Personality and Politics in the Comprehension of Gender Stereotypes

by

Stephanie Jacqueline Hammond-Thrasher

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Linguistics
University of Alberta

Examining committee:

Juhani Järvikivi, Supervisor
Stephanie Archer, Examiner
Jennifer Dailey-O'Cain, Examiner

Abstract

Previous research indicates that knowledge about sociocultural norms affects language processing immediately and automatically. One such example is the Stereotype Effect, where sentences containing violations of gender stereotypes take longer to read and are rated as less appropriate than sentences without these violations. Gender stereotypes are embedded in both descriptive adjectives (e.g., *dominant* versus *submissive*) and occupational role nouns (e.g., *doctor* versus *nurse*). The current study takes the first steps to investigate gender stereotype processing at the multi-sentence (i.e., discourse) level, providing an experimental exploration of the language comprehension of both noun- and adjective- level gender stereotype clashes within three-sentence short stories. Participants (N = 215) read 80 short stories pairing male/female gender stereotyped adjectives and role nouns with pronouns either congruent or incongruent with the stereotypical gender of each role noun. This reading task was followed by a yes-or-no decision as to whether the last sentence (i.e., the sentence containing the pronoun) was a sensible continuation of the vignette. In line with previous research, sociocultural world knowledge played an important role in the processing of these social pragmatic stories, where vignettes containing violations of common gender stereotypes were perceived as less sensible than stories without these violations. Importantly, this Stereotype Effect was more pronounced for violations describing male agents fulfilling feminine occupations, indicating that statements that contradict world knowledge about female gender roles have a strong influence on language comprehension. The role of this sociocultural knowledge in the comprehension of gender stereotyped language differed based on the conditions of the stereotype violations. Double stereotype violations (i.e., those at both the adjective- and noun- levels) were perceived as the least sensible, followed by noun-level violations and then adjective-level violations. This indicates that the syntactic role of a piece of gender stereotyped language predicts the degree to which a violation of its gender

influences language processing. Finally, individual differences in political ideology and Honesty-Humility predicted the processing of these stories, indicating that people possessing certain political and personality profiles may allocate more or less resources to sociocultural world knowledge during language comprehension. Notably, high degrees of conservatism predicted more sensitivity to adjective-level violations of male gender roles, while high degrees of Honesty-Humility predicted less sensitivity to gender stereotype violations overall.

Preface

This Thesis is an original work by Stephanie Jacqueline Hammond-Thrasher. The research project, of which this Thesis is a part, received ethics approval from the University of Alberta Research Ethics Board, Project Name “Sensitive or impervious? The effects of gender, perspective, and personality on language processing,” Pro00122792, September 12th, 2022.

Dedication

To Em, Pepita, and my parents: you are the reason I exist.

Acknowledgments

I cannot thank Dr. Juhani Järvikivi enough for the excellent, attentive supervision, insight, inspiration, and endless support. Thank you for your mentorship, trust, and patience. I would not be here today without your encouragement. Thank you to the other members of my Committee, Stephanie Archer and Jennifer Dailey-O'Cain, for your valuable feedback and insights. I am immensely grateful for the support of our Department: thank you to my mentors, instructors, and administrators for lending your time and support. Special thanks to Evangelia Daskalaki, Stephanie Archer, and Jorge Rosés Labrada for the inspiration and friendship over the years. And, of course, thank you to my funding partnerships: this research was hugely supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) Partnership Grant (Words in the World, 895-2016-1008) and by my own SSHRC Canada Graduate Scholarship - Masters funding.

Thank you to my colleagues, Dalia Cristerna Román, Veranika Puhacheuskaya, and Hannah Lam, for helping me dip my toes into experimental design in *PsychoPy*. Kasandra and Tyler, academia (especially statistics!) would have been too daunting to face without you by my side. Your humanity, compassion, and never-ending jokes fuelled this work. Many thanks to Lindsay Griener, our talented Lab Coordinator at the Centre for Comparative Psycholinguistics (CCP), for enduring my frustrating data glitches and lending me support when things looked endlessly complicated. Of course, thank you to all the members of the CCP and fellow Graduate students for listening to my rambling, boiling me water for tea, and reminding me to breathe.

Now, I must thank my family and friends: Mum, Dad, and Freya, your endless comfort, enthusiasm, and vulnerability made all of this possible. You’re my everything. To my brave Kate: talking with you dissolved my imposter syndrome, opened my eyes, and always lifted me

up. Emma, Sam, Alessia… you gave me laughter and softness when I needed it most. To the boys, Ed, Yogi, Jono, and Japes (and, of course, Delaney): you kept me laughing and dancing throughout everything. Thanks to the Sharp family for being my cheerleaders and, of course, Eden for years of friendship, support, and mischief. To my beloved Bestefar, Bestemor, and Papa: I wish you I could hug you now. To Nana, Grandma, Grandpa Jerry, and my numerous relatives across the globe: I hope you can love this work enough for all the people I wish were here to share it. To Adam, Stan, and Nix: your pursuit of academia always inspired me. Thanks for showing me the appeal and cheering me on. Thank you, Andriy, for showing me that I deserve good things like this. To the powerful force that is Joanna: you taught me how to do this with softness and grace.

To my heart, Pepita. Thanks for your endless cuddles on my lap, tearing me away from my desk for walkies when I needed it most, and for saving my life. To Peaches and Bug: thank you for not completely destroying this work with your paws. You can sleep on my keyboard anytime, now. Lastly, to my Em: you mean more to me than any words can express. Thank you for keeping me young, loved, and inspired. I cannot wait to live our future together, always

Table of Contents

List of Tables

List of Figures

Chapter 1

1. Introduction

Communication is a complex task that minimally involves a speaker or writer conceiving a thought, translating it into a phrase that represents those thoughts, and producing a string of sounds or letters that expresses that thought. The listener or reader must then take in that string of sounds or letters, tie them to their linguistic representation, mentally understand this message, and conceive a response to the speaker or writer's statement.

The key elements used to comprehend communicative messages were historically believed to be primarily syntactic (i.e., structural) and lexical (i.e., word-level), and it was these features that were thought to be processed automatically prior to any semantic meaning of words within a phrase (Acheson & MacDonald, 2009). However, the field of psycholinguistics deftly employed online research methodologies to find that there are a multitude of other factors that immediately and spontaneously influence language processing. Evidence from these studies indicate that a wide variety of linguistic and extra-linguistic elements are involved in language comprehension, including communicative context, speaker traits (e.g., voice-based inferences, accentedness, and affect), and comprehender characteristics such as world knowledge, personality, political ideology, empathy, affect, gender identity, and disgust sensitivity (Calkins, 2022; Puhacheuskaya & Järvikivi, 2022; Hubert Lyall & Järvikivi, 2021; Grant et al., 2019; Stott, 2019; Koenig & Eagly, 2014; Van Berkum et al., 2013; Van den Brink et al., 2012; Paczynski & Kuperberg, 2011). Furthermore, these complexities are highly involved in the processing of socially-charged language, which refers to language containing socially constructed elements, such as gender stereotypes, irony, and lies (Hammond-Thrasher & Järvikivi, 2023; Puhacheuskaya & Järvikivi, 2022; Fiawornu, 2022).

In this thesis, I present a systematic exploration of the extent to which the processing of written sentences containing gender stereotypes is affected by a reader's own personality traits, political ideology, and gender identity. Many studies have indicated that sentences directly contradicting common gender stereotypes present cognitive costs to language processing (see Marrville, 2017, and Canal et al., 2015, for an overview). However, studies investigating gender stereotype processing have focused solely on interruptions to language processing rather than the comprehender-specific factors that may modulate these effects. For this reason, this thesis stems from a desire to understand the complex social factors that interact with gender stereotyping and the role of these intersections with language comprehension. Across a virtually-accessible experimental paradigm involving three-sentence vignettes, this research aims to add to the literature that asserts that language processing is a complex task implicating various social and personal elements.

Before outlining the current study, this chapter will provide an overview of literature examining the process of language comprehension, particularly during reading comprehension; the concept of stereotyping and gender stereotype constructions; previous research on gender stereotype processing; and comprehender-specific individual differences in personality traits, political ideology, and gender identity and their role in language processing. The following chapters will discuss the experiments, their analyses, and results in detail. Finally, I will discuss the general findings of this study and their implications for future research.

1.1 Language Processing

To accomplish the goals of this research, it is first important to establish a framework through which to view language processing as a whole. Within this Thesis, I follow the modern linguistic theory that language processing involves general cognitive processes and rather than a solely

linguistic mechanism, meaning that it relies on skills and knowledge also used for non-linguistic tasks, such as working memory, world knowledge, and perception (Glenberg et al., 2009; Diessel, 2019). In other words, language comprehension is a domain-general process.

Regarding the comprehension of individual sentences, research shows that individuals process input from auditory and/or visual language sources immediately upon their presentation; this information is quickly adapted into a mental representation of the discourse in a matter of milliseconds (Canal et al., 2015). Under this domain-general view, sentences are processed online and in an incremental fashion, enabling comprehenders to form assumptions about the language they are perceiving in real-time (de Hoop & Lamers, 2006). It is this assumption that I apply to this research: language processing is immediate, incremental, and involves cognitive resources beyond purely linguistic structures. That is, we know that language processing involves widely-distributed cognitive mechanisms throughout both cerebral hemispheres in order for individuals to reconcile lexical, syntactic, semantic, pragmatic, and contextual information during communication (Caplan, 1992). In the context of the current reading study, is important to understand that this evidence is consistent for both spoken and written language comprehension, meaning that similar parsing, comprehension, and inference-making mechanisms are implicated in both forms of language processing (Nieto et al., 2022; Lechevalier et al., 1989).

1.1.1 Situation Modeling

I have discussed the basic view of language processing taken in this paper; now, I will expand upon this in order to demonstrate what kind of information readers are able to cognitively model based on the language input to which they are exposed. Early evidence suggests that readers produce a simulated, mental model of the language they are processing in real-time; these are known as situation models (Zwaan et al., 1998). This means that, upon reading text, individuals

immediately begin to form a mental representation of the events described in the linguistic input. In fact, experiments involving sentence-picture verification tasks (i.e., reporting whether an image reflects a scenario presented linguistically) show that individuals based their mental situation models on inferences made about character reference, location, and orientation (Connell, 2007). For example, consider the sentences in (1).

