Gendered Language and Gender Essentialism: Exploring What Gender/Sex Terms Communicate About a Speaker's Attitudes and Beliefs About Gender

by

Solangel C. Troncoso

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology and Women's and Gender Studies)
at the University of Michigan
2025

Doctoral Committee:

Professor Susan A. Gelman, Chair
Associate Professor Zach Schudson
Professor Isis Settles
Professor Emerit Abigail Stewart

Solangel C. Troncoso

troncoso@umich.edu

ORCID iD: 0000-0002-5560-8537

© Solangel C. Troncoso 2025

Dedication

To my mom and dad, my husband, and my daughters.

Acknowledgments

I am beyond grateful for the community I come from and the community I have created for myself. Writing my acknowledgements has been an intimidating task because of how much love and support I am surrounded by. The Ph.D. process is known to be isolating, and I have missed important moments and events of those I love and have felt the distance of being away from home. Yet, my family, my friends, and my community have made all those things easier. You have all made this possible.

I want to thank my parents who valued education and let it be my main priority. Mom, thank you for your sacrifices and for allowing me to pursue my goals. None of this would be possible without you. Dad, thank you for your consistent encouragement and long chats. Erie, thank you for always being super excited for me and cheering me along the way. I am proud of you and being your big sister is one of my favorite roles. Jahaira, Madelina, and Fredin, I am proud to share your last name and I know Papi would be so proud to have not one, but two Drs. Troncoso in the family.

The Decena Family. I am so proud to have a big yet tight-knit family full of aunts (Cura, Nuni, Pepe, Damaris, Taty, Coca) and uncles (Chucho, Ernesto, Juambo, Norvo) who show up for me when I need it and to celebrate all of life's milestones. Tia Pepe, thank you for taking the steps that paved the way for me to go away for college. Your courage made things easier for me and for those who come after us. Cynthia, Destiny, Jenny and all the older cousins, I have loved growing up with you all and look forward to continuing our family traditions. To my little

cousins, I am so honored anytime one of you asks me anything. I hope my journey has eased yours in some way and I am so excited for you all. Thank you to my grandparents; I love you so much. Thank you for our family.

To my husband, my daughters, and the Mercado Family. Edward, I love you so much. I am so grateful for the life we have built together. You have listened to me express all my doubts, sometimes over and over again about the same thing and have been the best sounding board. You have reassured me and challenged me when I needed it. More than anything, you have been my safe place. Sofia and Amanda. Sofia, at almost two years old, you are already this assertive and courageous little being. I admire you so much. Amanda, at 4 months old, you are all smiles and working diligently on your milestones. I am so excited watching you both grow. I hope you take some lessons from my journey and more than anything that you choose your own paths and go after what you want in life. Yoly, Pancho, Nayalit & Jason, Mandi & Rosy, Kara & Jubilee and the Mercado family; thank you for sharing your son/brother/Tio Pito with me and for your love and support. You have taught me what it means to show up for your loved ones in tangible ways. Thank you for allowing me into your family.

My friends. I feel so lucky and grateful for you all. Thank you for supporting me and my goals. Thank you for the love and laughs and endless text messages and group chats. Thank you for the outings and the vacations and also for showing up when things are difficult. I am so lucky to have found you all at different stages of life and to be able to continue to celebrate with you all. Fara & Daniel, Jocelyn & Branden, Destiny & Brian, Joshua & Jessica Ashley & Xahn, Bianka & Alexis, Becca & Cris, Madeline & Jorge, Ana Patricia, Ivette, Brittany and many other group chats combinations, thank you for always being a call or text away, whether we were in the same state or not. Thank you for all the times you all helped us moved on a whim. In the time

that it has taken to complete this program, we have welcomed so many little ones; Ziel, Olivia, Jack, Mia, Maya and Liam, your parents are some of my favorite people and I am so glad Sofia and Amanda have you all.

Thank you to my committee, Dr. Gelman, Dr. Schudson, Dr. Settles, and Dr. Stewart. Each one of you has been a great source of support and guidance and it has meant a lot to have you in my corner all these years. Dr. Gelman, you are an amazing mentor and role model for how to guide students to accomplish their goals while also caring for them holistically. Thank you for your patience and for working with me through the changes, good and bad, that I have experienced in graduate school. You gave me space to shift and adjust and kept me on track. Your reassurance and confidence in me and my ideas have carried me to the finish line. Dr. Schudson, thank you for always being there to answer my questions and provide encouraging words. I am so grateful for the opportunity to learn from you. Dr. Stewart, thank you for pointing me in the right direction time and time again. Dr. Settles, thank you for opening your office (and Zoom) door when I needed it. I am truly grateful to have had a committee of people who are experts in their fields and great teachers and mentors who truly care about their students.

I have been part of amazing programs and fortunate to have been mentored by great people. Dr. Ceballo, thank you for welcoming me into your lab, Resilience in Context, and introducing me to different kinds of research and for your support. Dr. Martínez-San Miguel, thank you for choosing me to be your undergraduate research assistant; working with you was such a lovely introduction into the world of research. Dr. Price, your words have encouraged me long past our conversations. Dr. Sanchez, you are the reason I knew it was possible to pursue Psychology and Women's and Gender Studies; thank you for showing me the way. To all the professors and teachers who made an impact on me over the course of my educational career-

thank you for taking an interest in me. I work to pay it forward. Thank you to all the staff that make EOF and McNair programs possible. They change lives. RMF, LSPA, and BSPA thank you for providing me with community at UM. Thank you to my writing groups and everyone I have worked with on Zoom over the years. You made me feel connected during the pandemic and while I have been away from campus.

With love and deep gratitude,

Dr. Troncoso

Table of Contents

Dedication ii

Acknowledgments iii

List of Tables ix

List of Appendices x

Abstract xi

Chapter I: Introduction 1

- Linguistic Sexism 2
- Speaker Perception 4
- Language as a Political Tool Regarding Gender 6
- How Race Relates to Perceptions of Gender Terms 9
- Heteronormativity 10
- The Present Studies 11

Chapter II (Study 1) 16

- Method 17
- Results 25
- Discussion 32

Chapter III (Study 2) 37

- Method 39
- Results 43
- Discussion 58

Chapter IV: General Discussion 64

References.....70

Appendices.....81

List of Tables

Table 1 (2.1) Demographics 18

Table 2 (2.2) Feminism, Political Leaning, and Language Background 20

Table 3 (3.3) Preferred Gender Term 21

Table 4 (2.4) Coding Example 25

Table 5 (3.1) Demographics 40

Table 6 (3.2) Context x Word Interactions (Perceived speaker traits) 44

Table 7 (3.3) Context x Word Interactions (Perceived speaker demographics) 45

Table 8 (3.4) Individual Scale Responses as a Function of Participant Self-Identified Race 56

List of Appendices

Appendix A: Study 1 Survey Tasks.....81

Appendix B: Study 1 SCEBS, ASI, and Demographic Questions 83

Appendix C: Study 2 Speaker Snapshot Task 87

Appendix D: Study 2 SCEBS, ASI, and Participant Demographics 95

Abstract

The words we use provide a powerful tool that can shape and reflect people's attitudes and beliefs about gender. Some scholars and online media contributors suggest that the term *female* feels offensive or odd outside of a medical or biological context. Drawing on my prior research, I hypothesized that those using the words "female" or "male" (vs. "woman" or "man") would be perceived more negatively, especially in contexts where reference to gender may be unexpected (e.g., daily errands). I also hypothesized that participants who are less essentialist, less sexist, and more feminist would have the most negative associations. In this dissertation, I test this idea in two studies (N=760) examining how gendered terms may serve as a signal to others about the speaker's attitudes and beliefs about sex and gender.

Study 1 examined how Black and White U.S. participants perceive imagined speakers depending on the gendered label they used (*female/male*, *woman/man*, *girl/boy*, *lady/gentleman*). In several ways, a speaker using the terms *female* or *male* elicited distinct inferences or evaluations. Participants mentioned the speaker's gender/sex more when the speaker used the terms *female* or *male* than any of the other gender terms. Participants also had more negative associations about a speaker who used the terms *female/male* in the errands context.

Study 2 presented a stronger signal of how the hypothetical speaker talked and assessed a wider range of judgments about the speaker. I focused on the linguistic contrast between the terms *woman* vs. *female*. Overall, participants imagined the speaker who used the word *female* to be less likable, warm, competent, educated, formal, and polite, compared to the speaker who

used the word *woman*. Participants also felt less warmly about the speaker who used the word *female* and perceived them to be less respectful of women they referenced, compared to the speaker who used the word *woman*. Last, participants' own beliefs and attitudes were predictive of their evaluations of the speaker. Specifically, participants higher on essentialism and hostile sexism rated the speaker using the term *female* (relative to *woman*) more positively, and participants higher in feminism rated the speaker using the term *female* (relative to *woman*) more negatively.

I conclude that subtle shifts in language can function as a cue to others about a speaker's attitudes and beliefs about gender. Specifically, calling women *female* repeatedly can serve as a signal to others that the speaker holds more sexist, more essentialist, and less feminist attitudes toward women than using the term *woman* to refer to women. This research has broad implications for people in varied professional settings and in interpersonal contexts. Changing how people speak about women, on its own, may not change people's attitudes and beliefs, but it is an important step toward gender-fair language and gender equality more broadly.

Chapter I: Introduction

What can language about gender tell us about gender inequity? Feminist activists and scholars turned their attention to the negative impacts of sexist language in the 1970s, directing their efforts toward advocating for policy changes (Wang & Ngai, 2023). Early research focused on overtly sexist language practices, such as: occupation labels that assumed men as the default (e.g., "policeman", "stewardess"), arguing instead for a shift to gender-neutral labels (e.g., "police officer," "flight attendant"); titles for women that required indicating whether they were married or not (e.g., "Miss" and "Mrs."), arguing instead for an unmarked option (e.g., "Ms." for women, equivalent to "Mr." for men); and problematizing the use of the word "he" instead of "he or she" and "mankind" instead of humankind (Kleinman et al., 2021; Parks & Roberton, 2004).

Feminist scholars have also noted that some terms may seem inoffensive but nonetheless more subtly express sexist notions or assumptions (Kleinman et al., 2021; Lakoff, 1970). For example, Lipton and Hershaft (1984) examined the impact of assigning gender labels to artists on ratings of the artist's art piece. In their study, the authors posited that there is a difference between using the terms *woman* and *girl* to refer to women, labeling *woman* as "high status" and *girl* as "low status." They also included the term *person* as a gender-neutral option. They found that only the term *person* led to more positive ratings of women's art. More recently, scholars have called attention to the relation between people's political attitudes and their choice of

pronouns to refer to mixed-gender groups, such as "*they"* (gender-inclusive) or "*you guys*" (gender-biased) (Bedin, 2025; Bradley, 2020; Kleinman et al., 2021; Saguy & Williams, 2022).

In this dissertation, I am interested in examining if and how people's attitudes about sex and gender may be reflected in their reactions to terms referring to a person's sex (*female*; a term that also applies to non-human animals) or to a person's gender (*woman*; a term that exclusively refers to humans). The use of the term *female* may reflect gender essentialist beliefs, by prioritizing biological characteristics of sex rather than social aspects of gender. In my prior work, I tested this idea by examining how people interpret these words, and in what contexts they deem them appropriate. In that work, I found that there is a positive link between preference for the term "*females*" over the term "*women*" and higher endorsement of gender essentialism (Troncoso et al., 2022). The two studies in this dissertation are designed to examine a new question, namely, how we evaluate speakers based on their choice of gendered words to refer to others in daily life. Given that the language about sex and gender that we use may have real impacts on gender equity and discrimination, it is important to continue examining the ways that our everyday language use shapes and/or reflects our attitudes.

In the remainder of this introduction, I first briefly review relevant literatures to my overarching research question, regarding: linguistic sexism, speaker perception, language as a political tool, the role of race, and heteronormativity. I then present an overview of the two studies.

Linguistic Sexism

In the 1970s, there was a concerted effort to make language about gender more equitable. Robin Lakoff (1973), a pioneer in conceptualizing how sexist language reinforces gender inequity and the status quo, raised many questions in this space. For example, she asked why

there are so many more words referring to women than words referring to men. Lakoff argued that this "linguistic imbalance" reflects real-world imbalances. She speculated that people use euphemisms for words they are uncomfortable with, and thus the disparity in choice of words is due to general discomfort with the word *woman*. She further argued that the word *woman* invokes sexual thoughts that people want to hide and therefore it is often replaced with the euphemism, *lady*, that is perceived to be less linked to sex and therefore more polite. Lakoff (1973) argued that linguistic change and social change go hand-in-hand, and thus social change does not happen just by changing the language; instead, people must be receptive to the social change first. In other words, for sexist and derogatory language about women to change, people must be receptive to the need for the status of women as inferior to men to change. Although Lakoff raised many important ideas, this was primarily a theoretical framework with little empirical evidence.

It is clear that using overtly hostile terms to refer to women (such as the "B-word" or the "C-word") reflects negative attitudes toward women. It is less clear, however, how language that is not overtly hostile, or is even seemingly neutral, may function to communicate messages about the speaker's attitudes toward women. For example, some speakers may engage in an unreflective sexist practice, one that although not intending harm, alienates those who are sensitized to it (Kleinman et al., 2021). Kleinman et al. (2021) provide the example of a student who did not feel included when the teacher used the word "guys" to refer to both women and men in the class (e.g., "What do you guys think?"). Wang and Ngai (2023) note that some people may be using language that is sexist without realizing it because they are unfamiliar with the different uses or connotations of the language. As Wang and Ngai explain, linguistic sexism shows up across different contexts, including casual everyday conversations and professional

settings. Whether overtly or indirectly, linguistic sexism is closely related to our beliefs about gender and thus, how we speak about gender can reflect and reinforce negative attitudes. These not-overtly-hostile variations in words may be more difficult to combat, and so especially pernicious. For example, a speaker can deny that they meant anything wrong by using a particular term or may even be unaware of the messages implied (Wang & Ngai, 2023)--especially if there are mismatches across cultural groups in what a particular word connotes.

Speaker Perception

What sorts of judgments or evaluations do people make about speakers, based on how they talk? Although my interest is in how such evaluations relate to gendered terms, I will first briefly review a sampling of the sorts of inferences people make beyond gender per se. This work illustrates how powerful language is, as a cue to evaluations of others. People make inferences about others based on their accent (Kinzler, 2021; Labotka & Gelman, 2020, 2024), prosody (i.e., the intonation in a speaker's voice, which conveys information about their attitudes) (Vergis & Pell, 2020), and the words a speaker chooses to describe a third party in conversation (Ames et al., 2010). Fausey and Boroditsky (2010) found that linguistic framing or subtle linguistic cues (reporting an event in an agentive versus non-agentive framing) affect attribution of blame, such that participants assigned more blame when an event was described using agentive language ("she had ignited the napkin" vs. "the napkin ignited"). Berry et al. (1997) found that an individual's specific word choice has a great impact on the first impression they make. For example, participants attributed higher ratings of warmth when the speaker used the present tense, fewer negative emotion words, and fewer "unique" words. People also make judgments about others based on their choice to use profanity. DeFrank and Kahlbaugh (2019) found that participants rated speakers using profanity lower than those not using profanity, with

regard to intelligence, trustworthiness, proneness to anger, deviancy, politeness, offensiveness, aggressiveness, and likability. Overall, research shows that a range of linguistic features, including word choice, sentence structure, and profanity, make an impact on how people judge others.

In addition to word choice and context, speaker perception is also related to the listener's own identity, attitudes, and beliefs. For example, Boasso et al. (2012) examined how women perceive women who use sexist terms of endearment toward other women, as well as how they perceive women who receive these terms. They found that when viewing a video with a woman using a sexist term of endearment, the more a participant endorsed benevolent sexism, the more she disliked the speaker and the more she liked the person being spoken to. DeFrank and Kahlbaugh (2019) found gender effects related to the use of profanity. Specifically, they found that women judged men's swearing to be more offensive, and men judged women's swearing to be more offensive. Moreover, in general, profanity in mixed-gender dyads was seen as more offensive than in same-gender dyads. These findings highlight how gender of the speaker can be taken into consideration when judging a speaker's language use.

Importantly, these evaluations also extend to gendered speech. Lakoff (1975) proposed that women have a different way of speaking than men and that this affects how they are perceived. For example, according to this theory, women use more hedges, empty adjectives, tag questions, and compound requests (Lakoff, 1973; Mulac et al., 1985; Newcombe & Arnkoff, 1979). Newcombe and Arnkoff (1979) tested Lakoff's theory and found that the speech styles did yield different results, with "female forms" of speech rated as less assertive and more warm and polite than "male forms," regardless of the gender of the speaker. In other words, it's the speech style and not the gender of the speaker that is informing the ratings. Based on this

research, Mulac and Lundell (1985) coined the term the *gender-linked language effect*, in which people perceive differences in speech, which they attribute to the speaker's gender. In one study, Mulac et al. (1985) found that "women's speech" was rated higher on socio-intellectual status and aesthetic quality (nicer, more pleasant, sweeter, and more beautiful) and "men's speech" was rated higher on dynamism (more active, more aggressive, stronger, and louder). The authors propose that the word choice and manner of speaking that men and women use are what drive this perception and not just sex role stereotypes. Haleta (1996) conducted a study on students' first impressions of their professors and examined the impact of "powerful" (concise and direct) and "powerless" (use of hesitations and hedges) language. The author found that students rated professors who used "powerful" language more favorably on dimensions of dynamism, status, and credibility and professors who used "powerless" language as more uncertain. Students also rated men professors higher in status than women professors.

More recent literature shows that the gender-linked language effect is relevant today. For example, after analyzing data from TED talks, Meier et al. (2020) found that more "female-typical" language style predicted ratings such as "beautiful" and "courageous," as opposed to ratings such as "ingenious" and "informative," which were more often used for TED talks by men. Altogether, these studies suggest that people form perceptions of others based on the style of speech they use and that this may have consequences for both first impressions and lasting judgments of both men and women.

Language as a Political Tool Regarding Gender

Given the current political climate, in which issues related to gender and sex are often considered controversial topics (for example, bans on education about LGBT issues or drag shows and policies dictating who can use which bathrooms; Barbee et al., 2022; David, 2017;

James et al., 2016), the use of language as a political tool is an especially timely and important issue. In fact, in January 2025, under the Trump Administration, The White House put forth an executive order declaring that the United States will only recognize two sexes: male and female. The order states that the administration's goal is to "defend women's rights and protect freedom of conscience by using clear and accurate language and policies that recognize women are biologically female, and men are biologically male" (The White House, 2025). They provide the following definitions:

> (a) "Sex" shall refer to an individual's immutable biological classification as either male or female. "Sex" is not a synonym for and does not include the concept of "gender identity."
> (b) "Women" or "woman" and "girls" or "girl" shall mean adult and juvenile human females, respectively.
> (c) "Men" or "man" and "boys" or "boy" shall mean adult and juvenile human males, respectively.
> (d) "Female" means a person belonging, at conception, to the sex that produces the large reproductive cell.
> (e) "Male" means a person belonging, at conception, to the sex that produces the small reproductive cell.

