2016.00369>

Harrer, M., Cuijpers, P., Furukawa, T. A., Ebert, D. D. (2019). *dmetar: Companion R package for the guide 'Doing meta-analysis in R'*. R package version 0.0.9000, <http://dmetar.protectlab.org/>

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman & Hall/CRC Press.

*Harrison, L. (1975). Cro-magnon woman-in eclipse. *Science Teacher*, 42(4), 8-11.

*Harrison, L., & Passero, R. N. (1975). Sexism in the language of elementary school textbooks. *Science and Children*, 12(4), 22-25. <https://www.jstor.org/stable/43145994>

*Heise, E. (2000). Sind Frauen mitgemeint? Eine empirische Untersuchung zum Verständnis des generischen Maskulinums und seiner Alternativen [Are women included? An em-

pirical investigation into the understanding of the generic masculine and its alternatives]. *Sprache & Kognition*, 19(1/2), 3–13. <https://doi.org/10.1024//0253-4533.19.12.3>

Hellinger, M., & Bußmann, H. (Eds.). (2001). *Gender across languages: The linguistic representation of women and men*. John Benjamins Publishing Company.

Henry, L., Wickham, H. (2020). *purrr: Functional programming tools*. R package version 0.3.4, <https://CRAN.R-project.org/package=purrr>

*Hentschel, T., Horvath, L. K., Peus, C., & Sczesny, S. (2018a). Kick-starting female careers. *Journal of Personnel Psychology*, 17(4), 193–203. <https://doi.org/10.1027/1866-5888/a000209>

*Hentschel, T., Horvath, L. K., Peus, C., & Sczesny, S. (2018b). *Experimental investigation of women's job interest and success expectations* [Unpublished raw data]. University of Amsterdam, The Netherlands.

*Horvath, L. K. (2014). *Counteracting or contributing to gender equality? The impact of gender-fair language in the work context* [Unpublished doctoral dissertation]. University of Bern, Switzerland.

*Horvath, L. K., Merkel, E. F., Maass, A., & Sczesny, S. (2016). Does gender-fair language pay off? The social perception of professions from a cross-linguistic perspective. *Frontiers in Psychology*, 6, Article 2018. <https://doi.org/10.3389/fpsyg.2015.02018>

*Horvath, L. K., & Sczesny, S. (2016). Reducing women's lack of fit with leadership positions? Effects of the wording of job advertisements. *European Journal of Work and Organizational Psychology*, 25(2), 316–328. <https://doi.org/10.1080/1359432X.2015.1067611>

- *Huckauf, A., Hensel, L., Oberzaucher, E. & Ehlers, J. (2018) Okulomotorik und Lesefluss beim Lesen von Texten in geschlechterfairer Schreibung [Oculomotor skills and reading fluency when reading texts in gender-fair spelling]. In C. Schwender, S. Schwarz, B. P. Lange, & A. Huckauf (Eds.) *Geschlecht und Verhalten aus evolutionärer Perspektive* [Gender and behavior from an evolutionary perspective] (pp. 293–307). Pabst Science Publishers.
- Hussein, S., Ismail, M., & Manthorpe, J. (2016). Male workers in the female-dominated long-term care sector: Evidence from England. *Journal of Gender Studies*, 25(1), 35–49. <https://doi.org/10.1080/09589236.2014.887001>
- *Hyde, J. S. (1984). Children's understanding of sexist language. *Developmental Psychology*, 20(4), 697–706. <https://doi.org/10.1037/0012-1649.20.4.697>
- Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen [The representation of generic masculine personal names]. *Zeitschrift für Psychologie/Journal of Psychology*, 213(3), 167–175. <https://doi.org/10.1026/0044-3409.213.3.167>
- Jakiela, P., & Ozier, O. (2018). *Gendered language*. World Bank Policy Research Working Paper. <https://www.worldbank.org/en/news/feature/2019/01/24/gendered-languages-may-play-a-role-in-limiting-womens-opportunities-new-research-finds>
- *Jöckel, S., Dogruel, L., & Bachofer, R. (2021). Wirkung gendersensibler Ansprachen in Anmoderationen bei Erwachsenen und Heranwachsenden [Effects of gender-fair addressing for audiovisual journalistic media among adults and adolescents]. *Publizistik*, 66, 441–462. <https://doi.org/10.1007/s11616-021-00682-z>