(1) A. The eagle was in the sky.

B. The eagle was in the nest.

When participants were shown drawings of an eagle with outstretched wings, they had a greater chance of reporting that the image matched the former sentence versus when they saw an image of an eagle with folded wings (Zwaan et al., 2002). This indicates that readers produce simulated, spatial models based on the content of linguistic input and their world knowledge about the state of affairs in the real world. This situation modeling process has been implicated in further experiments across the dimensions of time, space, protagonists, causation, and motivation (Kaup et al., 2015).

1.1.2 Inference Restructuring

Now, the question is raised: what happens when a reader is presented with linguistic information that contradicts the current set-up of their ongoing situation model? In other words, how do readers update their situation models when they encounter information that requires the restructuring of their mental representations? Research indicates that updating models can occur both incrementally, where the model is altered as new information is processed, and globally, where additional information requires a new model that replaces a prior one (Kurby & Zacks, 2012). Here, a shift in the mental representation of a described situation can lead to major

changes in the interpretation of linguistic input, including re-evaluating the entire view of the situation and, at times, restarting the entire comprehension process. This disruption in language processing can manifest itself in longer time taken to read text as well as large changes in pupil size during reading (Hubert Lyall & Järvikivi, 2021). Overall, this disruption is further evidence for the incremental and immediate process of language comprehension and its inferential nature.

Despite the potential red flags that accompany forming linguistic predictions, inferential language processing has many potential benefits for communication. Take, for example, the sentence in (2), provided by Paczynski and Kuperberg (2012).

(2) The bass was strummed by the guitarist during the song.

When individuals read this sentence, the time taken to process it is significantly shorter than if the subject of the sentence is replaced without musical implications, such as *gravedigger*; comparatively, reading times for a subject such as *drummer*, which is semantically related to the musical theme of the sentence yet more distant than the subject *guitarist*, are slightly longer than for the sentence in (2) yet significantly shorter than when the subject is *gravedigger* (Paczynski & Kuperberg, 2012). Here, inference-based facilitation is evident: comprehension of the sentence in (2) is facilitated by predictions made in situation models upon encountering the musical theme of the onset of the sentence. Indeed, language processing is still facilitated for any subject of the sentence with any musical traits. Conversely, comprehension is inhibited when the agent of the action is revealed to be contextually inappropriate because it is inconsistent with the preceding sentential content. Hence, potential clashes in expectations about linguistic input pose threats to real-time language processing, but inference formation is still a rewarding process because it

facilitates the comprehension of linguistic input and commitment to an expected sentential outcome (Camblin et al., 2007).

1.2 Gender Stereotypes

I have discussed the process of language comprehension during reading in Section 1.1 of this Thesis. Now, it is necessary to outline the specific concept of concern for this study: gender stereotypes. In essence, gender stereotypes are generalizations about the performance of gender, outlining the characteristics, behaviors, and roles typical to men and women (Hentschel et al., 2019). Early research on gender stereotypes indicates that they are constructed through social action and derived from historically discrepant distributions of men and women into roles within the family, home, workplace, and society at large (Eagly, 1987). The foundational sociological framework of Social Role Theory explains that these discrepant distributions are socially constructed: gender roles enforce the performance of gender within stereotyped roles, which in turn enforces these gender roles yet again (Eagly, 1997). Hence, gender stereotypes are embedded in occupational role names within everyday communicative vocabulary.

Interestingly, Social Role Theory is perpetuated through language usage. Social Dominance Theory, coined by Sidanius and Pratto (1999), states that mechanisms at the institutional level reinforce existing inequalities. Under this view, gender stereotypes could, and are, reinforced by the use of gender stereotyped language. It is precisely this language that I focus on in this research. After a brief overview of the stability of gender stereotypes, I will introduce the specific nominal and adjectival linguistic constructions that embed gender stereotypes in their very cores.

1.2.1 Stability of Gender Stereotypes

While there is no doubt that modern society has progressed towards a state of gender equality, there is also no question that full equality has yet to be achieved (AAUW, 2022). Research continues to display stability in gender stereotypes over time: men are continually characterized as more agentic while women are characterized as more communal (Hentschel et al., 2019; Haines et al., 2016; Abele & Wojciszke, 2014). Similarly, occupational roles, such as those in (3) and (4), are perceived to be more masculine versus feminine, respectively (Marrville, 2017; Lippa et al., 2014).

(3) A. Truck driver

B. President

C. Doctor

(4) A. Florist

B. Social worker

C. Dental hygienist

Interestingly, certain gender stereotypes have been shown to be significantly more stable over time: in a comparison between gender roles across 1983 to 2014, Haines et al. (2014) indicated that the performance of feminine roles by male agents has become less preferred than the performance of masculine roles by female agents. Hence, despite changes in the participation of men and women in nontraditional roles, gender stereotypes are still prevalent components of modern society. It is precisely the processing of these gender stereotyped occupational role names that are the subject of this study.

1.2.2 Gender Stereotyped Language

Above, I introduced Social Role Theory and the concept of Social Dominance. Now, I will reference two linguistic constructs as embodiments of these institutional mechanisms that reinforce gender roles. Firstly, occupational role names, such as those in (3) and (4) above, bear information beyond the occupation to which they refer: these linguistic units carry social connotations about gender stereotypes based on behaviors constructed by and learned through socially common behaviors associated with men and women. For this reason, interdisciplinary research often explores gender stereotypes using materials containing gender stereotyped role names (see Banaji & Hardin, 1996, Gaucher et al., 2011, Marrville, 2017, and Hammond-Thrasher, 2020, for an overview). Not all occupational role names are stereotyped, though: there are many examples of role names that do not carry gendered connotations. Examples collected from the occupational stereotyped norms created by Gygax & Gabriel (2008) are presented below.

Table 1. Occupational Role Names by Gender Stereotype

Masculine	**Feminine**	**Neutral**
Butcher	Babysitter	Student
Police officer	Dancer	Spectator
Fire fighter	Receptionist	Concert goer

Importantly, these stereotyped role names must be distinguished from definitional role names, such as *aunt* and *uncle*, which carry irrefutable gendered meanings. In the examples above, the occupational role names do not host any explicit gendered definitions: they have nuanced gender

stereotyped meanings precisely due to the perceived distribution and performance of gender roles in society but yet still may apply to referents of any gender (Gygax et al., 2016). Notably, the Neutral role names bear no gender stereotype information nor do they reinforce social roles.

Interestingly, occupational gender stereotyping is influenced by more than these role names. That is, gender stereotypes are embedded in other linguistic elements, such as adjectives. This second linguistic construct is particularly salient in job advertisements within career spaces, which are heavily influenced by institutional mechanisms implicated by Social Dominance Theory. Indeed, gendered descriptions in job advertisements have been well-documented by scholars such as Gaucher et al. (2011). In their corpus study, the authors assert that adjectives employed within job descriptions and advertisements carry gender stereotype connotations that are subtle yet systematic. Results from their archival analyses indicate that job advertisements for male-dominated roles employed masculine stereotyped wording through the saturation of adjectives such as *strong, competitive,* and *dominant*; advertisements for female-dominated roles were contrastively saturated with adjectives such as *supportive, understanding,* and *interpersonal*. Yet again, linguistic evidence from career-level constructs points to implications on the maintenance of gender-based inequality.

1.2.3 Perception of Gender Stereotype Violations

Given the relative stability of gender roles and the prevalence of gender stereotyped language, which support the assumptions of theories such as Social Role and Social Dominance, it follows that individuals who violate these gender stereotypes would be rather salient. Importantly, this has been demonstrated in the realm of gender stereotyped occupational role names. Research indicates that individuals consciously perceive gender stereotype violations in the real world (see Diekman et al., 2004). A study performed by Heilman et al. (2004) found three key observations

about gender stereotype violations from their series of experiments. Firstly, when women are described as successful, they are less liked than males described in the same situations. Secondly, these negative and derogatory reactions only occur when these women's successes are in the scope of male stereotyped occupations, such as in roles like *CEO*, *Boss,* and *Leader*. Finally, the success of these theoretical women in male stereotyped roles can induce outcomes with implications on their careers, both in the realm of job performance evaluations and in recommendations for rewards and promotions. Importantly, both male and female participants showed the same reactions to these descriptions of successful women. An additional study by Rosette et al. (2015) found that participants viewed men as less competent than women when these men were described as fulfilling more female stereotyped occupations, such as in roles like *Florist* and *Social Worker*. Here, it is apparent that gender stereotype violations at the role name level are salient in the real world and pose serious implications for social performance within career spaces.

Importantly, the salience of gender stereotype violations has also been implicated in the realm of stereotyped adjectives. Gaucher et al. (2011) used their collection of gender stereotyped adjectives to test the consequences of highly stereotyped wording in job advertisements. Their results showed that job advertisements that included strong masculine stereotyped descriptions led participants to rate hypothetical male characters as more appropriate for these roles compared to female characters. Furthermore, female participants in this study found these masculine-sounding jobs less appealing and rated hypothetical female applicants as less successful at these jobs compared to female applicants to advertisements containing stereotypically feminine adjectives. Here, it is yet again apparent that gender stereotype

violations at the adjectival level are quite salient in real-world situations and hence pose serious implications for social performance within career spaces.

A further study conducted by computer scientists Hoyle et al. (2019) used machine learning to survey 3.5 million English novels written between 1900 to 2008 to establish a database of common adjectives used to describe female versus male protagonists. This massive sample provided irrefutable evidence of the gender bias of descriptive adjectives, with a brief sample of feminine versus masculine stereotyped adjectives in Table 2.

Table 2. Adjectives by Gender Stereotype

Feminine	**Masculine**
Lovely	Righteous
Classy	Honorable
Vibrant	Bigoted

Given this evidence, it is extremely clear that there is a strong presence of gender stereotyped adjectives in the English language which pose implications for language processing, particularly for the formation of inferential mental representations of discourse.