Of interest to my work is how much of the order is based on the language that we are to use to refer to sex, and to ignore the "concept of 'gender identity.'" The executive order states, "The erasure of sex in language and policy has a corrosive impact not just on women but on the validity of the entire American system." The demand to return to "sex" in language as an effort to reinforce sex as a biological and immutable concept speaks to the belief that language has the power to shape not just perceptions of reality but also reality itself. A part of my work in the following studies is to examine how much the terms that we use in everyday conversation communicate our attitudes and beliefs about gender.

Just as the terms that we use can serve as signals about our attitudes, the terms we refuse to use also serve this purpose. Researchers have studied the link between different types of

sexism and resistance to newer, more progressive terms (Bailey et al., 2024; Patev et al., 2019), to gender-biased language (Cralley & Ruscher, 2005), and to attitudes toward and ability to detect sexist language (Parks & Roberton, 2004; Swim et al., 2004). Bailey et al. (2024) examined support for binary gender pronouns and found that people who highly endorse gender essentialism show more support for binary pronouns. The authors call for more research to establish the possibility that "support for binary gender pronouns may provide an indirect means for people high in (binary) gender essentialism to express negative attitudes toward transgender and non-binary individuals" (Bailey, 2024, p. 52). Patev et al. (2019) found that participants who held more negative attitudes toward transgender individuals perceive gender-inclusive language as more difficult to use, even if provided clear rules for use (p. 348). For example, the use of the term "they" as a singular pronoun has gained a lot of attention and can be seen as an example of how people's belief systems about sex and gender are tied to language. Research shows that people with more sexist, or generally more negative, attitudes about transgender people are more likely to resist the use of the pronoun "they" (Bailey et al., 2024). People also make assumptions about others' political ideologies depending on whether they use "gender-fair language." For example, Tarillo (2023) found that when participants read about a candidate using gender-fair language, they rated the candidate as more liberal and more likely to be a Democrat.

In their review of the literature, Saguy and Williams (2022) point to further nuances; it is not only those with negative attitudes toward transgender people who might refuse to use the term "they," but also people who support the use of binary pronouns (as well as identities) within the transgender community. Similarly, in their work on resisting the term "you guys" to refer to mixed-gender groups, Kleinman et al. (2021) identified various reasons why people who identify as feminist and are open to gender-inclusive language would continue its use. The authors

conclude that participants sought to justify their own belief system rather than admit to doing something that might be seen as "unfeminist." They refer to this as a form of "identity work" that functions to hold on to a moral identity.

How Race Relates to Perceptions of Gender Terms

An individual's beliefs and attitudes reflect multiple intersecting identities that shape their lived experiences. There are many factors (e.g., age, gender, race, cultural background, geographic location, familiarity with a term) that can contribute to how an individual may perceive the use of any given gender term, including different ways of defining and experiencing womanhood and manhood (Brassel et al., 2020; Settles et al., 2008). One important factor that I briefly consider here is how race and gender intersect and how this may influence gender term perception. Historic and systemic racism in the United States has contributed to different gender roles, norms, and expectations for Black and White women and men. Still today, White, middle class, women are seen as the standard of femininity while Black women are expected to aim for this standard while contending with harmful stereotypical portrayals (such as hypersexual Jezebel) and culture-specific norms (such as the Strong Black woman schema) (Jerald et al., 2017; Jones et al., 2018; Leath et al., 2021; Settles et al., 2008).

Another factor that may influence how people perceive gender terms is the historical context of how those terms have been used toward a specific group. There is evidence of gendered terms being racialized and of racist use of gendered terms against Black people in the U.S. For example, the terms *girl* and *boy* were used during slavery and into the Jim Crow era as a way to infantilize and deny Black people proper titles (Bennett-Alexander, 2010; WLRN Public Media, 2017). The term *boy* in particular still has racist connotations (Henderson, 2003). The

term *lady* also has historical context in which White women were referred to as *lady*, yet Black women were denied this title (Bush, 1981).

Recent research indicates that the term *females* is also racialized. Joensuu (2023) conducted a corpus analysis using the Corpus of Contemporary American English to examine the derogatory use of the word *females.* The author found that the top collocates of the term *females* were 'Black' and 'White' in reference to racial identity. In other words, the terms Black and White were often paired with, or near, the term *females* in the analyzed text. In Joensuu's (2023) data, the word 'Black' (15) was used about twice as often as 'White' (7) as an adjective collocate preceding the term *females*. The author also notes that in one of the analyzed quotes, the same person used the phrases "white women" and "black females." Although more research is needed, these findings indicate the interweaving of race and gender in language.

Heteronormativity

Gendered language use may also reveal heteronormativity – that is, a social system that privileges heterosexual relationships as 'normal' and 'natural' (Barker, 2014; Habarth, 2015; Marchia & Sommer, 2019). Given that the current study analyzes a range of gendered terms for both women and men, there is an opportunity to test whether heteronormativity is reflected in participants' responses, and if this varies by context. For this study, I operationalized heteronormativity as a participant's assumption that the speaker was of the 'opposite gender' as the person they referenced in the quotes. For example, if a participant read the quote "A friend asked me what I am looking for in someone. I responded that I am looking for a woman who shares my values and interests." and described the speaker as a man, this was marked as a heteronormative response. Heteronormative language can have a harmful impact on individuals, especially in the LGBTQ community (Estes, 2017).

One area where people are paying attention to heteronormative language use is in digital spaces. Vásquez et al. (2022) developed *HeteroCorpus* to detect and mitigate the use of heteronormative language in Natural Language Processing (NLP) systems. For example, the authors developed a coding system that would identify tweets as either heteronormative or not. They explain that detecting heteronormative language is an urgent task in efforts to mitigate gender bias within the rapidly growing field of Artificial Intelligence as such biases have real-life consequences for gender, sexual, and racial minorities. Less is known about whether the use of specific gendered terms (e.g., *female* vs *woman*) is more likely to yield heteronormative responses.

The Present Studies

In my dissertation, I examine how people perceive the use of several gendered terms. More specifically, I examine how a person's choice of gendered labels such as "*female/male*" or "*woman/man*" may serve as a signal to others about their (the speaker's) attitudes and beliefs about sex and gender. I ask the following questions: What inferences do people make about a speaker, based on their choice of a gendered label (e.g., *woman* vs. *female*)? Regardless of the speaker's intent, their language choice may be assumed to reflect their views on gender and sex and thereby shape others' perception of them. Furthermore, I examine how context influences speaker perception and how participant's own attitudes and beliefs affect their judgments of the speaker.

This work builds on prior research indicating that the term *female* is commonly associated with biological content (e.g., mentions of the body, genitals, chromosomes) (Schudson et al., 2019; for popular writing on this issue, see Norris, 2019; Sommers 2025). Likewise, in my earlier research (Troncoso et al., 2022), I found evidence linking the word

"female" to gender essentialism, in two ways. First, participants judged the word *female* to be more closely linked to biological and formal contexts than the word *woman.* For example, "female" was judged to be an appropriate way of referring to women at a doctor's office or on an official form (Troncoso et al., 2022; see also Schudson et al., 2019), whereas *woman* was deemed more appropriate and casual. And second, people who spontaneously used the word *female* relatively more often than *woman* when referring to women, scored significantly higher in gender essentialism.

In my dissertation studies, I focus on the use of the gendered terms, such as *female* (or *woman*), across various contexts, including where they may be unexpected. Specifically, although the term *female* is broadly deemed appropriate for referencing sex, some people report that it feels offensive, odd, or that there is something off-putting about hearing a person refer to a group of women as *females* outside of a medical or biological context, at times serving as "red flag" against the person speaking (Bee, 2025; D'Oyley, 2013; Joensuu, 2023). Similarly, others have suggested that the use of the term *females* to refer to women is derogatory and possibly conveys sexist attitudes in certain contexts (Brown, 2015; Joensuu, 2023; Moody, 2021; Sommers, 2025). The term *females* has also been thought to be dehumanizing, used to compare women to cattle (Joensuu, 2023), and reducing women to reproductive organs (Sullivan, 2021). Moreover, the term *female(s)* may be used in place of more overtly negative words. For example, in a blog post titled Please Don't Refer to Women as Females, Sommers (2025) states that by using the term *females* the speaker has conveyed his negative attitudes about women and that she would actually prefer them to use a more overtly negative word rather than hiding behind a biological term. She states, "You've revealed how you truly perceive women around you — and the other choice words you probably use to talk about them. I'd honestly rather be referred to as

some of those words because they're at least actual insults and not a reconfiguration of a biological term." In another blog post, Bee (2025) writes "It's a simple term, and accurate at that, however the way it is used feels like it places women as inferior." In other words, there is an ongoing conversation online in which many have observed the term *female* used negatively and perceive the use of the term to reveal the speaker's beliefs and attitudes about women. However, these suggestions have not been tested directly or systematically. As such, one aim of my dissertation is to understand how use of these labels (e.g., *female/male* vs. *woman/man*) are perceived by others. For example, one question of interest is: Do people perceive someone who repeatedly refers to women as *female(s)* as more gender essentialist and/or sexist?

In **Study 1**, I provided participants with a set of hypothetical speakers, each of whom used a gendered or gender-neutral term to describe another person. For each speaker, participants wrote a description of that person, and these responses were coded along various dimensions to assess speakers' judgments and evaluations. Gender is arguably the most salient social category (Rhodes & Mandalaywala, 2017) and as such, we have many terms to choose from to refer to a person's gender identity. In Study 1, I focused on the terms *women, men*, *female, male*, *girl, boy*, *gentleman, lady*, and *person*. As noted above, prior research shows that although there is often conflation of the definitions of sex and gender, people are more likely to associate the terms *female* and *male* with biology and the body than the terms *woman* and *man*, although all the gender/sex terms have complex meanings and associations (Schudson et al., 2019; Troncoso et al., 2022). Additionally, research shows that people associate sex with more essentialist features and gender with more social constructionist features (Lloyd & Galupo, 2019). Given these differences, one primary goal of my studies is to examine differences in how people perceive those using the terms *female* and *male* vs. *woman* and *man*. A secondary aim in Study 1 is to

understand how a speaker is perceived when using one of the following four terms perceived: *lady*, *gentleman*, *girl* and *boy*. The term *lady* has been theorized as a potential euphemism for the word *woman* (Lakoff, 1975) and may also be more associated with other factors such as the social status of the person being referred to, or the age of the person speaking. Although the term *gentleman* does not have the same connotations as *lady*, I included it as a comparison term. Lastly, the terms *girl* and *boy* both infantilize adults, with *girl* being used more often in a sexist manner (Cralley & Rusher, 2005; Lipton & Hershaft, 1984) and *boy* being used a racial epithet (Bennet-Alexander, 2010; Henderson, 2003). The term *person* will serve as a neutral comparison term.

Another factor to consider when judging the use of a word is the context in which it is used. For example, although the term *female* may seem appropriate on a form in a doctor's office, it may seem less appropriate when used to refer to a woman casually. With this in mind, I examined two contexts: dating and errands. Gender may be more salient in a dating context and thus may elicit more attention to the gendered labels. In contrast, the errands context, in which people are describing everyday activities, provides a test of attention to these labels in a more neutral context.

In **Study 2**, I again asked participants to evaluate hypothetical speakers based on the gendered label that they used. This study focused on the two labels of key interest, namely *female* and *woman*. Additionally, I designed the stimuli to provide a stronger signal of the hypothetical speaker's tendency to use a given label, by showing that each speaker used a particular label to refer to multiple people in different contexts (e.g., one speaker called four different people a "female," whereas another called four different people a "woman"). Finally,

Study 2 provided a series of targeted close-ended questions to assess participants' judgments more fully and systematically.

Overall, these two studies add to our understanding of the kinds of messages and information communicated by certain gendered language use, and more specifically will inform how choice of a term may be perceived to reflect the speaker's beliefs and attitudes about sex and gender. In both studies, I recruited equal numbers of Black and White participants, in order to examine the generality of the patterns obtained, given the intersection of gender and race in language perceptions.

Chapter II (Study 1)

Language choices communicate our attitudes and beliefs, whether intentionally or unconsciously. Study 1 is an exploratory study that examines how participants perceive imagined speakers depending on the gendered label a speaker was said to have used. The aim of this study was to assess what messages about the speaker are conveyed by the word they choose to use (e.g., 'female,' 'woman,' 'lady') – not what the terms themselves mean or connote.

I used an open-ended method in which participants were asked to describe, in their own words, the speaker of each of a set of quoted remarks, based on their first impressions. I provided a general prompt at the beginning, noting that they could mention anything they thought was relevant, including age, gender, race, and personality. Even with these cues, the open-ended method provided an opportunity for participants to decide what to focus on.

Specifically, I aim to understand:

1. What characteristics of a speaker are implied depending on the gendered label (*woman/man*, *female/male*, *lady/gentleman*, *girl/boy*) that they use? For example, what inferences do participants make about a speaker's age, race, gender, likeability, and competence, as a function of the gendered label used?
2. How do participants' evaluations of a speaker relate to participants' own beliefs about gender, as measured by the Social Constructionist and Essentialist Beliefs Scale (SCEBS)? Two hypotheses, for example, are that participants who more strongly endorse essentialist beliefs would have more positive responses about

speakers who use the terms *female/male*, and that participants who more strongly endorse social constructionist beliefs would have less positive responses about speakers using the terms *female/male*.

3. How do participants' evaluations of the speaker relate to their own sexist (or non-sexist) attitudes, as measured by the Ambivalent Sexism Inventory (ASI)? One hypothesis, for example, is that participants who more strongly endorse sexist attitudes would have more positive responses to the speaker using the labels *female/male*.

Additionally, I am interested in heteronormativity in participants' responses and analyzed this by tallying how often each participant assumed that the speaker differed in gender from the person they were describing. This could reflect participants' assumptions regarding the likelihood that a person will interact with someone of a different gender (e.g., in a dating context), and/or an assumption that a person will only choose to mention another person's gender if it differs from their own.

Method

I obtained IRB approval for this study by the University of Michigan, and pre-registered the exploratory predictions, method, and analysis plan for this study with AsPredicted (#130994).

Participants

I recruited 180 participants (90 Black, 90 White) from Amazon Mechanical Turk via Cloud Research (Litman et al., 2017) in May 2023. Eight people were excluded for not completing any of the tasks, four were excluded for providing off-topic responses on the open-ended task, and four were excluded for not responding on more than one item on the open-ended

task, resulting in a final sample size of 164 (83 Black participants, 81 White participants). I tested an initial *n* of 20 participants to ensure the study ran smoothly. These participants were included in the final sample.

Participants ranged in age from 21 to 77 years old ($M = 40.66$, $SD = 12.44$). Participants self-identified as a woman ($n = 84$), a man ($n = 78$), non-binary ($n = 1$), and "another gender" ($n = 1$). Given that online convenience sampling, such as Mturk, tends to over-represent White participants (Levay et al., 2016), I specifically recruited roughly equal numbers of White (49.4%, $n = 81$ participants self-identified as White or European American) and Black (50.6%, $n = 83$ self-identified as Black or African American) participants. Six people identified with more than one race. In the Black sample, one person also identified as White or European American and another person identified as both White or European American and Hispanic or Latinx. In the White sample, two people identified as Hispanic or Latinx, one person identified as Native American or Alaskan American, and one person identified as Asian or Asian American. I also asked participants about their sexual orientation, combined family income, and level of educational attainment (see Table 2.1 for a complete list of demographics). Lastly, I asked participants about their preferred gender term, whether they identify as feminist, their political leaning, and whether English is their first language (see Table 2.2 and Table 2.3 for complete list of additional questions).

Table 1 (2.1) Demographics

	Full Sample	Black Participants	White Participants
	n (%)	n (%)	n (%)
Participants (*N*)	164	83	81
Age in years, M (*SD*)	40.7 (12.4)	40.2 (12.7)	41.12 (12.3)
Gender *N (%)*			

Women	84 (51.2)	47 (56.6)	37 (45.7)
Men	78 (47.6)	35 (42.2)	43 (53.1)
Non-binary	1 (0.6)	1 (1.2)	0 (0.0)
Another gender	1 (0.6)	0 (0.0)	1 (1.2)
Race/Ethnicity *N* (%)			
Asian or Asian American	1(0.6)	0 (0.0)	1 (1.2)
Native American or Alaska Native	1(0.6)	0 (0.0)	1 (1.2)
Black or African American and White or European American	1(0.6)	1 (1.2)*	0 (0.0)
Hispanic or Latinx	3(1.2)	1 (1.2)	2 (2.4)
White or European American	81 (49.4)	x	x
Black or African American	83 (50.61)	x	x
Sexual Orientation *N* (%)			
Aromantic	1 (0.6)	0 (0.0)	1 (1.2)
Asexual	1 (0.6)	0 (0.0)	1 (1.2)
Asexual, heterosexual	1 (0.6)	1 (1.2)	0 (0.0)
Bisexual	10 (6.1)	3 (3.6)	7 (8.6)
Gay	1 (0.6)	1 (1.2)	0 (0.0)
Heterosexual / straight	132 (80.5)	69 (83.10	63 (77.8)
Lesbian	2 (1.2)	1 (1.2)	1 (1.2)
Pansexual	3 (1.8)	2 (2.4)	1 (1.2)
Missing	13 (8.0)	6 (7.2)	7 (8.6)
Family's combined yearly income *N* (%)			
Less than $15,000	18 (11)	10 (12.0)	8 (9.9)
$15,000 - $25,000	14 (8.5)	4 (4.8)	10 (12.3)
$25,000 - $35,000	12 (7.3)	6 (7.2)	6 (7.4)
$35,000 - $50,000	33 (20.1)	19 (22.9)	14 (17.3)
$50,000 - $75,000	46 (28.0)	22 (26.5)	24 (29.6)
$75,000 - $100,000	21 (12.8)	12 (14.5)	9 (11.1)
$100,000 - $150,000	15 (9.1)	9 (10.8)	6 (7.4)
Above $150,000	5 (3.0)	1 (1.2)	4 (4.9)
Education *N* (%)			
High School graduate	28 (17.1)	10 (12.0)	18 (22.2)
Some college	38(23.2)	27 (32.5)	11 (13.6)
Finished training other than college (e.g., vocational school)	5(3.0)	3 (3.6)	2 (2.5)
Graduated from 2-year college	15(9.1)	10 (12.0)	5 (6.2)
Graduated from 4-year college or university	48(29.3)	24 (28.9)	24 (29.6)
Some graduate or professional school	7(4.3)	1 (1.2)	6 (7.4)
Received masters, professional, or doctoral degree	23(14)	8 (9.6)	15 (18.5)

Table 2 (2.2) Feminism, Political Leaning, and Language Background

	Full Sample	Black Participants	White Participants
	n (%)	n (%)	n (%)
Participants (*N*)	164	83	81
Feminist *N* (%)			
Yes	36 (22)	14 (16.9)	22 (27.2)
No	102 (62.2)	52 (62.7)	50 (61. 7)
Maybe/unsure	26 (15.9)	17 (20.5)	9 (11.1)
Political Leaning: Economic *N* (%)			
Very liberal	27 (16.5)	9 (10.8)	18 (22.2)
Liberal	31 (18.9)	23 (27.7)	8 (9.9)
Somewhat liberal	24 (14.6)	17 (20.5)	7 (8.6)
Moderate	31 (18.9)	18 (21.7)	13 (16)
Somewhat conservative	21 (12.8)	8 (9.6)	13 (16)
Conservative	16 (9.8)	2 (2.4)	14 (17.3)
Very conservative	13 (7.9)	6 (7.2)	7 (8.6)
Political Leaning: Social *N* (%)			
Very liberal	34 (20.7)	16 (19.3)	18 (22.2)
Liberal	36 (22.0)	26 (31.3)	10 (12.3)
Somewhat liberal	24 (14.6)	12 (14.5)	12 (14.8)
Moderate	23 (14.0)	11 (13.3)	12 (14.8)
Somewhat conservative	20 (12.2)	8 (9.6)	12 (14.8)
Conservative	14 (8.5)	5 (6.0)	9 (11.1)
Very conservative	12 (7.3)	5 (6.0)	7 (8.6)
English as a first language? *N* (%)			
Yes	163 (99.4)	83 (100.0)	80 (98.8)
No	1 (.6)	0 (0.0)	1 (1.2)
Speak a language other than English? *N* (%)			
Yes	31 (18.9)	13 (15.7)	18 (22.2)
No	133 (81.1)	70 (84.3)	63 (77.8)

Table 3 (3.3) Preferred Gender Term

	Full Sample	Black Participants	White Participants
	n (%)	n (%)	n (%)
Preferred gender term: Women *n* (%)	84	47	37
Woman	70 (83.33)	38 (80.85)	32 (86.49)
Female	12 (14.29)	7 (14.89)	5 (13.51)
Girl	1 (1.2)	1 (2.13)	0 (0.0)
Lady	1 (1.2)	1 (2.13)	0 (0.0)
Preferred gender term: Men *n* (%)	78	35	43
Man	63 (80.77)	29 (82.86)	34 (79.07)
Male	13 (16.67)	5 (14.29)	8 (18.60)
Boy	0 (0.00)	0 (0.00)	0 (0.00)
Gentleman	2 (2.56)	1 (2.6)	1 (2.33)

Design

I used a within-subjects design with two different contexts (dating and errands), two "word genders" (i.e., whether the term was feminine [*female, woman, girl, lady*] or masculine [*male, man, boy, gentleman*]), and four "word types" (f*emale/male, woman/man, girl/boy, lady/gentleman*). A fifth word type (*person*) was analyzed separately from this fully balanced 2x2x4 design.