- Johnson, B. T., & Hennessy, E. A. (2019). Systematic reviews and meta-analyses in the health sciences: Best practice methods for research syntheses. *Social Science & Medicine*, 233, 237–251. <https://doi.org/10.1016/j.socscimed.2019.05.035>
- *Keith, N., Hartwig, K., & Richter, T. (2022). Ladies first or ladies last: Do masculine generics evoke a reduced and later retrieval of female exemplars? *Collabra: Psychology*, 8(1), Article 32964. <https://doi.org/10.1525/collabra.32964>
- *Khosroshahi, F. (1989). Penguins don't care, but women do: A social identity analysis of a Whorfian problem. *Language in Society*, 18(4), 505–525. <https://doi.org/10.1017/S0047404500013889>
- *Kim, J., Angst, S., Gygax, P., Gabriel, U., & Zufferey, S. (2022). The masculine bias in fully gendered languages and ways to avoid it: A study on gender neutral forms in Québec and Swiss French. *Journal of French Language Studies*, 1–26. <https://doi.org/10.1017/S095926952200014X>
- *Körner, A., Abraham, B., Rummer, R., & Strack, F. (2022). Gender representations elicited by the gender star form. *Journal of Language and Social Psychology*, 41(5), 553–571. <https://doi.org/10.1177/0261927X221080181>
- Körner, A., Glim, S., & Rummer, R. (2022). Examining the glottal stop as a mark of gender-inclusive language. PsyArXiv. <https://doi.org/10.31234/osf.io/fr74p>
- *Kollmayer, M., Pfaffel, A., Schober, B., & Brandt, L. (2018). Breaking away from the male stereotype of a specialist: Gendered language affects performance in a thinking task. *Frontiers in Psychology*, 9, Article 985. <https://doi.org/10.3389/fpsyg.2018.00985>
- Kurfer, T. (2022, July 22). Streit ums Gendern: Nein, die deutsche Sprache diskriminiert Frauen nicht [Controversy over gendering: No, the German language does not discriminate against women]. *Berliner Zeitung*. <https://www.berliner-zeitung.de/open->

source/streit-ums-gendern-nein-die-deutsche-sprache-diskriminiert-frauen-nicht-li.246245

Lindqvist, A., Renström, E. A., & Gustafsson Sendén, M. (2019). Reducing a male bias in language? Establishing the efficiency of three different gender-fair language strategies. *Sex Roles: A Journal of Research*, 81(1-2), 109–117. <https://doi.org/10.1007/s11199-018-0974-9>

Lucy, J. A., (2001). Sapir-Whorf Hypothesis. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13486-13490). Pergamon Verlag. <https://doi.org/10.1016/B0-08-043076-7/03042-4>

Lüdecke, D. (2019). *esc: Effect size computation for meta analysis* (Version 0.5.1). <https://doi.org/10.5281/zenodo.1249218>

MacArthur, H. J., Cundiff, J. L., & Mehl, M. R. (2020). Estimating the prevalence of gender-biased language in undergraduates' everyday speech. *Sex Roles: A Journal of Research*, 82(1-2), 81–93. <https://doi.org/10.1007/s11199-019-01033-z>

*MacKay, D. G. (1980). Language, thought and social attitudes. In H. Giles (Ed.) *Language: Social psychological perspectives* (pp. 89–96). Pergamon Press.