1.2.4 Stereotype-Based Inferencing

I have expressed that gender stereotyped language can take the form of both occupational role names and descriptive adjectives, both of which contain salient gender role implications within career spaces. Now, I turn back to language processing in order to connect these linguistic constructs to inferential language comprehension. In the previous section of this introduction, I introduced the concept that language processing relies on factors beyond simple linguistic

knowledge, meaning that readers make inferences about upcoming information in discourse using both the content of the linguistic input and the context of the communicative constructs. We know that situation models rely on world knowledge tied to these segments of linguistic input: this means that readers form expectations about discourse protagonists in a similar fashion. This means that, upon encountering a protagonist in discourse, readers pay particular attention to the description of this character to build robust situation models that allow them to seamlessly form inferences about their upcoming actions; these situation models include information about a character's gender. I will first begin with a discussion of the inferential processing that occurs when readers encounter gender stereotyped occupational role names. Then, I will discuss assumptions about the inferential processing of gender stereotyped adjectives.

1.2.4.1 Inferential Processing of Occupational Role Names

Because gender stereotypes are encoded heavily in language use, this means that individuals automatically make inferences regarding character gender upon the presentation of gender stereotyped occupational role names. Evidence suggests that the gender of a character is incorporated into the mental representation of a text even in cases when it is not explicitly stated. A study by Oakhill et al. (2005) asked participants to say whether two terms could refer to the same person. For example, participants saw the terms *uncle* and *nurse*, the former of which is definitionally male and the latter of which is stereotypically female. Here, response times to the question were inhibited when the gender of the stereotype did not match the definitional gender of the paired role name; the example just mentioned saw slower response times than when *uncle* was paired with *doctor*. Here, there were no images of female nurses provided, but the participants still hesitated to group *uncle* and *nurse* together when the words were presented as a pair due to restructuring of their mental representation of *nurses* as stereotypical females.

Inference formation based on gender stereotype information is done automatically and immediately and can even override grammatical processing. A study performed by Osterhout et al. (1997) asked gender-balanced participants to read sentences containing gender stereotype (e.g., *nurse* - *he*) and gender definition (e.g., *uncle* - *she*) violations while recording their brain activity. Both forms of clashes elicited immediate and significant cognitive activity, with gender stereotype clashes predicting persisting effects even when participants judged the sentences to be completely acceptable. A further study by Guerra et al. (2021) used an eye-tracking paradigm to investigate participants' visual gazes to potential protagonists while hearing sentences containing gender stereotype clashes. Observed gaze patterns indicated that participants used gender stereotype information immediately to predict the agents of actions mentioned in sentences. An additional study performed by Molinaro et al. (2016) found that gender stereotypes predominantly guide language processing even more noticeably than structural expectations. Manipulating stereotype- and grammar-based cues in their sentence processing study, the authors found that gender stereotype knowledge overrode syntactic cues, where participants used this information to parse information within each sentence rather than relying on grammatical constructions. The three studies summarized above highlight the immediacy and automaticity with which gender stereotype information is activated during language comprehension and the overall importance of this information during language processing.

It is vital to note that gender-based inference in English has been widely studied through the use of anaphoric constructions, where the subject of a sentence is referenced later down in the sentence through a referential pronoun. Examples of these constructions are seen in (5), where the subject of the sentences is *the secretary* and the referential pronouns are *she* and *he*, respectively.

(5) A. The secretary yawned because she was tired.

B. The secretary yawned because he was tired.

Here, a referent in the beginning of the sentence is referred to by a pronoun lower in the sentence. The presentation of the occupational role noun prompts the immediate mental representation of the *secretary*, including this noun's stereotyped gender; upon encountering the pronoun, readers are prompted to link the pronoun's definitional gender to its antecedent (i.e., *secretary*) in the previous sentential context. Studies show that sentences like (5) A. are processed faster than sentences like (5) B. because the stereotypical gender of *secretary* is female (Banaji & Hardin, 1996; Marrville, 2017). That is, sentential set-ups like the ones in (5) B. force readers to build mental representations that are explicitly challenged upon the presentation of an undeniable pronoun. These anaphoric constructions present psycholinguists with a valuable tool to explore gender stereotype processing. Importantly, this added processing difficulty that occurs when a pronoun's explicit gender is incongruent with the stereotyped gender of a previously-mentioned occupational role name will be referred to henceforth as the Stereotype Effect.

Furthermore, the Stereotype Effect has been systematically investigated through offline psycholinguistic measures. A recent study by Hammond-Thrasher and Järvikivi (2023) presented participants with complete sentences constructed across four conditions, each of which are exemplified in (6).

(6) A. Congruent female-stereotyped sentences (e.g., *The dietician recommended a supplement because she knew it was good for health.*).

B. Congruent male-stereotyped sentences (e.g., *The plumber dislodged the obstruction because he used a strong solvent to loosen it*.).

C. Incongruent female-stereotyped sentences (e.g., *The librarian flipped through the book because he was looking for a certain page*.).

D. Incongruent male-stereotyped sentences (e.g., *The farmer cooked a big feast because she was entertaining the neighbors*.).

Participants were asked to rate sentences such as those in (6) on 5-point Likert scales for both their Correctness and Appropriateness. The authors found that Incongruent sentences, like those in (6) C. and D., were rated significantly lower on both scales compared to Congruent sentences. This provides strong evidence for the strength of the Stereotype Effect: sentences containing clashes with gender stereotyped role names are strongly salient, even during self-report ratings. More interesting, however, is the fact that the authors also found that sentences describing male agents fulfilling female stereotyped roles (i.e., female stereotyped role names paired with male pronouns, such as in (6) C. were rated significantly lower across both scales. This result points to the particular salience of incongruent female-stereotyped sentences, suggesting that individuals perceive men occupying typically feminine occupations as more salient than women occupying typically masculine occupations. This is vital evidence in the consideration of the Stereotype Effect: not only does it indicate the strong salience of gender stereotype clashes in language use, but it also suggests a nuanced gendered imbalance in the relative strength of the Stereotype Effect. The authors attributed this character gender-specific Stereotype Effect to Social Role

Theory: stereotypes about males appear to be undergoing less change than stereotypes about male due to the larger-scale social changes in behavior that occur with the expansion of women's roles in the workplace.

Importantly, there is limited evidence for this character gender-specific Stereotype Effect. One EEG study by Grant et al. (2019) presented participants with sentences describing feminine (e.g., fashion) and masculine (e.g., sports) topics spoken by a male and a female speaker. Half of the participants heard stereotype-congruent sentences (i.e., for the male speaker, semantic errors about fashion and no errors about sports) while the other half heard stereotype-incongruent sentences (i.e., for the male speaker, semantic errors about sports and no errors about fashion). Results found a larger N400 ERP signature in participants listening to stereotype-incongruent sentences compared to congruent ones. Notably, the authors also found that incongruent stimuli presented in a male voice elicited greater N400 effects compared to those spoken by a female voice. Again, this effect must be further investigated in future research.

1.2.4.2 Inferential Processing of Gender Stereotyped Adjectives

As we know, gender stereotypes are encoded heavily in language use. Given that linguistic input is processed immediately and incrementally in order for readers to form robust situation models of ongoing discourse, it can be assumed that gender stereotyped adjectives influence the perception of characters as they are described in strings of words. That is, a character described as possessing stereotypically feminine traits would likely lead comprehenders to produce a more feminine protagonist within their mental representation of discourse; conversely, an ambiguous character described as possessing stereotypically masculine traits would likely influence the production of a more masculine protagonist within a reader's situation model. For example, consider the sentences in (7).

(7) A. The compassionate student walked across campus.

B. The dominant student walked across campus.

Given the evidence that gender stereotypes are salient in language use, readers would likely infer that the sentential protagonists in (6) are more feminine versus masculine, respectively. While this is a sound hypothesis, there is little research into the online processing of gender stereotyped adjectives. This necessitates both an investigation of the relative salience of gender stereotyped adjectives (i.e., whether gender stereotyped adjectives are perceived as strongly feminine or masculine by readers) and an exploration of their influence on inferential language processing.

To perform such research, it is paramount to understand that adjectives influence language comprehension through immediate, incremental incorporation into mental representations of discourse: Sedivy et al. (1999) performed a series of experiments to produce this evidence. The authors presented participants with simple instructions containing descriptive adjectives coupled by visuals that included two objects. When told to *Pick up the blue comb*, participants quickly selected the blue comb in the visual field when presented with a blue comb and a red pen. However, participants were significantly slower to select the blue comb when it was presented alongside a blue pen. Eye-tracking evidence from the same study indicated that, in the first scenario, participants made a short glance at the red pen during their processing of the instructions; in the second scenario, participants made longer glances at the blue pen, often making repeated glances between the two objects. This evidence suggests that comprehenders incorporate descriptive adjectives into their incremental language processing in an immediate fashion and experience a disruption in this comprehension when the descriptive adjective can

ambiguously refer to two visual referents. That is, adjectives are incorporated during language processing into a comprehender's mental representation of discourse, allowing them to make inferences about suggested traits of a referent. In the context of the present study, this is vital information when investigating the inferential processing of gender stereotyped adjectives.

1.2.4.3 Inferential Processing of Stereotyped Adjective-Noun Pairings

It is vital to acknowledge that there has been little research conducted on the inferential processing of gender stereotyped adjective-noun pairings. Indeed, the majority of psycholinguistic explorations of the processing of gender stereotypes has focused on occupational role names within the confines of the aforementioned Stereotype Effect; no such effect has been produced for gender stereotyped adjectives. However, two facts are known: firstly, both adjectives and nouns are incorporated into readers' situation models during language processing; secondly, both adjectives and nouns carry gender stereotype information, and this information has been shown to influence inferential processing in real-time. Given this knowledge, it is reasonable to assume that gender stereotyped adjective-noun pairings may pose implications of interest for inferential language processing. Consider, for example, the sentences in (8).

(8) A. The ditsy florist walked quickly because she was running late.