Materials and Procedure

Participants were tested individually by means of a Qualtrics survey. After completing a consent form, participants were presented with the Speaker Judgment task, in which they evaluated a series of researcher-generated quotes. They then completed the Social Constructionist and Essentialist Beliefs Scale (SCEBS) and the Ambivalent Sexism Inventory (ASI), and lastly a set of demographic questions.

Speaker Judgment Task. I generated 18 quotes (see Appendix A). Each quote included one of the following nine words: *woman, man, female, male, girl, boy, lady, gentleman, person.* There was a feminine and masculine version of each gendered word: *woman/man*, *female/male*, *girl/boy*, and *lady/gentleman*, allowing for examination of whether the “feminine" and "masculine" versions of a term are perceived similarly. I also included the term *person* as a gender-neutral term to compare to the gendered labels.

Each quote referenced one of two contexts: errands or dating. For the “errands” context, participants read about people describing an everyday event such as going to the bank or driving to work. These quotes were intended to sound like snippets of a conversation you might overhear of someone describing everyday activities that would be typical for many people and thus serve as a neutral context. For the “dating” context, participants read about people describing a dating preference or experience, such as being set up on a date. For both contexts, the quotes were equivalent in length and designed to be equivalent in valence (no quote was intended to be overtly positive or negative about either the speaker or the person they were describing), in order to minimize effects of the statement itself and instead allow for assessing effects of the gendered label used. The quotes were presented in a randomized order. Across participants, each term was rotated through all 18 quotes, using a Latin Square design.

Participants were provided with the following instructions:

> “You will read a series of quotes. Each screen will include one quote. Each quote is from a different person. We ask you to provide your first impression of who the speaker is most likely to be. For example, you can include age, gender, race, personality – whatever you think is relevant. We ask that you include at least 2-3 descriptors for each quote. You do not need to write complete sentences.”

Then, on each screen participants saw a single quote, for example:

> "I was set up on a blind date with a female my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation."

They then received the following instructions:

> "**Describe the speaker who said the quote** based on your first impressions, your best guess as to who the speaker is most likely to be. For example, you can include age, gender, race, personality – whatever you think is relevant. Include at least 2-3 descriptors for each quote. You do not need to write complete sentences."

Before the start of the tasks, participants completed an attention check where they had to answer the question "Is each quote from a different person?" If participants answered "Yes," they were allowed to continue to the next task. If participants answered "No," they were provided with the instructions again and then asked the same question. If they answered incorrectly a second time, they were automatically directed to an exit screen and excluded from the survey.

Social Constructionist and Essentialist Beliefs Scale (SCEBS) (Schudson & Gelman, 2023). Participants completed the Social Constructionist and Essentialist Beliefs Scale, which consists of 16 items on a 7-point scale from 1 ("Strongly disagree") to 7 ("Strongly agree"). SCEBS measures three aspects of beliefs and attitudes about gender: essentialism, social constructionism, and realism. Sample items include: "If society were different, the categories 'women' and 'men' could be very different than they are" (social constructionism), "There are innate differences in attitudes and behaviors associated with different genders" (essentialism), and "Gender is a real category" (realism). In the present study, alphas were as follows: Social Constructionism: .91, Essentialism: .82, Realism: .74.

Short Version of the Ambivalent Sexism Inventory (ASI) (Glick & Whitehead, 2010). Participants completed a short version of the Ambivalent Sexism Inventory, which consists of 12 items on a 5-point scale from 1 ("Strongly disagree") to 5 ("Strongly agree"). ASI measures benevolent and hostile sexism. Items include "Many women have a quality of purity that few men possess" (benevolent sexism), and "Many women get a kick out of teasing men by seeming sexually available and then refusing male advances" (hostile sexism). In the present study, benevolent sexism alpha was .88 and hostile sexism alpha was .90.

Coding

Two undergraduate research assistants and I coded for the following: a) age was mentioned [0,1], b) age in years, c) gender was mentioned [0,1], d) gender term provided, e) race, ethnicity, or nationality was mentioned [0,1], f) race, ethnicity, and nationality term provided, g) economic level was mentioned [0,1], h) economic level provided [1 = working class/low income, 2 = middle class, 3 = upper class/rich, 4 = other (e.g., specific job type mentioned)], i) education level was mentioned [0,1], j) education level provided [1 = high school education or less/ trade school, 2 = 4 year degree/ some college 3 = graduate or professional school], k) positive and l) negative comments about the speaker. When an attribute was mentioned, it received a '1'; otherwise, it received a '0'. See Table 2.4 for an example of a coded response. For each coding category, I coded 20% of responses along with two independent coders and calculated Kappas for each coder and myself (pairwise). All the Kappas had at least "substantial" (0.63–0.80) levels of interrater reliability, and most had "near perfect" (0.81 and above) levels (Landis & Koch, 1977), except for the code "socio-economic level" which had values of .45 and .59. For this reason, the socio-economic level code was not analyzed further.

Table 4 (2.4) Coding Example

"White female, early 50's, and inattentive."

Item coded	Code
Age was mentioned	1
Age in years	51
Gender was mentioned	1
Gender term provided	Female
Race, ethnicity, or nationality was mentioned	1
Race, ethnicity, or nationality term provided	White
Economic level was mentioned	0
Economic level provided	(left blank if economic level was not mentioned)
Education level was mentioned	0
Education level provided	(left blank if education level was not mentioned)
Positive	0
Negative	1

Results

Below, I present the results of the Speaker Judgment task and how these relate to the scales and demographics. Initial analyses with participant gender as a factor revealed no interactions between gender and word, so I removed participant gender as a factor from the reported analyses. I did not control for family-wise error rate in the analyses given the exploratory nature of this study (Nicholson et al., 2022). Since both studies employed a novel experimental manipulation, it was important to maximize statistical power to see the small effects which were expected.

Speaker Judgment Task

For each coding dimension (mention of speaker gender, mention of speaker age, etc.), I conducted a repeated-measures ANOVA, including word type (4: *female/male*, *woman/man*, *girl/boy*, *lady/gentleman*), word gender (2: feminine, masculine), and context (2: dating, errands).

Positive/negative responses. I separately examined positive responses and negative responses as a function of word type, word gender, and context. There were no significant effects involving positive comments. For negative comments, there was a significant main effect of context, with participants including more negative comments in the errands context ($M = .18$) than the dating context ($M = .14$), $F(1, 162) = 12.33$, $p < .001$, $\eta_p^2 = .07$. There was also an interaction between context and word type, $F(3, 486) = 2.83$, $p = .04$, $\eta_p^2 = .012$. Post-hoc tests revealed that, in the errands context, participants included more negative comments when the imagined speaker used the terms *female/male* ($M = .22$) than *woman/man* ($M = .15$), $p = .012$. In contrast, there were no word type differences in the dating context.

Mention of speaker gender. I examined how often participants mentioned speaker gender, depending on the word type, word gender, and context. There was a significant main effect of context, with participants mentioning gender more in the dating context ($M = .86$) than in the errands context ($M = .78$), $F(1, 162) = 28.51$, $p < .001$, $\eta_p^2 = .15$. There was also a significant main effect of word type, $F(3, 486) = 4.11$, $p = .007$, $\eta_p^2 = .03$. Post-hoc tests revealed that participants mentioned gender more when the imagined speaker used the terms *female/male* ($M = .85$) than *woman/man* ($M = .82$, $p = .041$), *girl/boy* ($M = .80$, $p = .004$), or *lady/gentleman* ($M = .80$, $p = .011$).

Mention of speaker age. I examined how often participants mentioned the age of a speaker as a function of word type, word gender, and context. Participants mentioned speaker

age more in the dating context (M = .72) than in the errands context (M = .69), $F(1, 162) = 6.06$, $p = .015$, $\eta_p^2 = .04$. They also mentioned age slightly more for masculine terms (M = .71) than feminine terms (M = .69), $F(1, 162) = 4.85$, $p = .029$, $\eta_p^2 = .03$. Participants also mentioned age more for some word types than others, $F(3, 486) = 5.10$, $p = .002$, $\eta_p^2 = .03$. Post-hoc tests revealed that participants mentioned age more often when the imagined speaker used the terms *lady/gentleman* (M = .72) than when they used the terms *female/male* (M = .67), $p = .002$. They also mentioned age more often when the imagined speaker used the terms *girl/boy* (M = .73) than *female/male* (M = .67, $p < .001$) or *woman/man* (M = .69, $p = .039$). Lastly, there was an interaction between context and word gender, $F(1, 162) = 4.01$, $p = .047$, $\eta_p^2 = .02$, such that for the errands context, participants provided an age more for the masculine terms (M = .71) than for the feminine terms (M = .66), but there was no effect in the dating context.

Perceived speaker age (in years). I next examined how the perceived age of a speaker reflected the gender term presented. Note that these analyses only included those trials on which a speaker spontaneously mentioned age. I found a main effect of context; participants thought the speaker was younger in the dating context (M = 32.02) than in the errands context (M = 36.58), $F(1, 58) = 35.61$, $p < .001$, $\eta_p^2 = .38$. There was also a main effect of word type, $F(3, 174) = 7.27$, $p < .001$, $\eta_p^2 = .11$. Participants thought the speaker was older when the imagined speaker used the terms *lady/gentleman* (M = 37.05) than *female/male* (M = 33.63), *woman/man* (M = 33.55), or *girl/boy* (M = 32.97), all ps < .001. There was also a significant interaction between context and word type, $F(3, 174) = 16.28$, $p < .001$, $\eta_p^2 = .22$, indicating that participants thought the speaker was older when using the terms *lady/gentleman* than any of the other terms and regardless of context, all ps < .001.

Mention of the speaker's race, ethnicity, or nationality. I also examined how often participants mentioned the speaker's race, ethnicity, and/or nationality as a function of word type, word gender, and context. Race, ethnicity, or nationality were mentioned more in the dating context (M = .31) than the errands context (M = .28), $F(1, 162) = 4.47$, $p = .036$, $\eta p^2 = .02$. There was also an interaction between word gender and word type $F(3, 485) = 4.13$, $p = .007$, $\eta p^2 = .03$, indicating that race, ethnicity, and/or nationality were mentioned more for *female* (M = .32) than *male* (M = .27), $p = .045$).

Heteronormativity analyses. I examined heteronormativity in participants' responses by coding how often participants reported that the speaker's gender differed from the gender of the person mentioned in the quote. For example, if the person mentioned in the quote was a woman, and the participant indicated that the speaker was a man, they received a '1'. I conducted a repeated-measures ANOVA, including word type (4: *female/male*, *woman/man*, *girl/boy*, *lady/gentleman*), word gender (2: feminine, masculine), and context (2: dating, errands). There was a main effect of context, $F(1, 69) = 57.37$, $p < .001$, $\eta_p^2 = .45$, such that participants provided more heteronormative responses in the dating context (M = .94) than in the errands context (M = .77). However, it should be noted that heteronormative responses were very common in both contexts. There was also an interaction between context and word type $F(3, 207) = 2.81$, $p = .041$, $\eta_p^2 = .04$. Post-hoc analyses revealed that, in the errands context, *female/male* (M = .83) elicited more heteronormative responses than *woman/man* (M = .71), p = .01. By contrast, in the dating context, *girl/boy* (M = .96) elicited more heteronormative responses than *female/male* (M = .91), $p = .03$.

Comparison of Black and White Participants' Responses on the Speaker Judgment task

I examined the dependent variables by sample to determine how the two groups of participants (Black and White) compared, and here report only those effects involving race. The only significant effect involved the negative responses (i.e., how often participants wrote something negative about the speaker). Specifically, I conducted a repeated-measures ANOVA including word type, word gender, and context on participants' negative responses, and obtained a 3-way interaction involving word type, context, and sample, $F(3, 483) = 2.86$, $p = .039$, $\eta_p^2 = .02$. To examine this interaction further, I conducted post-hoc t-tests and found that in the dating context, White participants ($M = .19$) responded more negatively to the use of the terms *girl/boy* than Black participants ($M = .07$), $p = .003$. There were no other significant differences between the samples.

Comparing Gendered Terms to "Person"

As an exploratory analysis, I also compared the responses involving the gender-neutral term *person* to responses involving each of the gendered terms. I conducted a repeated measures ANOVA, including context (2: dating, errands) and word type (5: *female/male*, *woman/man*, *girl/boy*, *lady/gentleman*, and *person*). To avoid redundancy with the analyses above, here I only report those results involving significant effects of *person*.

Mention of speaker gender. I examined how often participants mentioned gender as a function of word type and context. There was a main effect of word type, $F(4, 652) = 23.81$, $p < .001$, $\eta_p^2 = .13$. Post-hoc tests revealed that gender was mentioned less for *person* ($M = .68$) than *female/male* ($M = .85$, $p < .001$), *woman/man* ($M = .82$, $p < .001$), *girl/boy* ($M = .80$, $p < .001$), or *lady/gentleman* ($M = .81$, $p < .001$). There was also an interaction between context and word type, $F(4, 652) = 4.23$, $p = .002$, $\eta_p^2 = .03$. In the dating context, gender was mentioned less for

person than for all the other word types, *p*s < .001; and in the errands context, they mentioned gender less for *person* than every other word type except *girl/boy* (*p* = .056), *p*s < .02.

Mention of speaker age. I examined how often participants mentioned the age of a speaker as a function of word type and context. I found a main effect of word type, indicating that participants mentioned speaker age more when the imagined speaker used the terms *girl/boy* (*M* = .73) than when they used the term *person* (*M* = .68), $F(4, 652) = 3.66$, $p < .006$, $\eta_p^2 = .02$.

Perceived speaker age (in years). I examined the perceived age of the speaker as a function of the word type and the context. There was a main effect of word type, such that participants perceived the speaker using the term *person* to be younger (*M* = 33.40) than the person using the terms *lady/gentleman* (*M* = 37.03), and older than the person using the terms *girl/boy* (*M* = 26.83), *p*s < .001.

How Participant Scores on ASI, SCEBS, And Feminism Related to Their Judgments of the Speakers

First, I conducted Pearson correlations between: (1) positive and (2) negative judgments of speakers using the terms *female* and *male* and: ASI-hostile sexism, ASI- benevolent sexism, SCEBS-essentialism, SCEBS-social constructionism, and SCEBS-realism. I found a negative correlation between ASI-hostile sexism and negative judgments for use of the terms *female* and *male*, $r(164) = -.17$, $p = .025$. The more a participant endorsed hostile sexism, the less negatively they judged the use of the terms *female* and *male*. There were no other significant correlations between positive or negative judgments of the imagined speaker and the remaining scales. Next, I examined how (1) positive and (2) negative judgments of those using the terms *female* and *male* related to participants' self-identified feminism. For this analysis, I combined the number of responses of participants who said "Yes" (*n* = 36) and "Maybe/not sure" (*n* = 26) and

compared those to participants who answered "No" to the question of whether they identify as feminist (n = 102). This was done in order to have sufficient numbers of participants in the two groups and is consistent with my prior research (Troncoso et al., 2022). I conducted an independent-samples t-test to compare participants' feminism score and their positive minus negative evaluations of the words *female* and *male*. The results indicated no significant difference in this difference score between non-feminists (M = .13, SD = .29) and feminists (M = .04, SD = .39), $t(162) = 1.69$, p = .093.

Next, I conducted Pearson correlations between heteronormativity scores for the dating and errands contexts (separately) and: ASI-benevolent sexism and ASI-hostile sexism, SCEBS-essentialism, SCEBS-social constructionism, SCEBS-realism, and self-described feminism. This revealed significant correlations between heteronormativity dating scores and SCEBS-realism, r (157) = .28, p < .001, and heteronormativity errand scores and SCEBS-realism, r (151) = .22, p = .008. In other words, the more the participants viewed gender as "real" the more heteronormative their responses for both the dating and errands contexts.

How Participant Scores on ASI, SCEBS, and Feminism Related to One Another

There was a strong positive correlation between SCEBS-essentialism and ASI-hostile sexism, r (164) = .84, p < .001, and SCEBS-essentialism and ASI-benevolent sexism, r (164) = .85, p < .001, showing that higher essentialism scores corresponded to higher levels of ambivalent sexism. SCEBS-realism correlated positively with: SCEBS-essentialism, r (164) = .53, p < .001, ASI-benevolent sexism, r (164) = .32, p < .001, but not ASI-hostile sexism, r (164) = .12, p = .132. SCEBS-realism correlated negatively with SCEBS-social constructionism, r (164) = -.52, p < .001. In other words, the more participants viewed gender as "real" (as

measured by SCEBS-realism), the more likely they were to endorse essentialist thinking, and sexist attitudes, and the less likely they were to endorse social constructionist views of gender.