Makowski, D., Lüdecke, D., Patil, I., Thériault, R. (2023). *Automated results reporting as a practical tool to improve reproducibility and methodological best practices adoption*. <https://easystats.github.io/report/>

Marini, Z., & Case, R. (1994). The development of abstract reasoning about the physical and social world. *Child Development*, 65(1), 147–159. <https://doi.org/10.2307/1131372>

*Maschmann, I. T., Fleischmann, A., & Schneider, I. (2017). *Three experiments on gender-inclusive language* [Unpublished raw data]. University of Cologne, Germany.

- *McConnell, A. R., & Fazio, R. H. (1996). Women as men and people: Effects of gender-marked language. *Personality & Social Psychology Bulletin*, 22(10), 1004–1013.
<https://doi.org/10.1177/01461672962210003>
- McFadden, C. (1981, August 02). On Language; IN DEFENSE OF GENDER. *New York Times*. <https://www.nytimes.com/1981/08/02/magazine/on-language-in-defense-of-gender.html?searchResultPosition=9>
- McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science*, 11(5), 730–749.
<https://doi.org/10.1177/1745691616662243>
- *Merkel, E. (2013). *The two faces of gender-fair language* [Unpublished doctoral dissertation]. University of Padua, Italy.
- Merkel, E.* & Roessel, J.* (2023). Getting gender in or out of mind – Why it is worth to mind gender in language [Accepted for publication]. [*shared first authorship]
- Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical gender in German influences how role-nouns are interpreted: Evidence from ERPs. *Discourse Processes*, 56(8), 643–654. <https://doi.org/10.1080/0163853X.2018.1541382>
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7), Article e1000097. <https://doi.org/10.1371/journal.pmed1000097>
- *Mora, Y. L., Smeding, A., Quinton, J.-C., Klein, O. (2022). *The small dot challenge: Influence of gender-fair language on representations and stereotypes of science and care jobs* [Manuscript in preparation]. University of Bruxelles, Belgium.

- *Moulton, J., Robinson, G. M., & Elias, C. (1978). Sex bias in language use: "Neutral" pronouns that aren't. *American Psychologist*, 33(11), 1032–1036.
<https://doi.org/10.1037/0003-066X.33.11.1032>
- *Moyano, S., de Lemus, S., Kerleaux, V., Krawietz, O., & Dasgupta, N. (2022). *More than words: Gender inclusive language reduces identity threat for women in maths* [Manuscript in preparation]. University of Granada, Spain.
- Müller, K., Wickham, H. (2021). *tibble: Simple data frames*. R package version 3.1.4,
<https://CRAN.R-project.org/package=tibble>
- *Nissen, U. K. (1997). Do sex-neutral and sex-specific nouns exist: The way to non-sexist Spanish? In F. Braun & U. Pasero (Eds.), *Kommunikation von Geschlecht* (pp. 222–241). Centaurus.
- *Nissen, U. K. (2013). Is Spanish becoming more gender fair? A historical perspective on the interpretation of gender-specific and gender-neutral expressions. *Linguistik online*, 58(1/13), 99–117. <https://doi.org/10.13092/lo.58.241>
- Olos, L., & Hoff, E.-H. (2006). Gender ratios in European psychology. *European Psychologist*, 11(1), 1–11. <https://doi.org/10.1027/1016-9040.11.1.1>
- Payr, F. (2022, December 05). Gendersender [TV channel using gender-inclusive language]. *Frankfurter Allgemeine Zeitung*. <https://www.faz.net/aktuell/feuilleton/medien/kritik-am-gendern-ard-und-zdf-reagieren-nicht-auf-aufruf-18511532.html>
- *Prentice, D. A. (1994). Do language reforms change our way of thinking? *Journal of Language and Social Psychology*, 13(1), 3–19.
<https://doi.org/10.1177/0261927X94131001>
- Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and gender-