B. The ditsy florist walked quickly because he was running late.

Our knowledge of the Stereotype Effect predicts that (8) B. will take longer to read because readers must update their situation models of the sentence when they encounter the pronoun *he*, which is incongruent with the gender stereotyped role name, *florist*. When paired with the female

stereotyped adjective, *ditsy*, readers would theoretically have stronger evidence within their situation models for the femininity of the sentence's subject, hence producing a stronger Stereotype Effect because of the greater mismatch between the adjective-role name pairing and the pronoun.

What is perhaps more puzzling is the question of what happens when readers encounter multiple gender stereotype mismatches within a sentence. Consider the sentence in (9).

(9) The dominant florist walked quickly because he was running late.

Here, this sentence directly contrasts those in (8): the adjective is stereotypically masculine, the role name is stereotypically feminine, and the explicit gender of the sentence's subject is male. In other words, there exists a mismatch between both the gender stereotyped adjective and the role name as well as the gender stereotyped role name and the pronoun. In this situation, the existing evidence indicates the presence of the Stereotype Effect when solely considering the role name and pronoun pairing. However, evidence also indicates that the information encoded by the gender stereotyped adjective will also affect processing. Perhaps the masculinity encoded on the adjective will conflict readers, lengthening reading times; contrastively, perhaps it is precisely the masculinity encoded on the adjective that will facilitate readers' processing of the pronoun, hence speeding up reading times. Here, real-time research methods are necessary to determine the exact time course of processing for gender stereotyped adjective-noun pairings in order to illuminate the phenomena involved in the comprehension of these nuanced sentences. This is precisely the approach taken by the present study.

1.3 Individual Differences in Language Processing

Throughout Section 1.2 of this paper, I provided a detailed overview of gender stereotypes, including the distribution and salience of gender stereotypes under the views of Social Role and Dominance Theories and a survey of gender stereotyped language and its influence on inferential language processing. Now, I turn to the incredibly relevant role of individual differences in language comprehension. Individual differences refer to the unique, differential traits of individuals (Eysenck, 1966). In social science research, these traits can include the variables of age, sex, gender identity, personality, political ideology, intelligence, familiarity, and more. Revelle et al. (2010) sum up the realm of individual difference research tactfully below:

> "Individual differences in how we think, individual differences in how we feel, individual differences in what we want and what we need, individual differences in what we do. We study how people differ and we also study why people differ."

For psycholinguists, the question of how these individual differences affect our ability to communicate is particularly important. Indeed, recent evidence suggests that the comprehension of sociocultural elements of language (i.e., language that is conditioned by social norms, such as gender stereotyped adjectives and nouns, idioms, and verbal irony) is affected by individual differences in participants' personality traits, political beliefs, and overall identity (see Van Berkum et al., 2009, and Van den Brink et al., 2012, for an introduction). I will begin by providing an overview of the three individual difference measures I investigate in this research and provide previous evidence of their significance for language processing.

1.3.1 Personality

Personality psychology seeks to understand the many dimensions of human personality traits, which are considered to be enduring characteristics that encompass an individual's major interests, drives, values, abilities, emotional patterns, and more (Friedman & Schustack, 2016). Historically, personality psychologists found consensus on an overarching personality structure known as the five-factor approach, which measures personality across the five dimensions of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (McCrae & John, 1992). However, recent criticisms of this approach indicate that the scale's factor analysis methodology lacks precision: the dimensions are generally acknowledged to be too broad to be cohesive, meaning that the assessment lacks internal validity (Block, 1995). For this reason, more recent endeavors to operationally investigate personality traits have yielded methodologically strong assessments with robust internal and external validity.

A tool recently utilized by psycholinguists to measure individual differences in personality traits is the HEXACO Personality Index - Revised (HEXACO PI-R), which assesses six major dimensions of personality: Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience (Ashton & Lee, 2009). Numerous meta-analyses indicate that all dimensions exhibit strong reliability and self-observer agreement; importantly, there exist no strong correlations between the six dimensions, meaning that they accurately assess individual facets of personality (see Zettler et al., 2020 and Moshagen et al., 2019). Because of its methodological strength and its frequent usage in linguistic experimentation, it is the HEXACO PI-R that I employ in this research. An overview of the six personality dimensions encapsulated by the HEXACO PI-R assessment is presented in Table 3,

which is adapted from Ashton & Lee (2009). Importantly, low scores on each dimension indicate the opposite of each description below.

Table 3. Descriptions of Each Dimension of the HEXACO PI-R

Dimension	**Description**
Honesty-Humility	Individuals with high scores feel little temptation to break rules, feel no special entitlement to high social status, wealth, or luxury, and avoid manipulating others for personal gain.
Emotionality	Individuals with high scores rely on emotional support from others, feel immense empathy and sentimental attachments to others, and experience frequent fear of physical dangers and anxiety in response to life stress.
Extraversion	Individuals with high scores experience positive self-esteem, assume leadership roles with confidence, enjoy social interactions, and generally maintain enthusiasm and energy.
Agreeableness	Individuals with high scores are naturally forgiving and lenient to others' wrongdoings, have the ability to compromise, cooperate, and manage their temper easily under stress, and do not feel tempted to judge others.
Conscientiousness	Individuals with high scores feel tempted to organize their time and physical surroundings, strive for accuracy and perfection at all times, and use discipline and deliberation when making any decision.
Openness to Experience	Individuals with high scores are passionate about the beauty of art and nature, maintain curiosity about intellectual topics, do not struggle to use their imaginations, and are interested in unusual ideas and people.

Early psycholinguistic research implicated that personality traits such as empathy, openness, and extraversion can, to a degree, predict language processing in real-time. Van den Brink et al. (2012) used an electroencephalogram (EEG) paradigm, where participants' neural activity was measured while they listened to several sentences. Their results showed that words (e.g., *teddy bear*) in sentences clashing with a character's socially expected stereotypes (e.g., *I*

cannot sleep without my teddy bear in my arms, spoken by a male adult voice) elicited a strong N400 effect, which is an event-related brain potential (ERP) that indexes cognitive effort during incremental integration of stimuli to the context in which it occurs; for example, when listeners encounter semantic or world knowledge violations (Hagoort et al., 2004). In other words, listeners' language processing was disrupted upon hearing the word *teddy bear* because of its contextual mismatch with the speaker's adult male identity. Participants in this study also completed Baron-Cohen and Wheelwright's (2004) Empathy Quotient Questionnaire (EQQ), which measures an individual's ability to empathize with others. Key to the authors' findings was the fact that the N400 ERP for these stereotype violations was predicted by the listeners' EQQ scores: high empathizers showed a large N400 effect compared to low empathizers, meaning their online language processing was more disrupted than their less empathetic counterparts. Importantly, participants' EQQ scores did not predict their N400 ERPs for semantically incongruent sentences (e.g.,*Dogs sometimes chase teas*). The authors presume that individuals with a strong ability to empathize are more attentive to socioculturally relevant information during language comprehension, resulting in heightened sensitivity to stereotype violations.

Subsequent research has shown that personality traits beyond empathy affect the processing of sentences containing sociocultural violations. Hubert Lyall and Järvikivi (2021) presented participants with spoken sentences during an eye-tracking paradigm. Results indicated that more introverted participants, as measured by the five-factor approach, showed greater pupil dilation in response to sociocultural clashes, such as hearing the sentence *I sometimes buy my bras at Hudson's Bay* spoken by an adult male speaker. These results suggest, yet again, that

individuals with different personality profiles exhibit different patterns of cognitive resource usage during online language comprehension.

A study by Hammond-Thrasher and Järvikivi (2023) was the first of its kind to systematically explore the role of personality traits, measured using the HEXACO PI-R, during gender stereotyped language comprehension. In this study, participants read thirty-two experimental sentences containing gender stereotyped role names resolved by explicit pronouns (e.g., *The dancer waved in the air because she/he was proud of the performance*); participants were then asked to rate the sentences on 5-point Likert scales for Acceptability and Correctness. Results from this study confirmed that the Stereotype Effect is indeed perceptible during offline (i.e., not in real-time) language comprehension: participants rated sentences containing incongruent explicit pronouns significantly lower across both rating dimensions, with sentences describing male agents fulfilling stereotypically feminine roles rated lowest overall. The results also indicated that personality traits predict these offline ratings: participants who scored high in Openness to Experience and Conscientiousness were more sensitive to stereotype violations, especially these character gender-specific clashes, than their less Open and Conscientious counterparts. These findings were the first of their kind to implicate individual differences in personality traits on the offline perception of gender stereotype violations: the Stereotype Effect and its predictability by personality characteristics are indeed perceptible beyond milliseconds-long reading times.

Overall, it is evident that individual differences in personality traits can predict the inferential processing of sentences containing sociocultural language. Note that the specific personality profiles involved in these gender stereotype processing studies have yet to become explicit to researchers: there is a great need for replication and further systematic evaluations of

these effects. It is also important to note that disruptions in online language processing and low ratings of Acceptability during offline language comprehension do not equate to individuals with certain personality profiles finding sociocultural language violations as incorrect, inappropriate, or impossible; rather, these individuals are more perceptive of sociocultural world knowledge during language comprehension. In other words, it has been shown that individuals with certain personality profiles strongly rely on sociocultural world knowledge, such as gender stereotype information, during language processing: hence, they perceive sociocultural violations with increased salience.

1.3.2 Political Ideology

Political science is a broad social science discipline that includes the study of political ideology, which can be understood as a set of socially determined ethical ideals associated with an individual's belief in how society should work (Bouchard & Mcgue, 2003). Often deemed "the most elusive concept in the whole of social science," the study of political ideology (i.e., aligned with some position on a general political spectrum) must be distinguished from political strategies, which are the positions around which a political party or leader is grounded (McLellan, 1995). In other words, the social science of understanding political ideology is not grounded in any one political framework (e.g., United States political parties such as Republicans or Democrats); instead, it is interested in understanding free-standing political ideas about society and its organization. Importantly, research has indicated that political ideology is grounded in both culture and family inheritance, which may be due to behavioral neighborhoods (i.e., surrounding environments of political beliefs) or pure genetics (see Settle et al., 2009, Bouchard & Mcgue, 2003, Eaves & Eysenck, 1974, for an overview). Regardless of their origin, political ideologies appear to be stable from adolescence onward (Krosnick, 1991).