There was a negative correlation between feminist identity and: SCEBS-essentialism, r (164) = -.27, $p < .001$, and SCEBS-realism, r (164) = -.28, $p < .001$, and a positive correlation between feminism and SCEBS-social constructionism, r (164) = .31, $p < .001$. In other words, the more feminist a participant was, the less they endorsed essentialist and realist beliefs about gender, and the more they endorsed social constructionist beliefs about gender. Lastly, there was a negative correlation between feminism and ASI-hostile sexism, r (164) = -.34, $p < .001$ but not ASI-benevolent sexism, r (164) = -.14, $p = .077$, indicating that the more feminist the participant identified as, the less likely they were to endorse hostile sexist attitudes.

Discussion

Study 1 examined if a speaker's choice of gendered term, such as *female* or *woman,* conveys messages to listeners about the speaker's identity, as well as their beliefs and attitudes about sex and gender. Participants were prompted to provide a description of who they imagined saying a quote that included a gendered (or gender-neutral) term. I found a number of indications that participants are sensitive to a speaker's word choice when making social inferences.

Did participants make inferences about a speaker based on word choice? A primary goal of this study was to examine comparisons across the type of word used: *female/male*, *woman/man*, *girl/boy*, and *lady/gentleman*. To this end, I found several ways that a speaker using the terms *female* or *male* elicited inferences or evaluations from participants. Participants included a description of the speaker's gender or sex more when they read a quote in which the speaker used the terms *female* or *male* than any of the other gender terms. This may be due to the term's close ties to biological sex, which may heighten attention to gender differences (Schudson

et al., 2019). Relatedly, participants provided more heteronormative responses for the terms *female/male* than for *woman/man* in the errands context. By contrast, they provided fewer heteronormative responses for the terms *female/male* than *girl/boy* in the dating context.

Participants also had more negative responses about a speaker who used the terms *female/male* in the errands context than in the dating context. This is in line with prior research finding that the terms *female/male* are more scientific or formal and less appropriate for casual use such as describing errands (Schudson et al., 2019; Troncoso et al., 2022). At the same time, the terms *female* and *male* were found to heighten attention to gender (perhaps because these terms are biological and thus draw attention to the body), which may be why they were judged to be more appropriate in the dating context. Lastly, participants mentioned race, ethnicity, and/or nationality more when the speaker used the term *female* than the term *male*. This is an interesting finding when thinking about the intersection of gender and race and who may be more racialized; however, race, ethnicity, and/or nationality was one of the less mentioned factors in the open-ended responses. More research would be needed to examine how gendered terms elicit inferences about a speaker's race.

I also found differences in terms of context. Participants mentioned gender, age, race/ethnicity and nationality more in the dating context, and they perceived the speaker to be younger, mentioned SES less frequently, and provided fewer negative comments about the speaker than in the errands context. Personal characteristics are likely viewed as more relevant in a dating context, where many people evaluate potential partners based on attributes such as these. These findings point to the importance of context in the interpretation of gendered language.

Did participants' own attitudes relate to their scores on the speaker judgment task? I analyzed how participant's scores on SCEBS, ASI, and feminism related to their negative and

positive judgments of the speaker. I found a negative relation between ASI-hostile sexism and negative judgments; the more a participant endorsed hostile sexism the less negative were their judgments about the speaker using the terms *female* and *male*. This finding suggests that these words have a sexist connotation. None of the other scales showed significant results.

How did the samples compare? Due to historical and current differences in how Black and White Americans are gendered and socialized (Brassel et al., 2020; Settles et al., 2008), I also examined if there were any differences by participant race. I only found one significant difference: White participants had more negative responses to the terms *girl* and *boy* than Black participants, specifically in the dating context. In the case of the term *boy*, I found this result surprising, given the historical context of the use of the term *boy* as a racial epithet to refer to Black American men (Bennet-Alexander, 2010; Henderson, 2003) and indeed would have predicted the opposite. However, participants were not provided with information about the race of the speaker or the race of the person being referenced. Adding this information could lead to a fuller picture of what terms are offensive and how this may depend on who are the people in the interaction. Another possible interpretation is that because this was only the case in the dating context, it could be that in this context, many participants may have assumed that the speaker and the person being described were of the same race, therefore reducing vigilance regarding historically racist use of terms. It would be important to replicate this result before drawing strong conclusions. The use of the term *girl* has been found to be infantilizing to women (MacArthur, 2015) and thus it makes sense that it would elicit negative responses, but it is not clear why White participants would have a more negative response to it than Black participants. It could be the case that Black participants, because of structural racism, experience more overtly

negative interactions, such as being called slurs, and in comparison, a generally non-overtly offensive terms such as "girl" would not seem as a threat (Davis et al., 2024).

How did use of gendered terms compare to "person"? I also compared use of gendered terms to the gender-neutral term *person* for each dependent variable and found differences in mention of speaker gender, mention of speaker age, and perceived speaker age. As expected, gender was mentioned less for *person* than any other gendered term overall and even when factoring in context, gender was mentioned less for *person* than any term in any context except for *girl/boy* in the errands context. This is an important finding as we think about issues of gender-fair language, suggesting that gender-neutral language may reduce emphasis on gender in written language.

Did participants make assumptions about heteronormativity? I also conducted an exploratory analysis of heteronormativity to examine if participants assumed that the speaker was of a gender other than that of the person they were speaking about. I found that participants provided more heteronormative responses in the dating context than in the errands context, but overall, heteronormativity was present for both contexts at high levels. It makes sense that heteronormativity would be higher in the dating context given that we live in a heteronormative society where heterosexual relationships are privileged. It is notable, however, that this was found even in the errands context, where there is not a clear reason to believe that the speaker and person referenced would differ in their gender. Future research could explore the reasons for this effect.

Limitations

The open-ended method of this study had both benefits and limitations. On the positive side, it allowed participants to choose which aspects of the speaker to focus on, and participants

provided some rich and informative descriptions (although these were not formally analyzed beyond coding for positive or negative content). For example, one participant (a White woman) wrote in reference to the word *female* in a dating context, "This language is really stilted and awkward - and the social posture expressed is pretty loner behavior based. The personality is probably a little male dominant mindset with not great respect for women." Another (a Black woman) wrote in reference to the word *female* in a dating context, "Single white female, lesbian, 28---- straight men don't talk that way." Another (a Black man) wrote in reference to the word *female* in an errands context, "Male, 30, Black, he seems outgoing, traditional gender roles, chauvinistic."

On the negative side, the gendered term used by each speaker was just one word presented in a longer vignette, which appeared to have led many participants to focus on the content of the vignette rather than the term itself, as in the following example:

> "Gender: The speaker uses the word "female" to describe the person they bumped into, so they are likely male. Age: The speaker seems to be in their 20s or 30s, as they are going to a store to buy drinks before a gathering. I would guess that they are in their early 30s. Personality: The speaker seems to be clumsy and accident-prone. They tripped and stumbled into a woman, but they apologized and continued shopping."

Another limitation of this method is that I only had data on what people chose to talk about, thereby limiting data on some key dimensions of interest, especially participants' beliefs about the speaker's attitudes toward women. For example, often participants wrote something like "40 year old man" and did not provide any other descriptions. Study 2 is designed to address these limitations.

Chapter III (Study 2)

For Study 1, I designed an open-ended task where participants read a series of quotes, each purported to be said by a different person, and then wrote a short response for each, describing who they imagined as the speaker. Each quote included one of the following terms: *woman, man*, *female, male*, *girl, boy*, *gentleman, lady*, or *person.* This method allowed me to examine how people's judgments vary across a range of different words but had limitations. A key limitation of Study 1 is that participants saw each word only once (i.e., each speaker only used the word once) and this was a potentially weak signal of how the speakers typically refer to others. Additionally, given the varied scenarios, participants may have made inferences related to the content of the quotes rather than the use of the gender term. For example, for the quote that involved the speaker bumping into someone, many participants described the speaker as clumsy. Lastly, the dependent variable being an open-ended description was also a limiting factor. Although the open-ended descriptions meant I could see what participants chose to focus on, it also meant that if they did not mention something I did not have information for that dimension. For example, if participants did not mention age, I had no information about what they inferred about the speaker's age. If they did not mention anything about the speaker's attitudes, I had no information about this.

The goals of Study 2 are to present a stronger signal of how the hypothetical speaker talks and to assess a wider range of judgments about the speaker, consistently across the whole sample. To accomplish this, I focused on the linguistic contrast that is my main interest: the use

of the terms *woman* vs. *female*. More specifically, I designed a task in which each speaker said the same term (either *woman* or *female*) four different times (four different quotes describing different scenarios). In this way, the participant has evidence that the term used is characteristic of how that person speaks about women. In other words, Study 1 allowed us to see how participants respond to a wider range of words, while Study 2 allows us to see how participants respond to the way a specific person speaks and thus strengthen the signal coming from the speaker's word choice. Lastly, in Study 2, participants were asked closed-ended questions, allowing for a more systematic assessment of every dimension and a clearer image of what kind of person comes to mind with each term. My research questions in Study 2 were the following:

1) Do we make inferences about people based on the words they use to refer to women? Specifically, if a speaker *repeatedly* refers to women as "a female" instead of as "a woman" (indicating that this is a typical or characteristic way that the person speaks), does this affect participants' perceptions of the speaker's identity (gender, race, and age), traits (likability, warmth, competence, education level, formality, and politeness), impression of how much they respect the women they refer to, and attitudes and beliefs (feminist identity, political affiliation)? Does it affect participants' warmth toward the speaker? I predicted that participants would have more negative feelings about the speaker repeatedly using the term *female* than the speaker repeatedly using the term *woman*, such that the speaker using the females would be rated as less likable, warm, competent, educated, formal, and polite than the speaker using the term *woman*. I also predicted they would be rated as less feminist, less liberal, and less respectful, and that participants would feel less warmly toward them. Given the findings of Study 1 on heteronormativity, I also predicted that participants would be more likely to assume that the speaker was a man. I did not have specific predictions about perceptions of age and race.

2) Are our inferences of the speaker influenced by the context in which they make linguistic choices? Specifically, how does perception of the speaker vary if they refer to women as *female* (or as *woman*) while describing a dating situation versus while describing an everyday activity such as running an errand? I am positing the "errands" context as a neutral comparison to the "dating" context, for which I predict the use of a gender term will be more salient. Based on the findings from Study 1, I expect that participants will have more negative perceptions of the speaker who uses the term *female* in the errands context than the dating context.

Finally, I was interested in how participants' evaluations of a speaker related to their own beliefs about gender (as measured by the Social Constructionist and Essentialist Beliefs Scale (SCEBS), the Ambivalent Sexism Inventory (ASI), and the Feminist Identification Scale (FIS)). I predicted that more essentialist, more sexist, and less feminist participants would have more positive views of the speaker saying *female* relative to the speaker who said *woman*, than participants who were less essentialist, less sexist, and more feminist.

Method

I obtained IRB approval for this study by the University of Michigan and I pre-registered my predictions, method, and analysis plan with AsPredicted (#210478).

Participants

I recruited 600 participants from Prolific. I used this platform on the recommendation of colleagues, who prefer it to MTurk. A power analysis for a repeated measures within-between interaction ANOVA indicated that the minimum sample size to yield a statistical power of at least .95 with an alpha of .05 and a small effect size ($f = 0.14$) derived from a previous study is 288 (Faul et al., 2009). The final sample consisted of 596 participants. Four participants were automatically excluded due to incomplete or incorrect completion codes, which resulted in their

data not being imported to Qualtrics. Participants were not excluded for any other reason. The final sample consisted of 299 White participants and 297 Black participants. Participants were asked about their gender, their sexual orientation, and their age in an open-ended format. They were also asked about their race/ethnicity, highest level of education, annual family income, feminism identification, and political affiliation. Lastly, participants were asked what they thought the survey was about and if there was anything they would like to share with the researchers. See Table 3.1 for complete demographics.

Table 5 (3.1) Demographics

Participant Demographics	Full Sample	Black Participants	White Participants
	n (%)	*n (%)*	*n (%)*
Gender			
Woman	340 (57)	177 (60)	163 (55)
Man	239 (40)	107 (36)	132 (44)
Non-binary	4 (1)	2 (1)	2 (1)
Transmasculine	1 (0)	1 (0)	0 (0)
Another option	12 (2)	10 (3)	2 (1)
Sexual Orientation			
Asexual/demisexual/aromantic	8 (1.34)	2 (0.67)	6 (2.01)
Bisexual/pansexual	56 (9.40)	20 (6.73)	36 (12.04)
Gay/lesbian	18 (3.02)	9 (3.03)	9 (3.01)
Heterosexual	460 (77.2)	236 (79.5)	224 (74.92)
Queer	2 (0.33)	2 (0.67)	0 (0)
Nonresponse/unsure	8 (1.34)	1 (0.34)	7 (2.34)
Gender/sex term	44 (7.38)	27 (9.1)	17 (5.7)
Education			
Completed some high school	6 (2)	4 (1.3)	2 (0.7)
Graduated from high school	82 (27.5)	41 (13.8)	41 (13.7)
Completed some college	118 (39.6)	63 (21.2)	55 (18.4)
Finished training other than college (e.g., vocational school)	22 (7.4)	16 (5.4)	6 (2)
Graduated from 2-year college	53 (17.7)	23 (7.7)	30 (10)

Graduated from 4-year college or university	201 (67.4)	91 (30.6)	110 (8.1)
Completed some graduate or professional school	27 (9.1)	18 (6.1)	9 (3.0)
Received masters, professional, or doctoral degree	87 (29.2)	41 (13.8)	46 (15.4)
Family yearly income			
Less than $15,000	42 (14.1)	19 (6.4)	23 (7.7)
$15,000 - $25,000	45 (15.1)	19 (6.4)	26 (8.7)
$25,000 - $35,000	50 (16.8)	26 (8.8)	24 (8)
$35,000 - $50,000	85 (28.5)	50 (16.8)	35 (11.7)
$50,000 - $75,000	139 (46.7)	75 (25.3)	64 (21.4)
$75,000 - $100,000	84 (28.2)	44 (14.8)	40 (13.4)
$100,000 - $150,000	87 (29.2)	46 (15.5)	41 (13.7)
Above $150,000	64 (21.5)	18 (6.1)	46 (15.4)
Political Affiliation			
(1) Liberal	97 (32.6)	49 (16.5)	48 (16.1)
2	111 (37.3)	57 (19.2)	54 (18.1)
3	61 (20.4)	31 (10.4)	30 (10)
4	163 (54.7)	96 (32.3)	67 (22.4)
5	68 (22.8)	26 (8.8)	42 (14)
6	42 (14.1)	14 (4.7)	28 (9.4)
(7) Conservative	54 (18.1)	24 (8.1)	30 (10)

Design

Study 2 includes word (2 levels: *female*, *woman*) as the only within-subject factor, and three between-subject factors: context (2 levels: dating, errands), block order (2 levels: *female*-1st, *woman*-1st), and participant race (2 levels: Black, White). I counterbalanced the order in which participants saw the terms: roughly half saw the *female* block first, and roughly half saw the *woman* block first. Participants read quotes of someone speaking about dating experiences (dating context) or of someone speaking about everyday errands (errands context). Additionally, for each context, I created two sets of statements that the speaker produced, labeled as Set A and

Set B, respectively, and counterbalanced which set was associated with each word (*female* and *woman*) and block order. This ensured that there was not a confound between quote content and either word or block order. See Appendix C for a complete list of sentences.

Materials and Procedure

After completing a consent form, participants read the following instructions for the Speaker Snapshot task:

> **Instructions**: In this task you will be asked to provide your impressions of two different people (one at a time), after reading statements each person has made.
>
> First we will ask you to read a set of four (4) statements made by Speaker A and provide your impressions of Speaker A.
>
> Next we will ask you to read a set of four (4) statements made by Speaker B and provide your impressions of Speaker B.
>
> Be sure to answer the questions based on your first impressions of each speaker. There are no right or wrong answers; we are only interested in your impressions of each speaker.

For each block, after reading the four statements made by that speaker (either A or B), participants answered questions about that speaker. First, participants were asked to provide their impression of the speaker on a Likert scale from 1 (not at all) to 7 (very) on the following dimensions: likable, warm, competent, educated, formal, polite, feminist, and sexist. Then they were asked to provide their impression of the speaker's political affiliation on a scale from 1 (liberal) to 7 (conservative) and to rate, on a scale from 1 (not at all) to 7 (very), "How warm do you feel toward the speaker?" and "How much does the speaker respect the people they are talking about?" Lastly, participants were asked to provide their impression of the gender and age of the speaker (in an open-ended format) and the race of the speaker (in a multiple-choice format), and they were asked if there was anything else they would like to share about the speaker (short-response format). Participants completed SCEBS (alphas: Social Constructionism:

.90, Essentialism: .80, Realism: .59 and ASI (alphas: benevolent sexism: .83, hostile sexism: .92) and answered demographic questions about their own gender, sexual orientation, race, highest level of education completed, family's combined yearly income, and age. Realism was not analyzed further due to its low reliability.

I made a few minor changes in the wording of the demographic questions, compared to Study 1. Participants were asked to provide their gender with an open-ended question (rather than multiple choice), in order not to constrain their choices. To assess feminism, I used a more nuanced measure: the Self-Identification as a Feminist Scale (SIF) (Szymanski, 2004), a four-item measure asking participants to indicate the following on a scale of 1 (strongly disagree) to 5 (strongly agree): "1) I consider myself a feminist. 2) I identify myself as a feminist to other people. 3) Feminist values and principles are important to me. 4) I support the goals of the feminist movement." The alpha for this scale was .93. I used a more concise measure for political affiliation, a Likert scale from 1 (liberal) to 7 (conservative), in response to: "What is your political affiliation?" Lastly, because 99% of Study 1 participants spoke English as their first language and 81% did not speak another language, I did not ask about language in this study.

Results

I first present participants' evaluations of the speaker, measured by the responses on the Speaker Snapshot task, then their assumptions of the speaker's demographics, and finally how these relate to participants' own attitudes and beliefs (measured by the SCEBS, ASI, and SIF scales). Initial analyses with participant gender as a factor revealed no interactions between gender and word, so I removed participant gender as a factor from the reported analyses. As with

Study 1, I did not control for family-wise error rate in order to maximize statistical power (Nicholson et al., 2022).

Speaker Evaluations

I examined speakers' perceived characteristics and demographics as a function of word, context, block order, and participant race. For each dimension, I conducted an ANOVA with word (*female* or *woman*) as a within-subjects factor and context (dating or errands), block order (female-first, woman-first), and participant race (Black or White) as between-subjects factors. Below are the results. See Table 3.2 and Table 3.3 for all mean averages of the word-by-context interactions.