- less languages. *Sex Roles: A Journal of Research*, 66(3-4), 268–281.
<https://doi.org/10.1007/s11199-011-0083-5>
- Pustejovsky, J. (2022). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections*. R package version 0.5.8, <https://CRAN.R-project.org/package=clubSandwich>
- R Core Team. (2020). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- *Rothmund, J., & Scheele, B. (2004). Personenbezeichnungsmodelle auf dem Prüfstand [Person designation models put to test]. *Zeitschrift Für Psychologie / Journal of Psychology*, 212(1), 40–54. <https://doi.org/10.1026/0044-3409.212.1.40>
- Rücker, G., Krahn, U., König, J., Efthimiou, O., Papakonstantinou, T., Schwarzer, G. (2021). *netmeta: Network meta-analysis using frequentist methods*. R package version 1.5-0, <https://CRAN.R-project.org/package=netmeta>
- *Safir, M. P., Hertz-Lazarowitz, R., BenTsvi-Mayer, S., & Kupermintz, H. (1992). Prominence of girls and boys in the classroom: Schoolchildren's perceptions. *Sex Roles*, 27(9-10), 439–453. <https://doi.org/10.1007/BF00290002>
- *Sato, S., Gabriel, U., & Gygax, P. M. (2016). Altering male-dominant representations. *Journal of Language and Social Psychology*, 35(6), 667–685.
<https://doi.org/10.1177/0261927X15625442>
- *Scheele, B., & Gauler, E. (1993). Wählen Wissenschaftler ihre Probleme anders aus als WissenschaftlerInnen? Das Genus-Sexus-Problem als paradigmatischer Fall der linguisti-

schen Relativitätsthese [Do male scientists choose their problems differently from female scientists? The Genus-Sexus Problem as a paradigmatic case of the linguistic relativity thesis]. *Sprache & Kognition*, 12(2), 59–72.

Schmoll, H. (2022, February 24). Niedersachsens Schüler dürfen bei Abiturprüfung gendern [Lower Saxony's students allowed to use gender-inclusive language in the exam for university entrance]. *Frankfurter Allgemeine Zeitung*. <https://www.faz.net/aktuell/politik/inland/niedersachsen-erlaubt-gendern-bei-abiturpruefungen-17830931.html>

*Schneider, J. W., & Hacker, S. L. (1973). Sex role imagery and use of the generic "man" in introductory texts: A case in the sociology of sociology. *The American Sociologist*, 8(1) 12–18. <https://www.jstor.org/stable/27702068>

*Schütze, C. (2020). *Comprehension of gender-neutral forms and the pseudo-generic masculine in German: a visual world eye-tracking study: 'It goes without saying' that everyone is included?* [Unpublished Master's thesis]. University of Potsdam, Germany.

Schwarzer, G., Carpenter, J., Rücker, G. (2021). *metasens: Statistical methods for sensitivity analysis in meta-analysis*. R package version 0.6-0, <https://CRAN.R-project.org/package=metasens>

*Sniezek, J. A., & Jazwinski, C. H. (1986). Gender bias in English: In search of fair language. *Journal of Applied Social Psychology*, 16(7), 642–662. <https://doi.org/10.1111/j.1559-1816.1986.tb01165.x>

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In K. Fiedler (Ed.), *Social communication* (pp. 163–187). Psychology Press.

*Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen [Effects of the generic masculine language forms on the mental inclusion of women]