Social scientists have long been interested in measuring individuals' political ideologies for the purposes of psychological, linguistic, and sociological individual differences research. There is consensus that a two-factor approach for political ideology is reliable and possesses strong internal and external validity: political ideology must be measured both economically and culturally to assess individual differences in attitudes toward specific political topics (Alves & Porto, 2022). For this reason, questions on any political ideology questionnaire must address a broad range of economic and cultural topics regarding social organization. One such example is Wilson and Patterson's (1968) Conservatism Scale (C-Scale), which questions participants on their agreement with controversial topics ranging from economic theories to sociopolitical issues. The C-Scale has been widely cited as the first reliable and valid measure of political ideology within social science scholarship (see Feldman & Johnston). Derived from this assessment is the Political Ideology Questionnaire (PIQ) from Grenier's (2006) work in the School of Social Work at Louisiana State University (see full work in Ropple et al., 2020). The goal of this measure is to assess the political ideologies of adults in Western, democratic societies across a breadth of topics ranging from social (e.g., healthcare, mental health, child care, criminal justice, etc.) to economic (e.g., poverty, inequality, free market economies, welfare, etc.). With numerous reliability and validity related strengths, the PIQ is one of the premium tools for assessing political ideologies in North America (Ropple et al., 2020). Because of the context of the present research, it is this scale that is utilized in this study.

The PIQ uses six-point Likert scales for two batteries of questions; the former assesses how for-or-against participants are for a variety of socially-charged issues (examples are in Table 4.1), while the latter asks participants to rate how much they agree or disagree with several

political statements (examples are in Table 4.2). High scores indicate more Conservative beliefs (i.e., more traditional, rigid thinking), and questions are either normally or reverse coded.

Table 4.1 For-or-Against Topics in the PIQ

Topics	**Coding**
School prayer	Normal coding
Pro-choice (abortion)	Reverse coding
National healthcare system	Reverse coding
Death penalty for murder	Normal coding

Table 4.2 Agree-or-Disagree Topics in the PIQ

Topic	**Coding**
Government regulations are needed to control monopolies.	Reverse coding
If the rich continue to get richer and the poor get poorer, I would support a violent revolution to correct the inequality.	Reverse coding
The traditional family (married father and mother with children) must be preserved at all costs.	Normal coding
Helping the poor encourages laziness.	Normal coding

It is important to note, again, that high scores on the PIQ do not necessarily mean that individuals are explicitly aligned with any one political party (e.g.,Republican Party in the United States or the Conservative Party of Canada); they denote that these individuals possess more traditional, rigid ideals regarding sociopolitical and economic topics. Conversely, low

scores on the PIQ do not align individuals with a political party; they indicate that these individuals possess less traditional and more open-minded ideals about the same topics.

Early psycholinguistic evidence implicated political ideology in the processing of sociocultural language. In other words, psycholinguists have identified that political ideology predicts, to a certain extent, comprehenders' inferential processing of socially charged language. A study by Van Berkum et al. (2009) was the first of its kind to implicate political and moral views during language processing. In this study, participants were asked to read statements such as *Euthanisia is an acceptable* (versus *unacceptable*) *course of action* and rate their agreement. Regardless of the participants' explicit alignment with political parties, their political ideologies (i.e., ratings of sentences such as the one above) predicted their ERP signatures for statements that were incongruent with their politically associated beliefs. In fact, their results showed that value-based disagreement was salient extremely rapidly during processing: ERP signatures were detected within 200 to 250 milliseconds after the word indicating a clash with a participant's value system (e.g.,*acceptable* or *unacceptable* in the sentence above). In other words, participants' political ideologies predicted the strength of their unconscious reactions to socially-charged sentences.

A further study by Marrville (2017) showed that participants' political ideologies significantly predicted their sentence completions for verb phrases depicting interpersonal events and relationships. For example, participants were asked to complete sentences like *Kaleigh* (noun one) *thanked Nathan* (noun two)*, because* ____. The verb *thanked* is low in dominance and high in valence, meaning that it is often associated with female subjects; verbs like *criticize* are high in dominance and low in valence, meaning that they are often associated with male subjects. Results from this study showed that more conservative participants had more noun one

continuations for verbs associated with female subjects and more noun two continuations for verbs associated with male subjects; fewer conservative participants showed the reversed pattern. These results suggest that individuals' political ideologies modulate how they understand causal relationships in language-encoded events. Replication of this experiment by Niemi et al. (2020) led the authors to speculate that political ideologies that support the individual versus the group predict that individuals will place the results of a harmful action on the noun most associated with obedience to authority and loyalty.

In a similar vein, Puhacheuskaya and Järvikivi (2022) examined the comprehension of ironic comments spoken by individuals with native versus foreign accents. Interestingly, listeners' political ideologies significantly affected their ratings of severity of irony: more conservative participants detected irony less accurately overall, regardless of speaker accent; they also misinterpreted literal compliments as sarcasm. The authors speculate that less conservative individuals are more open-minded, meaning that they are more likely to detect playfulness and favor jocular interpretations over traditional judgements. Similarly, literal compliments are likely to be misinterpreted as sarcastic if speakers disfavor one another (Slugoski & Turnbull, 1988): the authors posit that more conservative individuals err on the side of caution during social interactions. These results indicate that political ideology significantly predicts the language comprehension of socially-constructed ironic comments, which is yet another piece of evidence suggesting that political ideology is involved in sociocultural language processing.

Little experimentation has been conducted to investigate the role of political ideology on gender stereotype processing. However, the recent work by Hammond-Thrasher and Järvikivi (2023) mentioned in Section 1.3.1 of this Thesis indicates that individual differences in political

ideology predict how strongly participants perceive gender stereotype violations. To avoid unnecessary restatements of fact, refer to the above section for an overview of the experiment. Results from this study indicated that less conservative participants rated sentences containing violations of gender stereotypes as less Acceptable and Correct overall. Furthermore, these same participants rated sentences describing explicitly male agents fulfilling stereotypically feminine roles as less Acceptable and Correct, yet again displaying the character gender-specific Stereotype mentioned above. Here, political ideology predicted individuals' sensitivity to gender stereotype violations. Importantly, this does not mean that less conservative individuals view gender stereotype clashes as unacceptable or incorrect: the results indicate that these individuals perceive gender stereotype violations with more salience, suggesting that they rely strongly on gender stereotype knowledge during inferential language comprehension. This effect is yet to be replicated but poses important implications for the processing of gender stereotyped language.

1.3.3 Gender Identity

Gender identity is an individual difference factor that must be differentiated from biological sex assigned at birth. While the latter refers to the typical sex assigned (i.e., female, male, or intersex) based on an individual's reproductive system and genetic factors, the former is a more nuanced category. Gender identity encapsulates an individual's internal and external experience of gender: this may align with sex assignment at birth but often differs (Ontario Human Rights Commission, 2023). In other words, gender identity is an individual's personal sense of being somewhere along the gender spectrum, which is typically depicted as varying from cisgender male to cisgender female with an infinite number of identities between. Díaz-Andreu (2005) explains the experience of gender as follows:

> "Because gender dynamics are included in all activities, even the most routine ones, gender constitutes an intimate element of the process of society. Gender is a basic structuring principle, for it frames the primary parameters that guide our understanding of the world and creates the rules that serve as a basis for individual behavior."

Gender is a vital component of an individual's understanding and expression of their identity. The degree to which individuals consciously rely on their gender identity as a lens for their interactions with the real world varies: for some individuals, gender is a known, primary component of their sense of self, while others may view gender as less important factor in their interpersonal relationships and everyday behaviors (Díaz-Andreu, 2005). For this reason, there are individual differences in how individuals identify their genders (i.e., the extent to which they are viewed as binary versus nuanced) and in the degree to which these identities are paramount to an individual's identity.

Furthermore, there are individual differences in gender identity that are less conscious. Social Role and Dominance Theories posit that gender identity plays a significant yet often unconscious role in everyday life: as summarized in Section 1.2 of this Thesis, gender stereotypes exist and are particularly salient because of the socially-constructed performance of gender and the institutional mechanisms, such as gender stereotyped language, that perpetuate these divisions (Eagly, 1987, and Sidanius & Pratto, 1999, respectively). Human ecology scholars posit that this framework, along with historical gender partitioning, is the precise reason for the persistent social debate regarding gender (see Peterson & Hyde, 2014, for an overview). Evidently, variation in gender identity is becoming increasingly visible in Western societies

(Statistics Canada, 2021). However, the acceptance of nuanced gender identities is still controversial, with both violent and nonviolent hate-crimes towards transgender and gender non-conforming individuals a prevalent, ongoing, and common tragedy (Human Rights Campaign, 2022). Hence, it is vital that social science research adopts a non-dichotomous approach to gender: non-refutable gender diversity must be visible and embedded in institutional frameworks, and methodology that does recognize existing gender identities fails to reliably account for individual differences.

I approach this work with the acknowledgement of my academic responsibility to participate in diversifying gender-based research in an effort to produce valid, reliable, and representative results. Hence, I opt for a non-binary measurement of gender identity by employing the Genderqueer Identity Scale (GQI; McGuire et al., 2018), which contains twenty-three statements rated on a five-point Likert scale. Sample statements are presented in Table 5 and are either normally- or reverse- coded, with high scores indicating less traditional gender identities.

Table 5. Sample Statements from the GQI

Statement	**Coding**
I don't want to be seen in the gender binary (as either male or female).	Normal coding
The way I show my gender will probably be mostly the same from day to day.	Reverse coding
The way I show my gender is important because I push society to question traditional gender roles.	Normal coding
In the future, I expect that people will rarely question my gender.	Reverse coding

The GQI assessment is praised for its ability to challenge the binary and social constructions while providing a reliable and valid alternative to asking for explicit gender labeling (Coco, 2021). Its use in psychological research is extensive (see Kallitsounaki & Williams, 2022; de Graaf et al., 2021; Bradford & Catalpa, 2019), making it a prime candidate for the present research.