Table 6 (3.2) Context x Word Interactions (Perceived speaker traits)

Traits	Dating		Errands	
	Female	Woman	Female	Woman
	M	*M*	*M*	*M*
	(SD)	*(SD)*	*(SD)*	*(SD)*
Likable	4.69*	5.19*	3.85*	4.84*
	0.09	*0.07*	*0.09*	*0.07*
Warm	4.57*	5.11*	3.78*	4.77*
	0.09	*0.07*	*0.09*	*0.07*
Competent	4.69	5.23	4.21	4.90
	0.09	*0.07*	*0.09*	*0.07*
Educated	4.70	5.29	4.22	4.93
	0.09	*0.07*	*0.09*	*0.07*
Formal	4.44	4.58	3.99*	4.52*
	0.10	*0.08*	*0.10*	*0.08*
Polite	4.89*	5.33*	4.11*	5.26*
	0.09	*0.07*	*0.09*	*0.07*
Feminist	3.00*	3.27*	2.82*	3.50*
	0.10	*0.10*	*0.10*	*0.10*
Sexist	3.23*	2.46*	4.36*	2.72*
	0.11	*0.10*	*0.11*	*0.09*
Political	4.18*	3.97*	4.69*	3.89*
	0.08	*0.07*	*0.08*	*0.07*
Warmth toward speaker	4.22*	4.82*	3.40*	4.37*

	0.10	*0.08*	*0.10*	*0.08*
Respect of person	4.48*	5.17*	3.73*	4.98*
	0.10	*0.08*	*0.10*	*0.08*

*Within each context, asterisks indicate that ratings are different between *woman* and *female*

Table 7 (3.3) Context x Word Interactions (Perceived speaker demographics)

Demographics	Dating		Errands	
	Female	Woman	Female	Woman
	M	*M*	*M*	*M*
	(SD)	*(SD)*	*(SD)*	*(SD)*
Age	29.6	30.07	35.96*	34.68*
	0.53	*0.42*	*0.53*	*0.42*
Race	0.83	0.92	0.81	0.85
	0.02	*0.02*	*0.02*	*0.02*
Gender	0.92	0.96	0.86*	0.68*
	0.02	*0.02*	*0.19*	*0.02*

*Within each context, asterisks indicate that ratings are different between *woman* and *female*

How likable is the speaker? There was a main effect of word such that participants rated the speaker who used the word *woman* ($M = 5.02$) as more likable than the speaker who used the word *female* ($M = 4.27$), $F(1, 588) = 140.23$, $p < .001$, $\eta_p^2 = .19$. There was a main effect of context such that participants liked the speaker more overall in the dating context ($M = 4.94$) than in the errands context ($M = 4.34$), $F(1, 588) = 41.82$, $p < .001$, $\eta_p^2 = .07$. There was also a main effect of sample such that Black participants ($M = 4.90$) rated the speaker as more likable than White participants ($M = 4.39$), $F(1, 588) = 31.03$, $p < .001$, $\eta_p^2 = .05$. Lastly, there was a word by context interaction, $F(1, 588) = 14.77$, $p < .001$, $\eta_p^2 = .02$. In both contexts, participants found the speaker to be more likable when they used the word *woman* than *female* (*p*s < .001), but the difference was greater in the errands context than in the dating context. None of the other effects were significant.

How warm is the speaker? Similarly, there was a main effect of word such that the speaker who used the word *woman* ($M = 4.94$) was perceived to be warmer than the speaker who used the word *female* ($M = 4.17$), $F(1, 588) = 126.02$, $p < .001$, $\eta_p^2 = .18$. There was a main effect of context such that speaker was perceived to be warmer in the dating context ($M = 4.84$) than in the errands context ($M = 4.27$), $F(1, 588) = 37.56$, $p < .001$, $\eta_p^2 = .06$. There was a main effect of sample, $F(1, 588) = 40.36$, $p < .001$, $\eta_p^2 = .06$, such that Black participants ($M = 4.85$) found the speaker to be warmer than did White participants ($M = 4.26$). Lastly, there was an interaction of word by context, $F(1, 588) = 10.74$, $p = .001$, $\eta_p^2 = .02$. In both the dating and errands contexts, the speaker who used the term *woman* was perceived as warmer than the speaker who used the word *female* (ps $< .001$), but the difference was greater in the errands context. None of the other effects were significant.

How competent and educated is the speaker? There was a significant main effect of word such that the speaker was seen as more competent when using the word *woman* ($M = 5.06$) than *female* ($M = 4.45$), $F(1, 588) = 90.42$, $p < .001$, $\eta_p^2 = .13$. There was a significant main effect of context such that speaker was seen as more competent in the dating context ($M = 4.96$) than in the errands context ($M = 4.56$), $F(1, 588) = 18.60$, $p < .001$, $\eta_p^2 = .03$. There was also a significant main effect of sample, such that Black participants rated the speaker as more competent ($M = 4.99$) than did White participants ($M = 4.52$), $F(1, 588) = 25.15$, $p < .001$, $\eta_p^2 = .04$. None of the other effects for competence were significant.

Similarly, for education, there was a significant main effect of word such that the speaker was seen as more educated when using the word *woman* ($M = 5.11$) than *female* ($M = 4.46$), $F(1, 588 = 118.63$, $p < .001$, $\eta_p^2 = .17$. There was a significant main effect of context such that the speaker was seen as more educated in the dating context ($M = 4.99$) than the errands context (M

= 4.58), $F(1, 588) = 21.68$, $p < .001$, $\eta_p^2 = .04$. Lastly, there was a significant main effect of sample, such that Black participants ($M = 5.04$) rated the speaker as more educated than did White participants ($M = 4.53$), $F(1, 588) = 33.38$, $p < .001$, $\eta_p^2 = .05$. None of the other effects for education were significant.

How formal is the speaker? There was a significant main effect of word such that the speaker was seen as more formal when using the word *woman* ($M = 4.55$) than *female* ($M = 4.21$), $F(1, 587) = 22.01$, $p < .001$, $\eta_p^2 = .04$. There was a significant main effect of context such that the speaker was seen as more formal in the dating context ($M = 4.51$) than in the errands context ($M = 4.25$), $F(1, 587) = 6.21$, $p = .013$, $\eta_p^2 = .01$. There was also a significant main effect of sample such that Black participants ($M = 4.66$) rated the speaker as more formal than did White participants ($M = 4.11$), $F(1, 587) = 28.90$, $p < .001$, $\eta_p^2 = .05$. Lastly, there was a word by context interaction such that, only in the errands context, participants found the term *woman* ($M = 4.52$) to be more formal than *female* ($M = 3.99$), $p < .001$, $F(1, 587) = 7.30$, $p = .007$, $\eta_p^2 = .01$.

How polite is the speaker? There was a significant main effect of word such that participants found the speaker to be more polite when they used the word *woman* ($M = 5.30$) than *female* ($M = 4.50$), $F(1, 588) = 121.93$, $p < .001$, $\eta_p^2 = .17$. There was a significant main effect of context such that the speaker in the dating context ($M = 5.11$) was perceived as more polite than the speaker in the errands context ($M = 4.69$), $F(1, 588) = 21.15$, $p < .001$, $\eta_p^2 = .04$. There was a main effect of sample such that Black participants ($M = 5.11$) rated the speaker as more polite than White participants ($M = 4.68$), $F(1, 588) = 21.69$, $p < .001$, $\eta_p^2 = .04$. There was also a word by context interaction, $F(1, 588) = 24.21$, $p < .001$, $\eta_p^2 = .04$, ps $< .001$**,** indicating that, for both contexts, the speaker using the word *woman* was seen as more polite than the speaker using the word *female*, but the difference was greater in the errands context. Lastly, there was a significant

word by block order interaction, $F(1, 592) = 8.76$, $p = .003$, $\eta_p^2 = .02$, indicating that, for both block orders, the speaker using the word *woman* ($M = 5.25$ *female*-first; $M = 5.34$ *woman*-first) was seen as more polite than the speaker using the word *female (M = 4.67 female first*; $M = 4.33$ *woman*-first) but the difference was greater in the *woman*-first block order. This may indicate a contrast effect, whereby *female* sounded less polite when contrasted with *woman*.

How feminist is the speaker? There was a significant main effect of word such that the speaker was seen as more feminist when they used the word *woman* ($M = 3.39$) than *female* ($M = 2.91$), $F(1, 588) = 35.09$, $p < .001$, $\eta_p^2 = .06$. There was a word by context interaction, $F(1, 588) = 6.35$, $p = .012$, $\eta_p^2 = .01$, such that the speaker was seen as more feminist when they used the word *woman* ($M = 3.27$ dating, $M = 3.50$ errands) than *female* ($M = 3.00$ dating, $M = 2.82$ errands) in both contexts, ps $\leq .017$, but the difference was greater in the errands context. There was a word by participant sample interaction, $F(1, 588) = 4.16$, $p = .042$, $\eta_p^2 = .01$, such that both Black and White participants rated the speaker as more feminist when using the word *woman* ($M = 3.36$ Black participants, $M = 3.41$ White participants) than *female* ($M = 3.05$ Black participants, $M = 2.77$ White participants) but the difference was greater among White participants.

How sexist is the speaker? There was a main effect of word for perceived sexism such that speaker was perceived as more sexist when using the word *female* ($M = 3.80$) than *woman* ($M = 2.59$), $F(1, 588) = 186.02$, $p < .001$, $\eta_p^2 = .24$. There was also a main effect of context such that participants in the errands context ($M = 3.54$) were seen as more sexist than those in the dating context ($M = 2.84$), $F(1, 588) = 35.25$, $p < .001$, $\eta_p^2 = .06$. Lastly, there was also a word by context interaction, $F(1, 588) = 24.87$, $p < .001$, $\eta_p^2 = .04$. In both contexts, the speaker was

seen as more sexist when saying female than woman, but the difference was greater in the errands context.

How liberal or conservative is the speaker? There was a main effect of word such that the speaker was seen as more conservative when using the word *female* ($M = 4.44$) than the word *woman* ($M = 3.93$) , $F(1, 588) = 46.02$, $p < .001$, $\eta_p^2 = .07$. There was a main effect of context such that participants perceived the speaker in the dating context ($M = 4.07$) as more liberal than the speaker in the errands context ($M = 4.29$), $F(1, 588) = 8.54$, $p = .004$, $\eta_p^2 = .01$. There also a word by context interaction, $F(1, 588) = 15.55$, $p < .001$, $\eta_p^2 = .03$, such that speakers were seen as more liberal when saying the word *woman* ($M = 3.97$ dating, $M = 3.89$ errands) than *female* ($M = 4.18$ dating, $M = 4.69$ errands) but the effect was greater in the errands context. Lastly, there was a word by participant sample interaction, $F(1, 588) = 7.14$, $p = .008$, $\eta_p^2 = .01$, such that both Black participants and White participants rated the speaker as more liberal when using the word *woman* than when using the word *female*, but the effect was larger among White participants.

How warmly does the participant feel toward the speaker? There was a significant main effect of word such that participants felt more warmly toward the speaker who used the word *woman* ($M = 4.59$) than the speaker who used the word *female* ($M = 3.81$), $F(1, 588) = 111.65$, $p < .001$, $\eta_p^2 = .16$. There was also a significant main effect of context such that participants felt more warmly toward the speaker in the dating context ($M = 4.52$) than in the errands context ($M = 3.88$), $F(1, 588) = 38.57$, $p < .001$, $\eta_p^2 = .06$. There was also a main effect of sample such that Black participants ($M = 4.52$) felt more warmly toward the speaker than White participants ($M = 3.88$), $F(1, 588) = 39.52$, $p < .001$, $\eta_p^2 = .06$. There was also a word by context interaction**,** $F(1, 588) = 6.24$, $p = .013$, $\eta_p^2 = .01$, such that participants felt more warmly

toward the speaker when they used the term *woman* ($M = 4.82$ dating, $M = 4.37$ errands) than when using the word *female* ($M = 4.22$ dating, $M = 3.40$ errands), but the effect was greater in the errands context.

How much does the speaker respect the people they are talking about? There was a significant main effect of word such that participants thought the speaker was more respectful toward the people they were referring to when using the word *woman* ($M = 5.07$) than *female* ($M = 4.11$), $F(1, 588) = 133.31$, $p < .001$, $\eta_p^2 = .19$. There was a significant main effect of context, such that participants thought that the speaker in the dating context ($M = 4.82$) was more respectful toward the target than the speaker in the errands context ($M = 4.36$), $F(1, 588) = 22.30$, $p < .001$, $\eta_p^2 = .04$. There was also a significant main effect of sample such that Black participants ($M = 4.86$) thought that the speaker was more respectful than did White participants ($M = 4.32$), $F(1, 588) = 29.84$, $p < .001$, $\eta_p^2 = .05$. There was also an interaction of word by context, $F(1, 588) = 11.50$, $p < .001$, $\eta_p^2 = .02$. Participants in both contexts thought the speaker was more respectful toward the target when using the word *woman* (M = 5.17 dating, $M = 4.98$ errands) than when using the word *female* ($M = 4.48$ dating, $M = 3.73$ errands), but the effect was greater in the errands condition, ps $< .001$.

Demographics of the Speaker

Perceived gender of the speaker. I coded responses into three categories: feminine identity terms (e.g., *woman*, *female*, *girl*; 57%), masculine identity terms (e.g., *man*, *male*, *boy*; 40%), and other (e.g., *non-binary, transmasculine*; 3%). Due to how rare the "other" responses were, I only analyzed the feminine (coded as 0) and masculine (coded as 1) identity terms. There was a significant effect of word such that participants thought the speaker was more likely to be a man when using the word *female* ($M = .90$) than when using the word *woman* ($M = .82$), $F(1,$

552) = 14.74, $p < .001$, $\eta_p^2 = .03$. There was a main effect of context such that participants thought that speaker was more likely to be a man in the dating context ($M = .94$) than in the errands context ($M = .77$), $F(1, 552) = 60.71$, $p < .001$, $\eta_p^2 = .10$. There was also a significant interaction of word and context, $F(1, 552) = 34.43$, $p < .001$, $\eta_p^2 = .06$, such that the effects of word described above was found only in the errands context (*female* $M = .86$; *woman* $M = .68$), $p < .001$.

Perceived race of the speaker. Because most participants (86%) indicated that the speaker was either Black (coded as 0) or White (coded as 1), I only coded for these two racial categories. There was a significant effect of word such that participants thought the speaker was more likely to be White when using the word *woman* ($M = .88$) than when using the word *female* ($M = .82$), $F(1, 467) = 8.53$, $p = .004$, $\eta_p^2 = .02$. There was a significant main effect of sample such that White participants were more likely than Black participants to infer that the speaker was White (Ms = .93 and .77, respectively), $F(1, 467) = 49.63$, $p < .001$, $\eta_p^2 = .10$. There was also a significant interaction of word and block order, $F(1, 467) = 4.61$, $p = .032$, $\eta_p^2 = .01$, such that for the woman-first block, participants thought the speaker was more likely to be White when they used the word *woman* ($M = .92$) than when they used the word *female* ($M = .81$), $p < .001$.

Perceived age of the speaker. There was a main effect of context such that the speaker was perceived to be older in the errands context ($M = 35.32$) than in the dating context ($M = 29.84$), $F(1, 587) = 91.58$, $p < .001$, $\eta_p^2 = .14$. There was also an interaction of word by context, $F(1, 587) = 6.04$, $p = .014$, $\eta_p^2 = .01$, such that in the errands context the speaker was perceived to be older when using the word *female* ($M = 35.96$) than *woman* ($M = 34.68$), $p = .011$. No significant difference was found in the dating context. None of the other effects were significant.

Relations between Speaker Ratings and Participants' Beliefs and Attitudes

I next examined how speaker ratings related to participants' beliefs and attitudes. Here, my key interest was not in participants' ratings for "female" and "woman" considered separately, but rather the extent to which participants differentially evaluated the two words. For example, did participants who had relatively more negative evaluations of the speaker who used *female* (vs. the speaker who used *woman*) hold lower levels of essentialist beliefs, or higher levels of feminist identification?

In order to provide this test, I first calculated difference scores by subtracting each participant's rating of the use of the term *female* minus their rating of the term *woman,* separately for each dimension. A higher score indicated a higher rating of the term *female* than *woman* on that dimension. I then conducted Pearson correlations between each of the individual measures (SCEBS-Essentialism, SCEBS-Social Constructionism, ASI-hostile sexism and benevolent sexism and FIS composite score) and participants' difference scores on each of the following dimensions: likeable, warm, competent, educated, formal, polite, feminist, sexist, political affiliation, warmth toward speaker, respectfulness of the speaker toward target, speaker gender, speaker race, and speaker age.

SCEBS-Essentialism. I found that the higher participants scored on the SCEBS-essentialism scale, the higher the difference score on the dimensions of likable [r (596) = .17, p < .001], warm [r (596) = .14, p < .001], competent [r (596) = .10, p = .013], educated [r (596) = .14, p < .001], polite [r (596) = .17, p < .001], and feminist [r (596) = .15, p < .001]. In other words, participants who scored more highly on SCEBS-Essentialism rated the speaker as more likable, warm, competent, educated, polite, and feminist when using the term *female* than when using the term *woman*. Additionally, the more essentialist they were, the more warmly they felt

toward the speaker [r (596) = .12, p = .004] and the more respectful they thought the speaker was toward the target [r (596) = .15, p < .001] when using the term *female* relative to using the term *woman*. On the other hand, the more essentialist the participant, the less they found the speaker to be sexist [r (596) = -.14, p < .001] and the less they thought the speaker's political affiliation [r (596) = -.18, p < .001] was conservative, when using *female* versus *woman*. There were no significant correlations for judgments of formality, gender, age, or race.

SCEBS-Social Constructionism. I found that the more highly participants scored on the SCEBS-Social Constructionism subscale, the lower the difference score on the dimension of education [r (596) = - .08, p = .049]. In other words, participants who scored highly on social constructionism rated the speaker as less educated when using the term *female* than when using the term *woman*. There were no other significant correlations between social constructionism and any of the dimensions.

ASI-Hostile Sexism. The higher participants scored on the ASI-hostile sexism subscale, the higher the difference score on the following dimensions: likable [r (596) = .19, p < .001], warm [r (596) = .16, p < .001], competent [r (596) = .14, p < .001], educated [r (596) = .18, p < .001], polite [r (596) = .18, p < .001], feminist [r (596) = .26, p < .001], warmth toward the speaker [r (596) = .19, p < .001], and respect toward the target [r (596) = .19, p < .001]. In contrast, the higher the score on the ASI-hostile sexism subscale, the lower the difference score on perceived sexism of the speaker [r (596) = -.13, p = .001] and political affiliation [r (596) = -.22, p < .001]. In other words, participants who scored more highly on ASI-hostile sexism rated the speaker using the term *female* more positively (e.g., more likable and educated, less sexist) relative to the speaker using the term *woman*. None of the other correlations with ASI-hostile sexism were significant.