- culine and alternative forms of language on the mental inclusion of women]. *Psychologische Rundschau*, 52(3), 131–140. <https://doi.org/10.1026//0033-3042.52.3.131>
- *Steiger-Loerbroks, V., & von Stockhausen, L. (2014). Mental representations of gender-fair nouns in German legal language: An eye-movement and questionnaire-based study. *Linguistische Berichte*, 2014(237), 57–80.
- *Stout, J. G., & Dasgupta, N. (2011). When he doesn't mean you: Gender-exclusive language as ostracism. *Personality & Social Psychology Bulletin*, 37(6), 757–769. <https://doi.org/10.1177/0146167211406434>
- *Switzer, J. Y. (1990). The impact of generic word choices: An empirical investigation of age- and sex-related differences. *Sex Roles*, 22(1-2), 69–82. <https://doi.org/10.1007/BF00288155>
- *Tavits, M., & Pérez, E. O. (2019). Language influences mass opinion toward gender and LGBT equality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(34), 16781–16786. <https://doi.org/10.1073/pnas.1908156116>
- Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research Synthesis Methods*, 4(2), 169–187. <https://doi.org/10.1002/jrsm.1070>
- *Vainapel, S., Shamir, O. Y., Tenenbaum, Y., & Gilam, G. (2015). The dark side of gendered language: The masculine-generic form as a cause for self-report bias. *Psychological Assessment*, 27(4), 1513–1519. <https://doi.org/10.1037/pas0000156>
- *Vervecken, D., Gyga, P. M., Gabriel, U., Guillod, M., & Hannover, B. (2015). Warm-hearted businessmen, competitive housewives? Effects of gender-fair language on adolescents' perceptions of occupations. *Frontiers in Psychology*, 6, Article 1437. <https://doi.org/10.3389/fpsyg.2015.01437>

- * Vervecken, D., & Hannover, B. (2015). Yes I can! Effects of gender fair job descriptions on children's perceptions of job status, job difficulty, and vocational self-efficacy. *Social Psychology*, 46(2), 76–92. <https://doi.org/10.1027/1864-9335/a000229>
- *Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (S)expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior*, 82(3), 208–220. <https://doi.org/10.1016/j.jvb.2013.01.008>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. *Proceedings of the 2nd ACM SIGHIT international health informatics symposium*, 819–824.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>
- Wickham, H. (2019). *stringr: Simple, consistent wrappers for common string operations*. R package version 1.4.0, <https://CRAN.R-project.org/package=stringr>
- Wickham, H. (2021). *forcats: Tools for working with categorical variables (factors)*. R package version 0.5.1, <https://CRAN.R-project.org/package=forcats>
- Wickham, H. (2021). *tidyr: Tidy messy data*. R package version 1.1.3, <https://CRAN.R-project.org/package=tidyr>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H.

- (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
<https://doi.org/10.21105/joss.01686>
- Wickham, H., Bryan, J. (2019). *readxl: Read excel files*. R package version 1.3.1,
<https://CRAN.R-project.org/package=readxl>
- Wickham, H., François, R., Henry, L., Müller, K. (2021). *dplyr: A grammar of data manipulation*. R package version 1.0.5, <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., Hester, J. (2020). *readr: Read rectangular text data*. R package version 1.4.0,
<https://CRAN.R-project.org/package=readr>
- *Wilson, L. C. (1978). Teachers' inclusion of males and females in generic nouns. *Research in the Teaching of English*, 12(2), 155–161. <https://www.jstor.org/stable/40170693>
- World Economic Forum. (2022). *Global gender gap report 2022*. https://www3.weforum.org/docs/WEF_GGGR_2022.pdf
- *Xiao, H., Strickland, B., & Peperkamp, S. (2022). How fair is gender-fair language? Insights from gender ratio estimations in French. *Journal of Language and Social Psychology*, 42(1), 82–106. <https://doi.org/10.1177/0261927X221084643>
- *Zacharski, L., & Ferstl, E. C. (2022). *Gendered representations of person referents activated by the nonbinary gender star in German: A word-picture-matching task* [Unpublished manuscript]. University of Freiburg, Germany.
- Zeileis, A., Grothendieck, G. (2005). zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14(6), 1–27. <https://doi.org/10.18637/jss.v014.i06>
- Zeileis, A., Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News*, 2(3), 7–10. <https://CRAN.R-project.org/doc/Rnews/>

Zoske, S. (2021, December 15). Noten für Gendern in bestimmten Fällen zulässig [Grades for gender-inclusive language in certain cases allowed]. *Frankfurter Allgemeine Zeitung*.
<https://www.faz.net/aktuell/rhein-main/noten-fuer-gendern-in-bestimmten-faellen-zu-laessig-17674497.html>