Traditional studies have viewed gender through a purely binary lens, measuring gender as a purely female-male dichotomy. However, recent evidence suggests that components of gender identity predict, to a degree, the processing of sociocultural language. An EEG study by Grant et al. (2019) presented participants with sentences describing feminine (e.g.,fashion) and masculine (e.g.,sports) topics spoken by a male and a female speaker. Half of the participants heard stereotype-congruent sentences (i.e., for the male speaker, semantic errors about fashion and no errors about sports) while the other half heard stereotype-incongruent sentences (i.e., for the male speaker, semantic errors about sports and no errors about fashion). Results found a larger N400 ERP signature in participants listening to stereotype-incongruent sentences compared to congruent ones. The authors also examined ERP changes across the course of the experiment: the degree to which these ERP signatures lessened over time was predicted by individual differences in sexism. In other words, participants who scored higher in sexism had less dramatic decreases in their neurological responses to stereotype violations, while participants low in sexism showed more dramatic decreases in detecting stereotype violations as they became more familiar with the speakers' voices. Hence, the degree to which participants endorsed traditional gender norms predicted their sensitivity to gender stereotype violations.

Another EEG experiment indicates that individual differences in gender identity modulate gender stereotype processing. Canal et al. (2015) recorded participants' ERPs while

they read sentences containing gender stereotyped role nouns that either matched or did not match explicit gendered pronouns. As expected, results showed strong ERP signatures in response to stereotype incongruent sentences. Interestingly, participants were asked to complete the Bem Sex-Role Inventory (BSRI; Bem, 1974), which asks individuals to rate self-descriptive traits associated with femininity (e.g.,"is affectionate") and masculinity (e.g.,"acts as a leader"). Results indicated that participants with more feminine self-descriptions showed a P600 signature for stereotype violations, which indicates that they experienced the incongruency as a syntactic agreement violation; less feminine participants showed a semantic processing effect, indicating more effort spent linking the pronouns with the stereotypical meaning of the role name. Here, participants' descriptions of their own sex roles predicted the specific type of difficulty they experienced when processing gender stereotype violations.

Overall, there is limited research investigating the role of individual differences in gender identity during sociocultural language processing, yet compelling evidence points to its ability to predict the comprehension of gendered statements. Despite the lack of consensus identifying a single assessment tool for gender identity in such studies, there presents a need for research that endeavors to connect self-descriptions of gender identity to these sociocultural language phenomena. Hence, this research assumes the task of linking GQI scores to gender stereotype processing.

Chapter 2

2. Current Study

In Chapter 1 of this Thesis, I presented experimental evidence outlining the processing of gender stereotyped language and its modulation by certain individual differences. While it is evident that gender stereotype violations are uniquely salient and that individual differences can, to a degree, predict the processing of this phenomenon, some important information about gender stereotype processing has yet to be investigated. Firstly, little is known about how gender stereotype violations are processed when embedded in discourse (Garnham et al., 2012). That is, are gender stereotypes salient across multiple sentences such that the Stereotype Effect (i.e., the increased salience of gender stereotype violations during language comprehension) arises in a discourse context? Secondly, gender stereotype violations at the adjective-level have yet to be investigated, especially when paired with gender stereotyped occupational roles. The question of whether adjectival gender stereotype violations are processed similarly to noun-based violations must be investigated because of the overwhelming evidence that gender stereotypes are embedded in broad linguistic contexts. That is, language comprehension literature must investigate whether gender stereotypes embedded in multiple linguistic constructions influence language comprehension because these stereotypes are located across many linguistic forms, such as both adjectives and nouns. Thirdly, the limited evidence of the character gender-specific Stereotype Effect (namely that sentences describing male characters fulfilling stereotypically feminine roles elicit a larger Stereotype Effect) necessitates replication in order to provide firm evidence of its existence. Does this Effect also exist for adjective-level gender stereotype violations? Finally, there is minimal evidence supporting the predictability of gender stereotype processing by

individual differences in personality, political ideology, and gender identity. More research must be performed to outline these effects. This study investigates the processing of noun- and adjective- level gender stereotype violations within discourse. Pairing individual difference methodologies with a reading paradigm, I collect reading time and yes-no comprehension question data in an effort to answer these specific questions. Based on previous literature, I aim to answer the following questions:

(1) Is the Stereotype Effect replicable in a multi-sentence (i.e., discourse) context? If so, does a character gender-specific Effect arise similar to that found in previous studies?
(2) Do violations of gender stereotyped descriptive adjective expectations elicit a similar Stereotype Effect to that of occupational role names?
(3) Do individual differences in personality, political ideology, and gender identity predict the comprehension of adjective- and noun- based gender stereotyped constructions?

2.1 Methodology

2.1.1 Participants

Ethics for this study were approved by the Research Ethics Office at the University of Alberta (Pro00122792). 211 participants were recruited from the Prolific participant pool, all of whom consented to participating in the study. All were native speakers of English, were born and residing in Canada or the United States of America, and had normal or corrected-to-normal vision. Every participant received adequate financial compensation for their participation as outlined by the Prolific user guidelines. Participant demographics are available in Appendix A.

2.1.2 Materials

A set of thirty-five female- and thirty-five male-stereotyped role names were gathered from Marrville's (2017) study. A second set of twenty-five female- and twenty-five

female-stereotyped descriptive adjectives were collected from Castillo-Mayén and Montes-Berges' (2014) study. In addition, thirty-five neural (i.e., non-gender stereotyped) role names and twenty-five adjectives were selected. All vocabulary items underwent preliminary ratings by fifty-five undergraduate students recruited through the SONA participant pool at the University of Alberta. Participants were presented with the entire vocabulary lists and asked to rate how feminine versus masculine they perceived the terms on a five-point Likert scale, with the extremes referring to extremely feminine or masculine, respectively, and the midpoints referring to neutrality. As expected, the rating results provided highly gender stereotyped items for use in this study (Appendix B). Specifically, the results yielded a total of twenty female-, male-, and neutral- stereotyped role names and the same number of adjectives, respectively.

There was a total of forty experimental and twenty filler short stories (Appendix C). These vignettes (i.e., small samples of discourse) were constructed according to the structure in Table 6.

Table 6. Vignette Construction Structure

Sentence One: Set-Up	**Sentence Two: Gender Stereotype Presentation**	**Sentence Three: Target Segment**
You are VERBing at LOCATION.	You see the ADJECTIVE ROLE-NOUN VERB the NOUN.	You watch PRONOUN VERB the NOUN.

Every vignette was constructed following the same paradigm: the first sentence began with a first-person perspective set-up that introduced the setting of the short story. The following sentence presented the experimental items: a descriptive adjective (either masculine or feminine

in stereotype gender) was presented in conjunction with a role name (either masculine or feminine in stereotype gender) to identify the protagonist in the discourse. The third sentence contained the target segment (i.e., the pronoun absolving the ambiguity of the protagonist's explicit gender) and closed off the end of the discourse segment.

Because this study follows a 2 x 2 x 2 design (i.e., there are two genders associated with the adjective, two genders associated with the role noun, and two genders associated with the pronoun), there were a total of eight experimental conditions. Table 7 presents samples of each condition, where the term 'Congruence' refers to whether the gender of other elements in the discourse match that of the gender stereotyped role name.

Table 7. Sample Vignettes for Each Experimental Condition

Condition	**Sentence One**	**Sentence Two**	**Sentence Three**
Female Congruent (no violation)	You are sitting in the waiting room in the clinic.	You see the soft dental hygienist look at the clock.	You watch her pick up the water pick.
Male Congruent (no violation)	You are driving a car down the highway.	You see the buff truck driver turn into the next lane.	You watch him check for traffic in the side mirror.
Female Incongruent (double violation)	You are walking by the daycare in your town.	You see the tough childcare worker play with the crayons.	You watch him draw on a piece of paper.
Male Incongruent (double violation)	You are carrying a box down the road.	You see the gentle construction worker climb up the scaffolding.	You watch her hammer in a long nail.
Female Congruent with Incongruent Adjective (adjective-level violation)	You are passing by the school down the street.	You see the forceful social worker pick up a teddy bear.	You watch her walk towards the big class.
Male Congruent	You are eating at the	You see the classy	You watch him nod

with Incongruent Adjective (adjective-level violation)	restaurant in the office building.	stockbroker speak into a phone.	at a nearby colleague.
Female Incongruent with Congruent Adjective (noun-level violation)	You are walking through the mall in your city.	You see the flexible social media influencer look at the flowers.	You watch him pose for a quick selfie.
Male Incongruent with Congruent Adjective (noun-level violation)	You are entering the airplane at the airport.	You see the robust pilot pick up the microphone.	You watch her point down the seating row.

Filler sentences followed the same structure but were neutral in their gender presentation to avoid any effects of gender stereotyping. A sample filler item is presented in Table 8.

Table 8. Sample Filler Item

Sentence One	**Sentence Two**	**Sentence Three**
You are watching the street outside your office.	You see the accurate observer read a street sign.	You watch them walk down the long street.

In addition to the experimental materials, individual differences questionnaires were incorporated into this study. Specifically, the HEXACO PI-R 60-Item version (Appendix D), the PIQ (Appendix E), and the GQI (Appendix F) were presented to participants in their original forms with no alterations.

2.1.3 Procedure

The experiment was designed using PsychoPy version 2022.02.04 (Peirce et al., 2019). A user-friendly approach was taken, incorporating the ethics information, consent questions, experimental trials, and individual differences questionnaires into the same interface. The PsychoPy experiment was uploaded to the Prolific platform using Pavlovia as the online host for its completion. Participants completed the study in a quiet location on either a laptop or desktop computer.

Experimental items were compiled and cross-balanced across four lists, each of which was completed by an equal number of participants; items were presented in a randomized order. Participants were presented with specific instructions asking them to read each sentence of every short story carefully. Each of the three sentences was presented on its own, with participants pressing the SPACE key on their keyboards to view the next sentence; once viewed, participants were unable to re-read previous sentences. Third-sentence reading times were collected for all items. After reading the full vignette, participants were asked to answer whether they found the final sentence to be a reasonable completion of the entire short story by pressing the X key to indicate "Yes" and the M key to indicate "No" following a similar paradigm to Garnham et al. (2012). Both reading time data for each sentence and these responses to the sensibility checks were recorded. Every ten vignettes, participants were asked to indicate whether they were still paying attention to the study with a similar yes-or-no response. It is important to note that each item was presented following a two-second fixation cross centered on the screen; all items matched in font size and positioning on screen.