ASI-Benevolent Sexism. The higher participants scored on the ASI-benevolent sexism subscale, the higher the difference score on the following dimensions: likable [r (596) = .14, p < .001], warm [r (596) = .15, p < .001], competent [r (596) = .11, p = .008], educated [r (596) = .14, p < .001], polite [r (596) = .19, p < .001], feminist [r (596) = .22, p < .001], warmth toward the speaker [r (596) = .13, p = .001], and respect toward the target [r (596) = .17, p < .001]. In contrast, the higher the score on the ASI-benevolent sexism subscale, the lower the difference score on perceived sexism of the speaker, r (596) = -.10, p = .012 and political affiliation, r (596) = -.18, p < .001. In other words, participants who scored more highly on ASI-benevolent sexism rated the speaker using the term *female* more positively (e.g., more likable and educated, less sexist) relative to the speaker using the term *woman*. No significant correlation with gender, age, or race.

Feminist Identification Scale (FIS). The more highly participants scored on the FIS scale, the lower the difference score on the following dimensions: likable [r (596) = -.09, p = .023], competent [r (596) = -.13, p = .002], educated [r (596) = -.17, p < .001], formal [r (596) = -.10, p = .02], polite [r (596) = -.13, p < .001], feminist [r (596) = -.16, p < .001], warmth toward the speaker [r (596) = -.12, p = .005], and speaker respect toward the target [r (596) = -.12, p = .008]. In contrast, the higher the score on feminism, the higher the difference score on sexism [r (596) = .08, p = .048]. In other words, participants who identified as more feminist rated the speaker who used the term *female* more negatively (e.g., less likable and competent, more sexist) relative to the speaker using the term *woman*.

How SCEBS-Essentialism, SCEBS-Social Constructionism, ASI-Hostile Sexism, ASI-Benevolent Sexism, and FIS Relate to One Another

As expected, there was a negative correlation between SCEBS-essentialism and SCEBS-social constructionism [r (596) = -.10, p = .016], and between SCEBS-Essentialism and the feminism composite score [r (596) = -.25, p < .001]. There was also a positive correlation between SCEBS-Essentialism and ASI-hostile sexism [r (596) = .49, p < .001] and ASI-benevolent sexism [r (596) = .47, p < .001]. In other words, the more essentialist the participant, the more sexist, less feminist, and less likely to view gender through a social constructionist perspective. There was also a negative correlation between SCEBS-Social Constructionism and ASI-hostile sexism, r (596) = -.14, p < .001, but no correlation between ASI-benevolent sexism, r (596) = .01, p = .725. There was a positive correlation between SCEBS-Social Constructionism and feminism composite score [r (596) = .45, p < .001]. There was a negative correlation between ASI-hostile sexism and feminism composite score [r (596) = -.35, p < .001] and a positive correlation between ASI-hostile sexism and ASI-benevolent sexism [r (596) = .46, p <.001]. Lastly, there was a positive correlation between ASI-benevolent sexism and feminism composite score [r (596) = -.12, p = .003].

Differences in Individual Scale Responses between Black and White Participants

I conducted an independent samples t-test to examine if there were any differences between Black and White participants' individual difference scores. See Table 2.4 for results.

Table 8 (3.4) Individual Scale Responses as a Function of Participant Self-Identified Race

	Black participants	White participants	*t*	*p*
SCEBS-essentialism	4.61 (1.10)	4.20 (1.20)	-4.497	< .001
ASI-hostile sexism	1.91 (1.40)	1.52 (1.31)	-3.581	<.001
ASI-benevolent sexism	3.01 (1.06)	2.24 (1.20)	-8.291	>.001

Open-Ended Responses

After completing multiple choice questions about each speaker, participants were asked if there was anything else they'd like to share about the speaker in a short, open-ended response format. Here I describe some qualitative observations of these data. Sixty-six participants (11%) specifically commented on the speaker's use of the term *female*. Most expressed negative sentiments, ranging from the use of the word *female* being "odd" or "strange" to a general dislike of the term, to specifically labeling someone who refers to women as *females* as sexist and misogynist. For example, one participant stated, "I HATE anyone who refers to women as 'female.'" Another participant said, "It's pretty weird to refer to woman [sic] as 'females' all the time." Another participant said "They use female instead of woman. Seems like they don't like women." The following response encompasses a shared sentiment among some participants; "Weird he keeps referring to women as 'females'. That's always a red flag of someone more incline [sic] to sexist behaviors and actions." Similarly, another participant stated, "referring to women as 'females' is a big red flag that they view women as an instrument they can use."

Some participants directly mentioned inferring the speaker's demographics, as well as their attitudes about women, from the use of the term *females*. Many assumed the speaker was a man. For example, one participant stated, "I figure this was a male speaker because of how many

times ‘female’ was said by them.” Another stated, “Referring to the person as ‘female’ seems like a strong indicator of a male voice.” Other participants had more pointed observations about who uses the term female. One person wrote, “The usage of ‘female’ leads me to believe Speaker A is a man who doesn't regularly interact with women.” A few people called the speaker an “incel,” referring to an online subculture of men who self-identify as involuntarily celibate and have hostile attitudes about women (Joensuu, 2023; Krendel, 2020). For example, one participant wrote, “He sounds like a bit of an Incel. Lots of guys use the word ‘female’ instead of woman as a dog whistle - to degrade women and treating them like objects instead of actual people…” Another participant factored age into their speculation: “referring to women as 'females' is a huge red flag that the speaker is a terminally online incel. Age range is overwhelmingly 18-39 and it is vanishingly rare to find anyone older than that who consistently refers to women as 'females'.”

Lastly, some participants commented on the use of the term *female* more broadly. For example, one participant wrote, “The speaker seems very kind but the only thing I don't like is how they refer to women as females. The term female has a negative connotation nowadays and I think that a good amount of women would take offense to him referring to them as females.” In contrast, another participant acknowledged the potential for the word *female* to be seen as offensive, even when disagreeing with this: “I know there’s this weird social ritual where we get offended at non-offensive words. There is nothing disrespectful or wrong about the word female.” Another speaker mentioned that the use of the term depends on the context. They wrote, “Some people think use of the word ‘female’ is insensitive or politically incorrect. It depends on who the speaker is talking to.” Altogether, these responses support the quantitative findings that participants do form assumptions about the speaker’s demographic traits, as well as their attitudes and beliefs about women, based on the use of the term *female*.

Discussion

In Study 2, I examined three key questions. First, what characteristics and traits are associated with a person's repeated use of the terms *female* and *woman*? Specifically, if a person characteristically uses the term *female* (vs. the term *woman*) to refer to women, what does this signal to others about that person's traits, attitudes toward the people they are speaking about, and demographic characteristics? Secondly, do our evaluations of a speaker as sketched out above vary by context (dating or errands), by the order in which they see the term used (*female*-first or *woman*-first), and by the race of the participant (Black or White). Lastly, how do these ratings and perceptions relate to participants' own beliefs and attitudes about gender and sexism (as measured by SCEBS, ASI, and SIF)?

Did Participants Make Inferences about the Speaker Based on the Use of the Terms *Female* and *Woman*?

Consistent with my predictions, results showed that participants made inferences about a speaker based on word choice. Overall, participants imagined the speaker to be less likable, warm, competent, educated, formal, and polite when the speaker used the word *female* than when they used the word *woman*. Participants also felt less warmly about the speaker and perceived them to be less respectful of the women they referenced when they used the word *female* rather than *woman*. Additionally, the speaker was perceived to be less feminist, more sexist, and more conservative when using the word *female* than the word *woman*. The finding that the speaker was perceived as more "formal" when using the term *woman* than *female* is of note. The term *female* is seen as more academic and scientific and thus this is what I expect when thinking of "formal." It would be helpful to know how participants are defining formal in this context. This brings to

light an important methodological point of using operationalized terms or providing a space for participants to explain what they mean (McClelland & Holland, 2016).

Similarly, perceptions of demographic characteristics of the speaker varied depending on the speaker's word choice. Specifically, participants imagined that the speaker was more likely to be a man when using the word *female* than when using the word *woman*. A number of participants explicitly noted that they thought the speaker was a man because of the repeated use of the word *female*–implying that 'only a man would talk like that.' It is possible that participants believed men would be more likely to essentialize women (see Smiler & Gelman, 2008, for evidence that this is the case) and thus use a more essentializing/biologizing term (Schudson et al., 2019; Troncoso et al., 2022). This is in line with online discourse in which some people speculate that the use of *female* as a term serves as a red flag indicating that a man may be sexist or misogynist (Bee, 2025; Joensuu, 2023; Moody 2021; Sommers, 2025).

Participants also made inferences about race based on word choice. Specifically, participants imagined the speaker to be White more often when using the word *woman* than when using the word *female*. A few participants expressed, in their short responses, that the use of the term *female* is more common in Black communities, an observation also stated by Newton–Small (2016). To my knowledge, the idea that Black Americans use the term *females* more than White Americans has not been tested. Future areas of research could examine potential relations between gendered terms and different dialects and gendered terms and code-switching. There are other associations with who uses the term more frequently, including a person's online presence and age or cohort. I return to this point in the General Discussion.

How Did Perceptions of the Speaker Differ by Context?

Aside from word choice, I found that context often made a difference, such that the speaker in the dating context received more positive and higher ratings than the speaker in the errands context. Specifically, participants found the speaker to be more likable, warm, competent, educated, formal, polite, and more liberal in the dating context than in the errands context. Participants in the dating context also felt more warmly toward the speaker and perceived them to be more respectful toward the women they referenced. Another way to look at this is that participants were more sensitive to the difference between the use of the terms *female* and *woman* in the errands context than in the dating context. Although in both contexts, the speaker using *female* was perceived more negatively than the speaker using *woman*, that difference was greater for those in the errands condition. One possible explanation for this is that the use of the term *female* might feel more jarring in that context, where gender is not expected to be as salient, in comparison to the dating context. Alternatively, it could also be that in the dating context, participants are primed to think about gender and perhaps are inclined to be more favorable toward someone speaking about their dating experiences than someone speaking about daily inconveniences (e.g., losing their wallet) or even a seemingly neutral interaction (e.g., asking for a book recommendation).

Context also influenced participants' perception of speaker demographics. Participants perceived the speaker to be a man and to be younger more often in the dating context than the errands context. Because the only gendered terms provided in this study were *female* and *woman*, it makes sense that, living in a heteronormative society that influences our perceptions of gender and sexual orientation and how they function together (Barker, 2014; Marchia & Sommer, 2019), participants assumed that the speaker was a man. This is further evidenced by the finding that participants showed more heteronormativity in the dating context because they assumed

heterosexual relationships. In terms of age, participants might assume that people in dating context are younger as people tend to form long term relationships as they age.

How Did Perceptions of the Speaker Differ by Participant Race?

There were some overall differences among Black and White Participants. Specifically, Black participants felt more warmly about the speaker and found the speaker to be more likable, warm, competent, educated, formal, polite, and to be more respectful of the women they referenced. One possible explanation for why Black participants rated the speakers more positively than White participants could be that Black participants perceived the use of the terms *woman* and *female* as neutral (or similar enough to each other) and therefore relied on other aspects of the quotes to make judgments. It would be interesting to examine this further; for example, are there differences in how gendered terms are defined and perceived depending on racial/ethnic identity? Lastly, although both groups overall assumed that the speaker was White, White participants were more likely than Black participants to think that the speaker was White. It is possible that without specific identity markers, such as use of African American Vernacular English (AAVE), participants defaulted to the majority group in the context of the United States.

I found that Black participants scored higher on essentialism, hostile sexism, and benevolent sexism. This is consistent with literature that Black people endorse benevolent sexism at higher rates than White people (Campbell et al., 2023; Davis et al., 2022; Hayes & Swim 2013). With specific regard to benevolent sexism, scholars posit that this is because of higher endorsement of traditional roles (Hayes & Swim; 2013), and threat perception (Davis et al., 2024). Future research could examine how race plays a role in endorsement of essentialism and hostile sexism.

How Did Participants' Attitudes and Beliefs Relate to their Ratings of the Speaker?

I examined how participants' own attitudes and beliefs relate to their evaluations of the speaker. Overall, participants who endorsed essentialism more highly and those who held more sexist attitudes evaluated the speaker using the term *female* more positively than the speaker using the term *woman*. This is consistent with the findings linking essentialism to the use of the term *females*. In line with this, participants who more strongly identified as feminist evaluated the speaker using the term *female* more negatively across most dimensions. Interestingly, there were fewer effects involving social constructionist beliefs. Specifically, participants who more strongly agreed with a social constructionist view of gender found the speaker to be less educated when using the term *female* than *woman*, but there were no other differences. This is notable compared to the other scales and worthy of further research.

Limitations

The current study is not without limitations. For one, participant responses to the question "Is there anything else you'd like to share about the speaker?" were very informative but response was optional and only a small number of participants completed the question. In future research, I might request that all participants elaborate on their judgments, so I can better understand why they rated the speakers the way that they did. For example, I could ask, "What informed your perceptions of the speaker?" or "What information did you use from the quotes to make your judgments?" It would also be interesting to add to the speaker rating dimensions to examine what other associations participants make according to word choice. For example, regarding the speaker's occupation, would participants assume that the speaker who used the word *female* worked in careers where we might expect to see its use more often (e.g., medical settings)? Lastly, I would consider adding more direct questions about the participant's experience with gendered terms. For example, "Do you ever hear the term *females* used

casually? If so, how?" "Do you find the use of the term *females* offensive, neutral, or positive? Please explain." There are additional limitations to address but as they apply to both Study 1 and Study 2, I will address those in the General Discussion.

Chapter IV: General Discussion

Across the two studies of this dissertation, I examined how speakers are evaluated based on their choice of gendered words to refer to others (Study 1) and to women specifically (Study 2). My main questions were: what impressions are formed of a speaker based on the gendered term they use? What demographic characteristics are assumed based on the term? And how do participants' attitudes and beliefs shape their evaluations of the speaker? I found that people indeed make a broad array of assumptions about a speaker's demographic and character traits based on the term used. Largely, calling women *female* resulted in the speaker being evaluated more negatively (less likable, less warm, less polite). Use of terms also had implications for perceived beliefs of the speaker; participants perceived the speaker to be more sexist, less feminist, and more conservative when using the term *female* to refer to women. Lastly, participants' own attitudes and beliefs about gender predicted their evaluations of the speaker. The more a participant endorsed sexist and essentialist attitudes, the more positively they rated the speaker who used the word *female*. In contrast, the more a participant identified as a feminist, the more negatively they rated the speaker who used the word *female*.

These results reveal that subtle shifts in language can function as a cue to others about a speaker's attitudes and beliefs about gender. Specifically, calling women *female* repeatedly can serve as a signal to others that the speaker holds more sexist, more essentialist, and less feminist attitudes toward women than using the term *woman* to refer to women. This supports observations made online by many who found that most often when they saw the term *female*, it

was in a negative statement about women (Moody, 2021; Sommers, 2025; Sullivan, 2021; Webling, 2021). This sentiment is shared by Sommers (2025): "I've never seen 'females' used by a man in a way that wasn't inherently misogynistic toward women, and again, I'm certainly not the only. 'Females are the best,' said literally by no man ever."

These findings may have practical implications for people in varied professional settings and in interpersonal contexts. For example, I found that the use of the term *female* made the speaker seem more conservative. This could have implications for political candidates, informing which terms they choose to use and to avoid. Conversely, as was the case in Tarillo (2023), using gender-fair language can make people seem more liberal. This research could have implications for the classroom. Students may be attuned to the gender terms their teachers and professors use and make inferences about their attitudes. In turn this may have consequences for how the student views the professor and how they choose to participate in the class (Haleta, 1996; Kleinman et al., 2021). In general, this research may have implications for interpersonal relationships. People may not know that the term they are using is communicating certain information. For example, some who use the term *female* in casual settings may do so because they are accustomed to using it in their job ("military speak" or "police speak"; Joensuu, 2023). They may not realize that they are using a term with negative connotations and that people are indeed making assumptions about their attitudes and beliefs (Kleinman et al., 2021; Wang & Ngai, 2023).

Comparison of the Two Studies

There were a number of consistencies across the two studies, reinforcing the findings of each. Most notably, the two studies showed the same pattern of more negative evaluations of the speaker using *female* than *woman*, despite the differences in method (open-ended descriptions in

Study 1; close-ended ratings in Study 2). Moreover, in both studies, participants' own attitudes and beliefs were related to their ratings of the hypothetical speakers. In Study 1, participants who endorsed hostile sexism more strongly, showed fewer negative judgments (i.e., fewer negative descriptions of the speaker) about the use of the terms *female* and *male*. Similarly, in Study 2 participants who endorsed essentialism more strongly found the speaker using the term *female* (as compared to *woman*) more likable. Overall, though, there were stronger relations with individual differences in Study 2 than in Study 1, likely because there was a stronger signal in Study 2 (more consistent usage of a given term by the speaker) and more targeted dimensions (i.e., explicitly asking if the speaker is likable). Despite these differences in method, in both studies a number of speakers expressed strong judgments about the speaker using the term *female*. For example, in Study 1, one participant described the speaker as "Probably a guy. Either a tad sexist or just plain a weird person like a robot or something, but who refers to someone as a 'female' like that? Also apparantly [sic] somewhat spineless and not one to insist on having his way." Another participant wrote, "This is definitely a man, the reference to a 'female' is a dead giveaway. He is probably conservative and solidly misogynist." See Open-Ended Responses in Study 2 for similar quotes.

Future Research

An important direction for future research is to examine the role of race in the perception of gendered terms. In the current studies, I only examined this question obliquely. Although I included both Black and White participants in equal numbers, I did not provide participants with information about the race of the characters involved in the speech events (the speaker, the person they were describing, or the person they were speaking to). I believe that this information may be critical to gaining insight into how race may play a role in people's interpretations of

gendered language (see also Kirkman & Oswald, 2020). Indeed, some participants spontaneously mentioned that race was a factor in their thinking. For example, in response to the question of whether they wanted to share anything else about the speaker, a White participant in Study 2 stated, “If the speaker was white, I would feel they are disrespectful, if the speaker was Black I would feel they were codeswitching and not as offensive.” Another participant wrote, “Since they kept referring to the woman as female, I assume they were a male. ‘Female’ is most often used by AA males (me being one). It's not formal, just more of a cultural nuance.” Interestingly, in these examples, "female" is described as more common in how a Black person might speak, and thus less revealing of their attitudes.

In the current studies, most participants assumed that the speaker was White, and it is unknown what they assumed about the race of the person being referenced. In future research, it would be important to provide the race and gender of both the speaker and the person referenced in order to examine how these interact. For example, how would the ratings of a White man calling a Black woman a *female* differ from a White man calling a White woman a *female*? Would there be any differences between the perception of Black women calling other women a *female* compared to Black men?