It is vital to note that binary sensibility confirmations directly measure participants' meta-linguistic judgements of vignette sensibility (i.e., conscious decisions about the relationship

between vignette contents and participants' beliefs of a story's ability to make sense; Tjuka et al., 2022). Contrastively, third-sentence reading times measure both the time taken to read the given sentence and the active decision to proceed to the sensibility check question (Racine, 2014). Here, it is impossible to claim that these reading times only pertain to solely the time taken to read a sentence: this data also encodes any decisional hesitation and attentional difficulties. Hence, reading time data may be indicative of the process of language comprehension, but cannot be taken at face value to solely encode reading time itself. For this reason, binary sensibility check responses are the primary data source for this study.

Upon completion of the experimental trial, participants were asked to complete the HEXACO PI-R, PIQ, and GQI, respectively. Each question on each assessment was presented individually on screen with a sliding scale with fixed points for ratings across five-point Likert scales. Items on these assessments were presented in the order in which the authors designed. Importantly, pre-screening questions were incorporated into the Prolific platform to ensure the participants' status as native speakers of English, record their birthplace and current residence, and block participation from those who previously participated.

2.2 Results

Before data analysis, filler items were removed. Analysis began with sensibility check responses (i.e., yes-or-no answers to whether the last sentence in a vignette was a sensible continuation of the short story); this was motivated by previous evidence from Hammond-Thrasher and Järvikivi (2023), where offline data supported the gender Stereotype Effect. For this reason, I considered sensibility check responses to be the primary subject of analysis for this study. Following this, reading time data (i.e., the time taken to read the third sentence, which contained the pronoun resolution, in each vignette) was analyzed following similar analyses to those used on the

sensibility check responses. Finally, I examined the effects of individual differences measures on the predictability of these results.

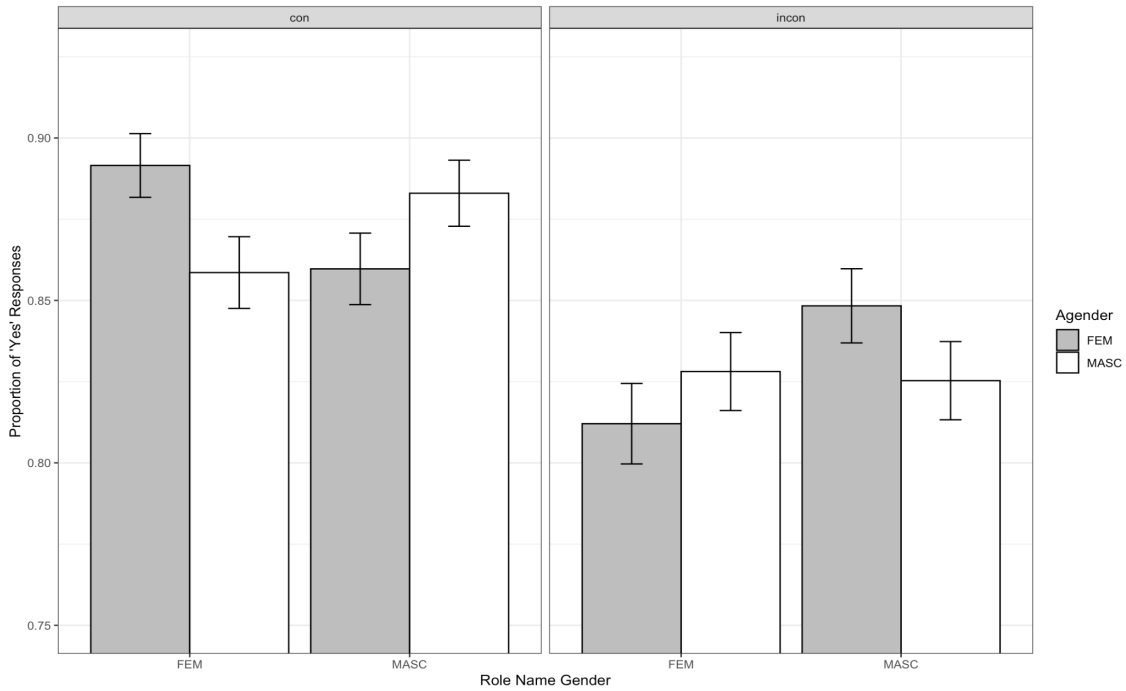
2.2.1 Sensibility Check Analysis

The results were analyzed using generalized additive mixed-effects modeling (GAMMs) in the R statistical environment (Wood, 2017; R Core Team, 2020). This paradigm was selected to ensure reliable testing of manipulated variables alongside the effects of individual differences without assuming linearity. Sensibility check responses, which were binomial in nature (i.e., 'Yes' or 'No'), were the dependent variable in these analyses. Analysis began by examining the fixed effects variables of interest as follows:

(1) Congruence (whether the stereotyped gender of the role was Congruent or Incongruent with the gender of the explicit pronoun);
(2) R Gender (whether the stereotyped gender of the role name was female or male); and
(3) A Gender (whether the stereotyped gender of the adjective was female or male).

By-subject and by-item factor smooths for Condition were added as far as the models converged using a backward-fitting procedure and comparing model fits (compareML() in package *itsadug*; van Rij et al., 2022). Pairwise comparisons were run using the function pairs() in package *emmeans* (R Core Team, 1996). The final model showed several significant effects (Figure 1).

Figure 1. Proportion of "Yes" Responses by Congruence (Congruent in Panel One, Incongruent in Panel Two), R Gender (Left and Right Bar Groupings for Feminine and Masculine Genders, respectively, in Panel One and Two), and A Gender (Dark Grey and White for Feminine and Masculine Genders, respectively, in Panel One and Two).



The left-most panel depicts the proportion of "Yes" key press responses for all variations of Congruent items, while the right-most panel depicts the proportion of "Yes" key press responses for all variations of Incongruent items. Statistical outputs are shown in Table 8. The rows provide parametric coefficients for Congruence, R Gender, and A Gender, with Congruent Female items as the reference level.

Table 9. Summary of Effects for Key Press Responses

	Estimate	**Standard Error**	**z-value**	**p-value**
(Intercept)	2.8344	0.2971	9.539	<2e-16 ***

CongruenceIncongruent	-0.7519	0.1471	-5.113	3.17e-07 ***
RGenderMASC	-0.4163	0.3328	-1.251	0.2110
AGenderMASC	-0.3453	0.1523	-2.268	0.02334 *
CongruenceIncongruent: RGenderMASC	0.6359	0.2032	3.129	0.00175 **
CongruenceIncongruent: AGenderMASC	0.4783	0.2029	2.357	0.01841 *
RGenderMASC: AGenderMASC	0.5909	0.2121	2.786	0.00534 **
CongruenceIncongruent: RGenderMASC: AGenderMASC	-0.9269	0.2855	-3.246	0.00117 **

Significance codes: 0 ‘***’ .001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘’ 1
Formula: keys ~ congruence * Rgender * Agender + s(Rname, bs = "re") +
s(pp, bs = "re")
Note: RName refers to Item and PP refers to Participant.

Firstly, there was a significant effect of Congruence: Congruent items were consistently judged to be more sensible than Incongruent items ($p = 8.33e-07$). In other words, there were significantly more “Yes” responses for Congruent items compared to Incongruent ones. Secondly, there was a significant two-way interaction between Congruence and R Gender. Pairwise comparisons showed that there were significantly more “Yes” responses for Incongruent Male items compared to Incongruent Female ones ($p = 0.0018$); the difference in “Yes” responses was greater for Congruent vs. Incongruent Female items compared to that of Congruent vs. Incongruent Male items ($p < 0.001$). This is in line with previous research indicating that sentences describing male agents fulfilling stereotypically feminine roles are judged to be relatively less sensible than those describing female agents fulfilling stereotypically masculine roles (Hammond-Thrasher & Järvikivi, 2023).

Finally, there was a significant three-way interaction between Congruence, R Gender, and A Gender. Here, items combining feminine adjectives with feminine R Genders and male pronouns (e.g., *flexible* + *social media influencer* + *he*) were responded to with the least number of "Yes" responses; contrastively, items combining masculine adjectives with masculine role names and female pronouns (e.g., *robust* + *pilot* + *she*) were responded to with significantly more "Yes" responses than the former ($p = 0.00117$). This result shows that an A Gender - R Gender coupling where both components are Incongruent with the pronoun results in the largest Stereotype Effect of all conditions, and that this is a character-gender specific effect aligned with previous results (Hammond-Thrasher & Järvikivi, 2023). In other words, two Incongruent items (i.e., both an adjective and a noun phrase) facilitate an amplified stereotype-based representation of a protagonist within discourse that then undergoes a significant revision when a clashing pronoun is encountered. Further, A Genders that were Incongruent with a statement's R Gender (i.e., pairing a male R Gender with a feminine A Gender (e.g., *classy* + *stockbroker* + *he*) or pairing a female R Gender with a masculine A Gender (e.g., *forceful* + *social worker* + *she*)) received significantly more "Yes" responses when coupled with pronouns Congruent with the same statement's R Gender ($p = 0.00117$); these items received a significantly lower proportion of "Yes" responses than fully Congruent items but this proportion was still significantly higher than other Incongruent items. This result shows that a clashing A Gender does interrupt stereotype-based integration during language comprehension; however, simultaneous Congruence between the R Gender and gendered pronoun still facilitates language comprehension.