Of course, race and gender are not the only dimensions along which people's language is assessed. I suspect that a variety of other factors may also play a role, including the imagined speaker's age, geographic location, and dialect. For example, because language changes over time and words take on different meanings, or are sometimes reclaimed, listeners may incorporate information about a person’s age in evaluating the use of a term, concluding, for example, ‘that’s the way people of that generation speak.’ Similarly, evaluations may differ if the

speaker made use of African American Vernacular English (AAVE), regional markers (e.g., use of the word “y'all”), or generational markers (phrases commonly used by Generation Z).

Another interesting avenue for research would be to systematically examine how response to the use of the term *female* compares to other problematic but not overtly hostile terms (e.g., terms of endearment such as *honey* and *sweetie*) and to more overtly offensive terms (e.g., “the *B* word”). Because the term *female* is one of the most common gender terms (along with *woman*) it would be interesting to see if, for example, people who have negative responses to terms like *sweetie* would also respond negatively to the term *female*. Alternatively, within the context of more overtly negative terms, the word *females* may be perceived as more neutral.

Lastly, social media would be a very interesting context to examine with regard to these questions. There are relevant studies that explore the use of gender terms in online spaces/corpora (Joensuu, 2023; Kendrel, 2020; Pearce, 2008; Vásquez, 2022). Future research could build on this by more closely analyzing the online conversations about people using the term *female(s)* to refer to women, focusing not only on the use of the terms, but also the content of what is expressed, and the metalinguistic discussion of the term.

Conclusion

How much of an impact does language about gender have? Feminist scholars have focused, and continue to focus, on issues related to gendered language because it is a powerful tool, even if addressing issues in language may not lead to change on its own. As Lakoff (1973) stated, for language about gender to change there has to be a shared understanding and acknowledgment that there is a problem (i.e., structural sexism). Additionally, people intentionally refuse to use certain terms in order to deny others’ identity and to reinforce their beliefs and perpetuate harm. For example, the refusal to use ‘they’ pronouns is one way that

some try to deny the existence of gender outside of a binary view and cause interpersonal harm (Bailey et al., 2024). Another example of this link between language and structural sexism can be seen in the Trump administration's focus on changing the language about gender. Trump's second executive order, once back in office, declared gender identity, as a concept, in language, and as part of people's lived experiences, irrelevant.

Language has the power to shape, reflect, and reinforce people's attitudes and beliefs about gender. The current studies suggest that people make inferences about a speaker based on the gender term they use and often have more negative perceptions of the speaker using the term *female* relative to the term *woman*. These findings have broad implications, including for people in leadership roles (e.g., political candidates and policy makers), schools and workplaces, and everyday interpersonal interactions. Calling a woman a *female,* or a group of women *females,* may feel dehumanizing to some listeners and reinforce negative gender roles (e.g., the idea that women's main function is to reproduce, given the term's ties to biology). Changing how people speak about women, on its own, may not change people's attitudes and beliefs, but it is an important step toward gender-fair language and gender equality more broadly.

References

Ames, D. R., Bianchi, E. C., & Magee, J. C. (2010). Professed impressions: What people say about others affects onlookers' perceptions of speakers' power and warmth. *Journal of Experimental Social Psychology*, *46*(1), 152–158. https://doi.org/10.1016/j.jesp.2009.09.011

Bailey, A. H., Dembroff, R., Wodak, D., Ikizer, E. G., & Cimpian, A. (2024). People's beliefs about pronouns reflect both the language they speak and their ideologies. *Journal of Experimental Psychology: General*, *153*(5), 1388–1406. https://doi.org/10.1037/xge0001565

Barbee, H., Deal, C., & Gonzales, G. (2022). Anti-transgender legislation—A public health concern for transgender youth. *JAMA Pediatrics*, *176*(2), 125–126. https://doi.org/10.1001/jamapediatrics.2021.4483

Barker, M. (2014). Heteronormativity. In *Encyclopedia of Critical Psychology* (pp. 858–860). Springer, New York, NY. https://doi.org/10.1007/978-1-4614-5583-7_134

Bedin, C. A. (2025). *"A Tale of Two Pronouns": Race, Place, Gender, and the Indexical Tangle of You Guys and Y'all* [M.A., University of California, Santa Barbara]. https://www.proquest.com/docview/3198955787/abstract/3BEC64E693DB4BC8PQ/1

Bee, K. (2025, April 16). Is it offensive when men call women "females"? *HubPages*. https://discover.hubpages.com/politics/calling-women-females

Bennett-Alexander, D. D. (2010). *The use of the term "boy" as evidence of race discrimination: Apparently the 11th circuit didn't get the memo?* (SSRN Scholarly Paper 1785095). Social Science Research Network. https://doi.org/10.2139/ssrn.1785095

Berry, D. S., Pennebaker, J. W., Mueller, J. S., & Hiller, W. S. (1997). Linguistic bases of social perception. *Personality and Social Psychology Bulletin*, *23*(5), 526–537. https://doi.org/10.1177/0146167297235008

Boasso, A., Covert, S., & Ruscher, J. B. (2012). Benevolent sexist beliefs predict perceptions of speakers and recipients of a term of endearment. *The Journal of Social Psychology*, *152*(5), 533–546. https://doi.org/10.1080/00224545.2011.650236

Bradley, E. D. (2020). The influence of linguistic and social attitudes on grammaticality judgments of singular 'they.' *Language Sciences*, *78*, 101272. https://doi.org/10.1016/j.langsci.2020.101272

Brassel, S. T., Settles, I. H., Jellison, W. A., & Dodson, C. (2020). Power and race in Black and White men's perceptions and experiences of manhood. *Translational Issues in Psychological Science*, *6*(4), 325–343. https://doi.org/10.1037/tps0000257

Brown, K. (2015, February 5). The problem with calling women "females." *Jezebel Magazine*. https://jezebel.com/the-problem-with-calling-women-females-1683808274

Bush, B. (1981). White 'ladies', coloured 'favourites' and black 'wenches'; some considerations on sex, race and class factors in social relations in white Creole Society in the British Caribbean. *Slavery & Abolition*, *2*(3), 245–262. https://doi.org/10.1080/01440398108574830

Cralley, E. L., & Ruscher, J. B. (2005). Lady, girl, female, or woman: Sexism and cognitive busyness predict use of gender-biased nouns. *Journal of Language and Social Psychology*, *24*(3), 300–314. https://doi.org/10.1177/0261927X05278391

David, E. (2017). Capital T: Trans visibility, corporate capitalism, and commodity culture. *TSQ: Transgender Studies Quarterly*, *4*(1), 28–44. https://doi.org/10.1215/23289252-3711517

Davis, T. M., Settles, I. H., & Jones, M. K. (2022). Standpoints and situatedness: Examining the perception of benevolent sexism in Black and white undergraduate women and men. *Psychology of Women Quarterly*, 03616843211043108. https://doi.org/10.1177/03616843211043108

DeFrank, M., & Kahlbaugh, P. (2019). Language choice matters: When profanity affects how people are judged. *Journal of Language and Social Psychology*, *38*(1), 126–141. https://doi.org/10.1177/0261927X18758143

Diaz-Leon, E. (2016). *Woman* as a Politically significant term: A solution to the puzzle. *Hypatia*, *31*(2), 245–258. https://doi.org/10.1111/hypa.12234

D'Oyley, D. L. (2013, February 7). *Who you calling a female?* The Root. https://www.theroot.com/who-you-calling-a-female-1790895152

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fausey, C. M., & Boroditsky, L. (2010). Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin & Review*, *17*(5), 644–650. https://doi.org/10.3758/PBR.17.5.644

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P., & Whitehead, J. (2010). Hostility toward men and the perceived stability of male dominance. *Social Psychology*, *41*(3), 177–185. https://doi.org/10.1027/1864-9335/a000025

Habarth, J. M. (2015). Development of the heteronormative attitudes and beliefs scale. *Psychology & Sexuality*, *6*(2), 166–188. https://doi.org/10.1080/19419899.2013.876444

Haleta, L. L. (1996). Student perceptions of teachers' use of language: The effects of powerful and powerless language on impression formation and uncertainty. *Communication Education*, *45*(1), 16–28. https://doi.org/10.1080/03634529609379029

Henderson, A. (2003). What's in a Slur? *American Speech*, *78*(1), 52–74. https://muse.jhu.edu/pub/4/article/40728

Hildebrand, L. K., Monteith, M. J., Carter, E. R., & Burns, M. D. (2022). Honey, sweetie, dear: Terms of endearment communicate, reflect, and reinforce sexism toward adult women. *Sex Roles*. https://doi.org/10.1007/s11199-022-01311-3

House, T. W. (2025, January 21). *Defending women from gender ideology extremism and restoring biological truth to the federal government*. The White House. https://www.whitehouse.gov/presidential-actions/2025/01/defending-women-from-gender-ideology-extremism-and-restoring-biological-truth-to-the-federal-government/

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The report of the 2015 U.S.transgender survey.* National Center for Transgender Equality. https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf

Jerald, M. C., Ward, L. M., Moss, L., Thomas, K., & Fletcher, K. D. (2017). Subordinates, Sex Objects, or Sapphires? Investigating Contributions of Media Use to Black Students' Femininity Ideologies and Stereotypes About Black Women. *Journal of Black Psychology*, *43*(6), 608–635. https://doi.org/10.1177/0095798416665967

Joensuu, S. (2023). *The plural females and its derogatory use: A corpus study on the sexist aspects in the use of the word* [Tampere University]. https://trepo.tuni.fi/bitstream/handle/10024/150851/JoensuuSanteri.pdf

Jones, M. K., Buque, M., & Miville, M. L. (2018). African American Gender Roles: A Content Analysis of Empirical Research From 1981 to 2017. *Journal of Black Psychology*, *44*(5), 450–486. https://doi.org/10.1177/0095798418783561

King, S. (2020). From African American Vernacular English to African American Language: Rethinking the study of race and language in African Americans' speech. *Annual Review of Linguistics*, *6*(1), 285–300. https://doi.org/10.1146/annurev-linguistics-011619-030556

Kirkman, M. S., & Oswald, D. L. (2020). Is it just me, or was that sexist? The role of sexism type and perpetrator race in identifying sexism. *The Journal of Social Psychology*, *160*(2), 236–247. https://doi.org/10.1080/00224545.2019.1634505

Kleinman, S., Copp, M., & Wilson, K. B. (2021). We've come a long way, guys! Rhetorics of resistance to the feminist critique of sexist language. *Gender & Society*, *35*(1), 61–84. https://doi.org/10.1177/0891243220979636

Krendel, A. (2020). The men and women, guys and girls of the 'manosphere': A corpus-assisted discourse approach. *Discourse & Society*, *31*(6), 607–630. https://doi.org/10.1177/0957926520939690

Labotka, D. (2020). *Children's evaluations of interlocutors in foreigner talk contexts* [Thesis]. http://deepblue.lib.umich.edu/handle/2027.42/155151

Labotka, D., & Gelman, S. A. (2024). The effect of foreigner talk on children's evaluations of addressees. *Cognitive Development*, *71*, 101486. https://doi.org/10.1016/j.cogdev.2024.101486

Lakoff, R. (1973). Language and woman's place. *Language in Society*, *2*(1), 45–80. http://www.jstor.org/stable/4166707

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Leath, S., Jones, M., Jerald, M. C., & Perkins, T. R. (2021). An investigation of Jezebel stereotype awareness, gendered racial identity and sexual beliefs and behaviours among Black adult women. *Culture, Health & Sexuality*, *0*(0), 1–18. https://doi.org/10.1080/13691058.2020.1863471

Leinonen, E. (2020). *Language, heteronormativity and gender in EFL teaching* [Stockholm University]. urn:nbn:se:su:diva-182669

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The demographic and political composition of Mechanical Turk samples. *Sage Open*, *6*(1), 2158244016636433. https://doi.org/10.1177/2158244016636433

Lipton, J. P., & Hershaft, A. M. (1984). "Girl," "woman," "guy," "man": The effects of sexist labeling. *Sex Roles*, *10*(3), 183–194. https://doi.org/10.1007/BF00287773

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Lloyd, A. E., & Galupo, M. P. (2019). What people with normative identities believe about sex, gender and sexual orientation. *Psychology & Sexuality*, *10*(3), 269–280. https://doi.org/10.1080/19419899.2019.1614088

Macarthur, H. (2015). *When women are called "girls": The effect of infantilizing labels on women's self-perceptions* [Penn State]. https://etda.libraries.psu.edu/catalog/25797

Marchia, J., & Sommer, J. M. (2019). (Re)defining heteronormativity. *Sexualities*, *22*(3), 267–295. https://doi.org/10.1177/1363460717741801

McClelland, S. I., & Holland, K. J. (2016). Toward better measurement: The role of survey marginalia in critical sexuality research. *Qualitative Psychology*, *3*(2), 166–185. https://doi.org/10.1037/qup0000056

Media, W. P. (2017, May 16). Black women on being called "girl" In the workplace. *WLRN*. https://www.wlrn.org/news/2017-05-15/black-women-on-being-called-girl-in-the-workplace

Meier, T., Boyd, R. L., Mehl, M. R., Milek, A., Pennebaker, J. W., Martin, M., Wolf, M., & Horn, A. B. (2020). Stereotyping in the digital age: Male language is "ingenious", female language is "beautiful" – and popular. *PLOS ONE*, *15*(12), e0243637. https://doi.org/10.1371/journal.pone.0243637

Moody, K. (2021, July 24). Stop calling women "females." *The Daily Nexus | The University of California, Santa Barbara's Independent, Student-Run Newspaper.* https://dailynexus.com/2021-07-24/stop-calling-women-females/

Mulac, A., Incontro, C. R., & James, M. R. (1985). Comparison of the gender-linked language effect and sex role stereotypes. *Journal of Personality and Social Psychology*, *49*(4), 1098–1109. https://doi.org/10.1037/0022-3514.49.4.1098

Newton-Small, J. (2016, April 20). Why we need to reclaim the word "female." *Time*. https://time.com/4300170/female-word/

Nicholson, K. J., Sherman, M., Divi, S. N., Bowles, D. R., & Vaccaro, A. R. (2022). The Role of Family-wise Error Rate in Determining Statistical Significance. *Clinical Spine Surgery*, *35*(5), 222. https://doi.org/10.1097/BSD.0000000000001287

Norris, M. (2019, May 30). Female trouble: The debate over "woman" as an adjective. *The New Yorker*. https://www.newyorker.com/culture/comma-queen/female-trouble-the-debate-over-woman-as-an-adjective

Parks, J. B., & Roberton, M. A. (2004). Attitudes toward women mediate the gender effect on attitudes toward sexist language. *Psychology of Women Quarterly*, *28*(3), 233–239. https://doi.org/10.1111/j.1471-6402.2004.00140.x

Patev, A. J., Dunn, C. E., Hood, K. B., & Barber, J. M. (2019). College students' perceptions of gender-inclusive language use predict attitudes toward transgender and gender nonconforming individuals. *Journal of Language and Social Psychology*, *38*(3), 329–352. https://doi.org/10.1177/0261927X18815930

Pearce, M. (2008). Investigating the collocational behaviour of MAN and WOMAN in the BNC using Sketch Engine. *Corpora*, *3*(1), 1–29. https://doi.org/10.3366/E174950320800004X

Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *WIREs Cognitive Science*, *8*(4), e1437. https://doi.org/10.1002/wcs.1437

Saguy, A. C., & Williams, J. A. (2022). A little word that means a lot: A reassessment of singular they in a new era of gender politics. *Gender & Society*, *36*(1), 5–31. https://doi.org/10.1177/08912432211057921

Schudson, Z. C., Beischel, W. J., & van Anders, S. M. (2019). Individual variation in gender/sex category definitions. *Psychology of Sexual Orientation and Gender Diversity*, *6*(4), 448–460. https://doi.org/10.1037/sgd0000346

Schudson, Z. C., & Gelman, S. A. (2022). Social constructionist and essentialist beliefs about gender and race. *Group Processes & Intergroup Relations*, 13684302211070792. https://doi.org/10.1177/13684302211070792

Sczesny, S., Formanowicz, M., & Moser, F. (2016). Can gender-fair language reduce gender stereotyping and discrimination? *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00025

Settles, I. H., Pratt-Hyatt, J. S., & Buchanan, N. T. (2008). Through the lens of race: Black and White Women's perceptions of womanhood. *Psychology of Women Quarterly*, *32*(4), 454–468. https://doi.org/10.1111/j.1471-6402.2008.00458.x

Smiler, A. P., & Gelman, S. A. (2008). Determinants of gender essentialism in college students. *Sex Roles*, *58*(11), 864–874. https://doi.org/10.1007/s11199-008-9402-x

Sommer, J. (2025, May 22). Please don't refer to women as "females." *InsideHook*. https://www.insidehook.com/internet/please-dont-refer-women-females

Sullivan, K. (2021, March 9). Casual use of "female" is more harmful than it seems. *Berkeleyhighjacket.Com*. https://berkeleyhighjacket.com/2021/opinion/casual-use-of-female-is-more-harmful-than-it-seems

Swim, J. K., Mallett, R., & Stangor, C. (2004). Understanding subtle sexism: Detection and use of sexist language. *Sex Roles*, *51*(3), 117–128. https://doi.org/10.1023/B:SERS.0000037757.73192.06

Szymanski, D. M. (2004). Relations among dimensions of feminism and internalized heterosexism in lesbians and bisexual Women. *Sex Roles*, *51*(3), 145–159. https://doi.org/10.1023/B:SERS.0000037759.33014.55

Tarillo, M. C. (2023). *Is He/She/They conservative? Political implications of the use of gender-fair language* [M.A., The University of North Carolina at Chapel Hill]. https://www.proquest.com/docview/2854293500/abstract/FAD37E801C74F66PQ/1

Troncoso, S. C., Schudson, Z. C., & Gelman, S. A. (2022). Women versus females: Gender essentialism in everyday language. *Journal of Psycholinguistic Research*. https://doi.org/10.1007/s10936-022-09917-0

van Anders, S. M., Schudson, Z. C., Abed, E. C., Beischel, W. J., Dibble, E. R., Gunther, O. D., Kutchko, V. J., & Silver, E. R. (2017). Biological sex, gender, and public policy. *Policy Insights from the Behavioral and Brain Sciences*, *4*(2), 194–201. https://doi.org/10.1177/2372732217720700

Vásquez, J., Bel-Enguix, G., Andersen, S. T., & Ojeda-Trueba, S.-L. (2022). HeteroCorpus: A corpus for heteronormative language detection. In C. Hardmeier, C. Basta, M. R. Costa-jussà, G. Stanovsky, & H. Gonen (Eds.), *Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP)* (pp. 225–234). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.gebnlp-1.23

Vergis, N., & Pell, M. D. (2020). Factors in the perception of speaker politeness: The effect of linguistic structure, imposition and prosody. *Journal of Politeness Research*, *16*(1), 45–84. https://doi.org/10.1515/pr-2017-0008

Wang, W., & Ngai, J. (2023). “You look like my 14-year-old daughter”: A corpus-based study of sexist language in everyday sexism Twitter stories. *Journal of Language Aggression and Conflict*. https://doi.org/10.1075/jlac.00090.wan

Webling, C. (2021, February 18). Please stop calling us “females” in your gaming reports. *Medium*. https://superjumpmagazine.com/please-stop-calling-us-females-in-your-gaming-reports-1785cb66949e

Appendices

Appendix A: Study 1 Survey Tasks

Task Instructions
You will read a series of quotes.

- Each screen will include one quote.
- **Each quote is from a different person.**
- We ask you to describe the **speaker who said the quote** based on your first impressions, your best guess as to who the speaker is most likely to be. For example, you can include age, gender, race, personality – whatever you think is relevant.
- We ask that you include at least 2-3 descriptors for each quote. You do not need to write complete sentences.

Describe the speaker who said the quote based on your first impressions, your best guess as to who the speaker is most likely to be. For example, you can include age, gender, race, personality – whatever you think is relevant. Include at least 2-3 descriptors for each quote. You do not need to write complete sentences.

Survey Items

Errands

1. "Last week I went to the bank to deposit my savings. The [gendered term/person] who helped me asked for my bank account number. Since I don't have my bank account number memorized, I pulled out my phone to log in to my bank app."

2. "I went to the library yesterday, looking for something new to read. I wasn't looking for anything in particular and after a while I figured I would ask the [gendered term/person] behind the counter for recommendations."

3. "I went to a concert. The [gendered term/person] at the concession stand did not have change for a $20 bill so I had to wait for someone to pay with smaller bills in order to get change back."

4. "I went to my local pharmacy to pick up a prescription. I was surprised that there was a new [gendered term/person] behind the counter."

5. "I lost my wallet last week. I called the restaurant to see if anyone returned it but the [gendered term/person] who answered the phone said no one found a wallet."

6. “I went to the store to buy drinks before a gathering. I tripped and stumbled into a [gendered term/person] buying beer. I apologized and continued shopping.”

7. “As I drove to work this morning, a [gendered term/person] signaled to let me know that my car had a broken taillight. I called the mechanic on the way to my office.”

8. “I went to the gym in my office building earlier today and noticed it was crowded. I needed a machine to do my workout but there was a [gendered term/person] using it, so I moved on to the next station.”

9. “I went to the furniture store to buy a couch and the [gendered term/person] who helped me let me know that the one I wanted was on backorder. I ended up going with my second choice.”

Dating

1. “A friend asked me what I am looking for in someone. I responded that I am looking for a [gendered term/person] who shares my values and interests.”

2. “My friend arranged a blind date for me to go on next week. It has been a while since I have met an interesting [gendered term/person] so I am excited about this.”

3. “I find it so funny when my friends try to set me up on dates. They insist that they found the perfect [gendered term/person] for me and get excited to hear about the date afterwards.”

4. “I don’t enjoy being set up on dates. I'd rather meet a [gendered term/person] on my own, because it’s easier to get to know someone without external pressure.”

5. “My single friends are obsessed with dating. If they thought less about dating maybe the right [gendered term/person] would just come along.”

6. “I was invited to a singles mixer last week. I wasn't planning to go but a friend encouraged me, saying it’s a great place to meet an outgoing [gendered term/person].”

7. “I was set up on a blind date with a [gendered term/person] my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation.”

8. “I want to download a dating app in hopes of meeting a compatible [gendered term/person] but I am overwhelmed by the different app options. I am going to do more research and see which app to start with.”

9. “I listen to many dating podcasts and a common piece of advice is that you have to put yourself out there to meet a [gendered term/person] that is a good match. I am just not sure how to do it.”

Appendix B: Study 1 SCEBS, ASI, and Demographic Questions

Now we want to learn more about your beliefs and attitudes. Please complete the following questions.

Social Constructionist and Essentialist Beliefs Scale, Schudson & Gelman 2022
(1) Strongly disagree to (7) Strongly agree

Instructions: Please indicate your agreement with the following statements about gender.

Social constructionism

1. If society were different, the categories "women" and "men" could be very different than they are.
2. How we talk about gender as a society affects what gender actually is.
3. Differences between women and men depend on how people talk about gender.
4. The way to be a woman or man changes across generations.
5. Different societies think about gender differently.
6. There are many different acceptable ways to define what it means to be a woman or a man.
7. If people work together, they can affect what it means to be a woman or a man in our society.

Essentialism

8. There are innate differences in attitudes and behaviors associated with different genders.
9. Differences between genders are due to different core essences.
10. It is possible to predict many aspects of a person once you know their gender.
11. Members of the same gender share many similarities with one another.
12. A person's core essence can't change; it is always affected by their gender regardless of what they do.
13. Differences between genders start early in life and do not change across development.

Realism

14. Gender is a real category.
15. "Women" and "men" are made-up categories. (Reverse-coded)
16. Gender is a natural category.

Ambivalent Sexism Inventory (ASI) [Short version], Glick & Whitehead (2010)
(0) Strongly disagree to (5) Strongly agree

Instructions: Below is a series of statements concerning men and women and their relationships in contemporary society. Please indicate the degree to which you agree or disagree with each statement.

1. Many women have a quality of purity that few men possess.
2. Women should be cherished and protected by men.
3. Women seek to gain power by getting control over men.
4. Every man ought to have a woman whom he adores.
5. Men are incomplete without women.
6. Women exaggerate problems they have at work.
7. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash.
8. When women lose to men in a fair competition, they typically complain about being discriminated against.
9. Many women get a kick out of teasing men by seeming sexually available and then refusing male advances.
10. Women, compared to men, tend to have a superior moral sensibility.
11. Men should be willing to sacrifice their own well being in order to provide financially for the women in their lives.
12. Feminists are making unreasonable demands of men.

Study 1: Participant Demographics

Now we ask that you complete the following demographics questions.

What is your gender identity?

Woman
Man
Non-binary
Another gender

I prefer that others refer to me as a: _____(women)

Woman
Female
Girl
Gal
Lady

I prefer that others refer to me as a: _____(men)

Man
Male
Boy
Guy
Gentleman

What is your sexual orientation? _____

What is your racial/ethnic background?

White or European American
Black or African American
Native American or Alaska Native
Asian or Asian American
Native Hawaiian or Pacific Islander
Hispanic or Latinx
Multi-racial (please specify) _____
Another racial/ethnicity (please specify) _____

What is the highest level of education you have completed?

Less than high school
Some high school
High school graduate
Some college
Finished training other than college (e.g., vocational school)
Graduated from 2-year college
Graduated from 4-year college or university
Some graduate or professional school
Received masters, professional, or doctoral degree

What is your family's combined yearly income?

Less than $15,000
$15,000 - $25,000
$25,000 - $35,000
$35,000 - $50,000
$50,000 - $75,000
$75,000 - $100,000
$100,000 - $150,000
Above $150,000

How old are you? _____

Do you identify as a feminist?

Yes
No
Maybe/unsure

Political Leaning

(1) Very liberal to (7) Very conservative

In terms of economic issues, how would you describe your political attitudes and beliefs?
In terms of social issues, how would you describe your political attitudes and beliefs?

Is English your first language?

Yes
No

If no: **At what age did you learn to speak English? _____**

Do you speak a language other than English?

Yes
No

If yes: **What other language do you speak? _____**

What do you think this survey is about? ____

Is there anything you would like to share with the researchers (comments, questions, concerns)? ____

Appendix C: Study 2 Speaker Snapshot Task

Instructions: In this task you will be asked to provide your impressions of two different people (one at a time), after reading statements each person has made.

First we will ask you to read a set of four (4) statements made by Speaker A and provide your impressions of Speaker A.

Next we will ask you to read a set of four (4) statements made by Speaker B and provide your impressions of Speaker B.

Be sure to answer the questions based on your first impressions of each speaker. There are no right or wrong answers; we are only interested in your impressions of each speaker.

Attention Check
Before we proceed to the task, please answer the following questions, to show that you understand the instructions.

> How many speakers will you be asked to judge? [options: 1, 2, 3; correct answer: 2]
>
> You will be provided with _____ statements that each speaker made. [options: 2, 3, 4; correct answer: 4]
>
> You should provide your first impressions of each speaker. [options: True/False; correct answer: True]

You must answer all three questions correctly before moving forward. [The survey is designed to not allow participants to move forward in the study until they answer all three questions correctly.]

<u>Introduction of Speaker A</u>

[Each participant read the following text:]

Speaker A

The four statements below were made by Speaker A on different occasions. Please read them carefully.

[And then read one of the following sets of statements:]

[V1_A_Errand_Female]

"Last week I went to the bank to deposit my savings. The female who helped me asked for my bank account number. Since I don't have my bank account number memorized, I pulled out my phone to log in to my bank app."

"I went to the library yesterday, looking for something new to read. I wasn't looking for anything in particular and after a while I figured I would ask the female behind the counter for recommendations."

"I went to a concert. The female at the concession stand did not have change for a $20 bill so I had to wait for someone to pay with smaller bills in order to get change back."

"As I drove to work this morning, a female signaled to let me know that my car had a broken taillight. I called the mechanic on the way to my office."

[V2_A_Errand_Woman]

"I lost my wallet last week. I called the restaurant to see if anyone returned it but the woman who answered the phone said no one found a wallet."

"I went to the gym in my office building earlier today and noticed it was crowded. I needed a machine to do my workout but there was a woman using it so I moved on to the next station."

"I went to the furniture store to buy a couch and the woman who helped me let me know that the one I wanted was on backorder. I ended up going with my second choice."

"I went to my local pharmacy to pick up a prescription. I was surprised that there was a new woman behind the counter."

[V3_A_Errand_Female]

"I lost my wallet last week. I called the restaurant to see if anyone returned it but the female who answered the phone said no one found a wallet."

"I went to the gym in my office building earlier today and noticed it was crowded. I needed a machine to do my workout but there was a female using it so I moved on to the next station."

"I went to the furniture store to buy a couch and the female who helped me let me know that the one I wanted was on backorder. I ended up going with my second choice."

"I went to my local pharmacy to pick up a prescription. I was surprised that there was a new female behind the counter."

[V4_A_Errand_Woman]

"Last week I went to the bank to deposit my savings. The woman who helped me asked for my bank account number. Since I don't have my bank account number memorized, I pulled out my phone to log in to my bank app."

"I went to the library yesterday, looking for something new to read. I wasn't looking for anything in particular and after a while I figured I would ask the woman behind the counter for recommendations."

"I went to a concert. The woman at the concession stand did not have change for a $20 bill so I had to wait for someone to pay with smaller bills in order to get change back."

"As I drove to work this morning, a woman signaled to let me know that my car had a broken taillight. I called the mechanic on the way to my office."

[V5_A_Dating_Female]

"A friend asked me what I am looking for in someone. I responded that I am looking for a female who shares my values and interests."

"My friend arranged a blind date for me to go on next week. It has been a while since I have met an interesting female so I am excited about this."

"I find it so funny when my friends try to set me up on dates. They insist that they found the perfect female for me and get excited to hear about the date afterwards."

"I want to download a dating app in hopes of meeting a compatible female but I am overwhelmed by the different app options. I am going to do more research and see which app to start with."

[V6_A_Dating_Woman]

"I listen to many dating podcasts and a common piece of advice is that you have to put yourself out there to meet a woman that is a good match. I am just not sure how to do it."

"My single friends are obsessed with dating. If they thought less about dating, maybe the right woman would just come along."

"I was set up on a blind date with a woman my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation."

"I was invited to a singles mixer last week. I wasn't planning to go but a friend encouraged me, saying it's a great place to meet an outgoing woman."

[V7_A_Dating_Female]

"I listen to many dating podcasts and a common piece of advice is that you have to put yourself out there to meet a female that is a good match. I am just not sure how to do it."

"My single friends are obsessed with dating. If they thought less about dating maybe the right female would just come along."

"I was set up on a blind date with a female my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation."

"I was invited to a singles mixer last week. I wasn't planning to go but a friend encouraged me, saying it's a great place to meet an outgoing female."

[V8_A_Dating_Woman]

"A friend asked me what I am looking for in someone. I responded that I am looking for a woman who shares my values and interests."

"My friend arranged a blind date for me to go on next week. It has been a while since I have met an interesting woman so I am excited about this."

"I find it so funny when my friends try to set me up on dates. They insist that they found the perfect woman for me and get excited to hear about the date afterwards."

"I want to download a dating app in hopes of meeting a compatible woman but I am overwhelmed by the different app options. I am going to do more research and see which app to start with."

[Participants then answered questions about the speaker and were presented with speaker B]

<u>Introduction of Speaker B</u>

Speaker B

The four statements below were made by Speaker B on different occasions. Please read them carefully.

[V1_B_Errand_Woman]

"I lost my wallet last week. I called the restaurant to see if anyone returned it but the woman who answered the phone said no one found a wallet."

"I went to the gym in my office building earlier today and noticed it was crowded. I needed a machine to do my workout but there was a woman using it so I moved on to the next station."

"I went to the furniture store to buy a couch and the woman who helped me let me know that the one I wanted was on backorder. I ended up going with my second choice."

"I went to my local pharmacy to pick up a prescription. I was surprised that there was a new woman behind the counter."

[V2_B_Errand_Female]

"Last week I went to the bank to deposit my savings. The female who helped me asked for my bank account number. Since I don't have my bank account number memorized, I pulled out my phone to log in to my bank app."

"I went to the library yesterday, looking for something new to read. I wasn't looking for anything in particular and after a while I figured I would ask the female behind the counter for recommendations."

"I went to a concert. The female at the concession stand did not have change for a $20 bill so I had to wait for someone to pay with smaller bills in order to get change back."

"As I drove to work this morning, a female signaled to let me know that my car had a broken taillight. I called the mechanic on the way to my office."

[V3_B_Errand_Woman]

"Last week I went to the bank to deposit my savings. The woman who helped me asked for my bank account number. Since I don't have my bank account number memorized, I pulled out my phone to log in to my bank app."

"I went to the library yesterday, looking for something new to read. I wasn't looking for anything in particular and after a while I figured I would ask the woman behind the counter for recommendations."

"I went to a concert. The woman at the concession stand did not have change for a $20 bill so I had to wait for someone to pay with smaller bills in order to get change back."

"As I drove to work this morning, a woman signaled to let me know that my car had a broken taillight. I called the mechanic on the way to my office."

[V4_B_Errand_Female]

"I lost my wallet last week. I called the restaurant to see if anyone returned it but the female who answered the phone said no one found a wallet."

"I went to the gym in my office building earlier today and noticed it was crowded. I needed a machine to do my workout but there was a female using it so I moved on to the next station."

"I went to the furniture store to buy a couch and the female who helped me let me know that the one I wanted was on backorder. I ended up going with my second choice."

"I went to my local pharmacy to pick up a prescription. I was surprised that there was a new female behind the counter."

V5_B_Dating_Woman

"I listen to many dating podcasts and a common piece of advice is that you have to put yourself out there to meet a woman that is a good match. I am just not sure how to do it."

"My single friends are obsessed with dating. If they thought less about dating, maybe the right woman would just come along."

"I was set up on a blind date with a woman my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation."

"I was invited to a singles mixer last week. I wasn't planning to go but a friend encouraged me, saying it's a great place to meet an outgoing woman."

[V6_B_Dating_Female]

"A friend asked me what I am looking for in someone. I responded that I am looking for a female who shares my values and interests."

"My friend arranged a blind date for me to go on next week. It has been a while since I have met an interesting female so I am excited about this."

"I find it so funny when my friends try to set me up on dates. They insist that they found the perfect female for me and get excited to hear about the date afterwards."

"I want to download a dating app in hopes of meeting a compatible female but I am overwhelmed by the different app options. I am going to do more research and see which app to start with."

[V7_B_Dating_Woman]

A friend asked me what I am looking for in someone. I responded that I am looking for a woman who shares my values and interests."

"My friend arranged a blind date for me to go on next week. It has been a while since I have met an interesting woman so I am excited about this."

"I find it so funny when my friends try to set me up on dates. They insist that they found the perfect woman for me and get excited to hear about the date afterwards."

"I want to download a dating app in hopes of meeting a compatible woman but I am overwhelmed by the different app options. I am going to do more research and see which app to start with."

[V8_B_Dating_Female]

"I listen to many dating podcasts and a common piece of advice is that you have to put yourself out there to meet a female that is a good match. I am just not sure how to do it."

"My single friends are obsessed with dating. If they thought less about dating maybe the right female would just come along."

"I was set up on a blind date with a female my friend knows from work. It went well. We ate at a nice restaurant and had a good conversation."

"I was invited to a singles mixer last week. I wasn't planning to go but a friend encouraged me, saying it's a great place to meet an outgoing female."

Questions about the Speaker (Same questions for Speaker A & B)

What is your impression of the speaker?
(1) Not at all to (7) Very

Likable
Warm
Competent
Educated
Formal
Polite
Feminist
Sexist

What is the speaker's political affiliation?
(1) Liberal to (7) Conservative

Please respond to the questions below.
(1) Not at all to (7) Very

How warm do you feel toward the speaker?
How much does the speaker respect the people they are talking about?

What is the gender of the speaker? *Open text box*

What is the age of the speaker? *Open text box, limit 0-99*

What is the race/ethnicity of the speaker?

Asian or Asian American
Native American or Alaska Native
Black or African American
White or European American
Hispanic or Latinx
Another option *Open text box*

Is there anything else you'd like to share about the speaker? *Open text box*

[At this point participants saw their second set of statements, followed by the same questions about the speaker. Then they move on to the measures and participant demographics.]

Appendix D: Study 2 SCEBS, ASI, and Participant Demographics

SCEBS [refer to page 84 for scale]

Ambivalent Sexism Inventory [refer to page 85 for scale]

Study 2: Participant Demographics

Now we would like to learn more about you. We ask that you complete the following demographics questions about yourself.

What is your gender? *open text box*

What is your sexual orientation? *open text box*

What is your race/ethnicity?

White or European American
Black or African American
Native American or Alaska Native
Asian or Asian American
Native Hawaiian or Pacific Islander
Hispanic or Latinx
Multi-racial (please specify) *Open text box*
Another race/ethnicity (please specify) *Open text box*

What is the highest level of education you have completed?

Completed less than high school
Completed some high school
Graduated from high school
Completed some college
Finished training other than college (e.g., vocational school)
Graduated from 2-year college
Graduated from 4-year college or university
Completed some graduate or professional school
Received masters, professional, or doctoral degree

What is your family's combined yearly income?

Less than $15,000
$15,000 - $25,000

$25,000 - $35,000
$35,000 - $50,000
$50,000 - $75,000
$75,000 - $100,000
$100,000 - $150,000
Above $150,000

How old are you? *Open text box*

Please indicate your agreement with the following statements.
(*1) Strongly disagree to (5) Strongly agree*
I consider myself a feminist.
I identify myself as a feminist to other people.
Feminist values and principles are important to me.
I support the goals of the feminist movement.

What is your political affiliation?
(1) Liberal to (7) Conservative

What do you think this survey is about?

Is there anything you would like to share with the researchers (comments, questions, concerns)?