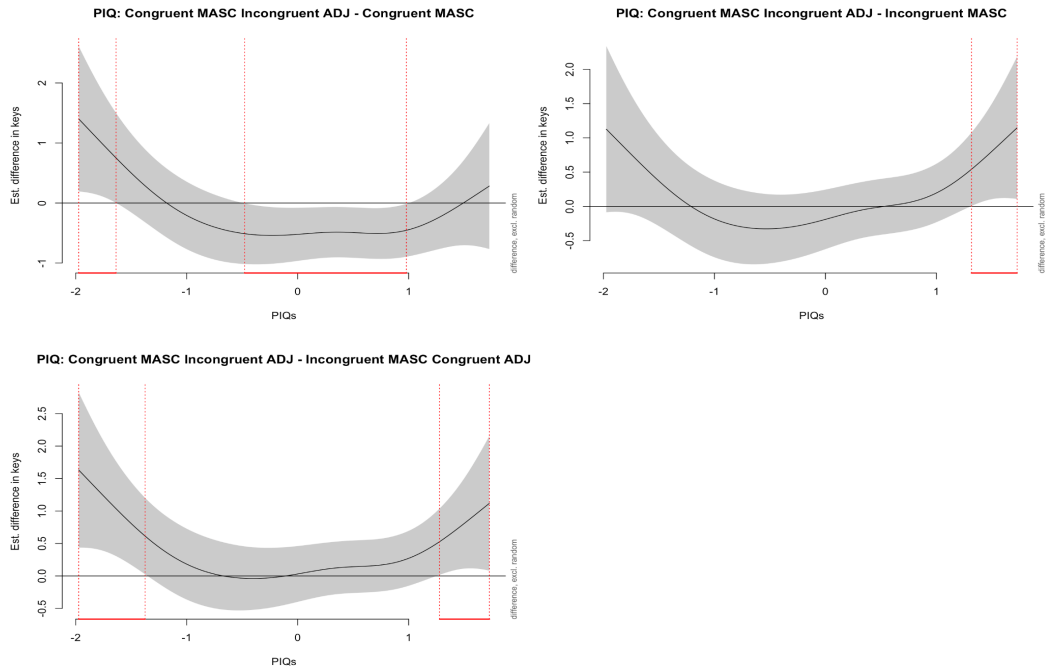
2.2.1.1 Individual Differences Effects on Key Press Responses

In order to inspect the effects of the individual difference measures on the key press responses while containing the three-way interaction mentioned above, I first formed an eight-level predictor variable, Condition, by merging Congruence, R Gender, and A Gender; levels are detailed in the Methods section of this thesis. This was necessary because GAMMs allow the modeling of behavioral data, such as responses to stimuli and individual difference measures, in a nonlinear fashion (Lin & Zhang, 1999). All individual difference predictors were tested separately and only one model was fitted for each to avoid inflation of the chance of false positives. Again, GAMMs were employed with the two-level (Yes/No) key press as the dependent variable and Condition as a parametric fixed predictor. Each model included a smooth term for the interaction between Condition and the given individual difference predictor (i.e., each of the seven HEXACO PI-R traits along with PIQ and GQI scores), random smooths for participant and item, as well as by-subject factor smooths for Condition. No significant effects were found for any HEXACO PI-R trait nor for GQI scores.

There was a significant effect indicating that PIQ scores predicted key press responses for one level of Condition ($p = 0.047$); the corresponding statistical model can be found in Appendix G, where the first rows provide parametric coefficients for Condition with Congruent Female as the reference level. The effective degrees of freedom (edf) indicate the degree of linearity of each regression line, where values greater than one represent non-linearity. The p-values indicate whether the partial effects, in effect lines, are significantly different from zero for any value on the x-axis (i.e., PIQ scores). In order to interpret the shape of the regression lines, visual inspection of plots is necessary. Figure 2 illustrates the differences in “Yes” key press responses between different levels of Condition and PIQ scores (i.e., the differences between the regression

lines for each pair as labeled). For each plot, gray shading marks 99% confidence intervals and the red dotted lines indicate points on the x-axis between which the regression lines are significantly different from one another. Throughout these plots, increasing values on the x-axis denote higher degrees of political Conservatism (labeled PIQ). Importantly, only those plots containing significant differences are depicted.

Figure 2. Difference Plots Depicting the Interaction Between PIQ Scores (X-Axis) and Condition (Panel Title) and Their Effects on Proportion of "Yes" Responses (Y-Axis).



Note: Increasing scores on the x-axis represent higher degrees of Conservatism (vs. Liberalism)

The upper left panel shows the interaction between PIQ scores and the proportion of "Yes" responses for Congruent Male vignettes containing Incongruent A Genders versus Congruent Male items (i.e., those pairing feminine adjectives, masculine role names, and masculine pronouns versus those pairing masculine adjectives, role names, and pronouns, respectively).

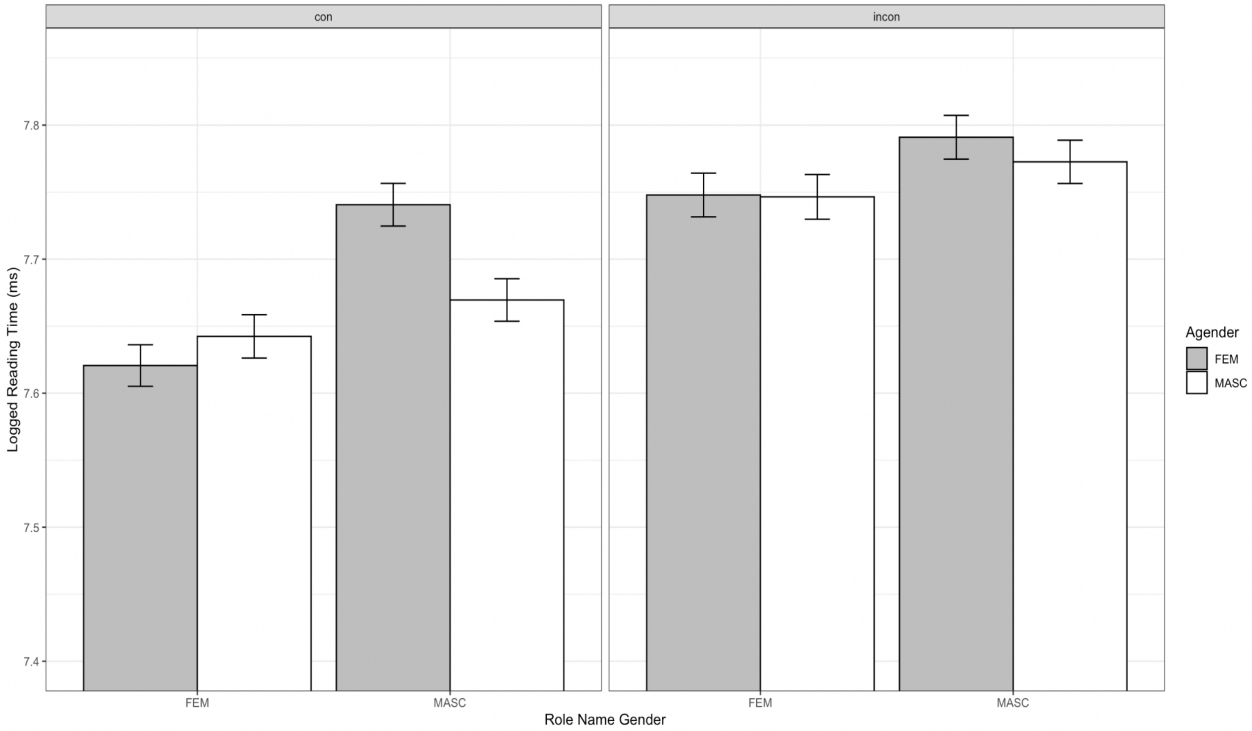
Here, participants who were lower in Conservatism (i.e., more Liberal) gave more sensibility confirmations to Congruent Male vignettes containing Incongruent A Genders than their more Conservative counterparts. The upper right panel shows the interaction between PIQ scores and the proportional of "Yes" responses for Congruent Male vignettes containing Incongruent A Genders versus Incongruent Male items (i.e., those pairing feminine adjectives, masculine role names, and masculine pronouns versus those pairing feminine adjectives, masculine role names, and feminine pronouns, respectively). Here, highly Conservative participants gave more sensibility confirmations for Congruent Male vignettes containing Incongruent A Genders compared to their less Conservative counterparts. The lower left panel shows the interaction between PIQ scores and the proportion of "Yes" responses for Congruent Male vignettes containing Incongruent A Genders versus Incongruent Male items containing Congruent A Genders (i.e., those pairing feminine adjectives, masculine role names, and masculine pronouns versus those pairing masculine adjectives and role names with feminine pronouns, respectively). Here, less Conservative (i.e., more Liberal) participants gave more sensibility confirmations to Congruent Male vignettes containing Incongruent A Genders than their more Conservative counterparts.

2.2.2 Sentence Reading Time Analysis

All participant data was cleaned such that reading time data exceeding two standard deviations from the mean by-participant were removed. The results were again analyzed using GAMMs following the same paradigm as above and including the same fixed effects variables and by-subject and by-item random intercepts as well as slopes for item and subject by CxR. In addition, all models included a smooth for previous (second) sentence log reading time (in order

to counter possible effects of autocorrelation). The final model showed several significant effects (Figure 3).

Figure 3. Sentence Three Reading Times (logged, ms) by Congruence (Congruent in Panel One, Incongruent in Panel Two), R Gender (Left and Right Bar Groupings for Feminine and Masculine Genders, respectively, in Panel One and Two), and A Gender (Dark Grey and White for Feminine and Masculine Genders, respectively, in Panel One and Two).



Again, the left-most plot depicts the log-transformed reading times (ms) for all variations of Congruent items, while the right-most plot depicts log-transformed reading times (ms) for all variations of Incongruent items. Statistical outputs are shown in Table 9. The rows provide parametric coefficients for Congruence, R Gender, and A Gender, with Congruent Female items as the reference level.

Table 10. Summary of Effects for Sentence Reading Times

	Estimate	**Standard Error**	**z-value**	**p-value**
(Intercept)	7.6466	0.0372	205.314	<2e-16 ***
CongruenceIncongruent	0.1164	0.0122	9.565	<2e-16 ***
RGenderMASC	0.07978	0.0440	1.813	0.0699
AGenderMASC	0.002476	0.0120	0.207	0.836
CongruenceIncongruent: RGenderMASC	-0.03616	0.0169	-2.139	0.0325 *
RGenderMASC: AGenderMASC	-0.04432	0.0169	-2.618	0.00885 **

Significance codes: 0 '***' .001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Formula: s3rt ~ congruence * Rgender + Rgender * Agender +
s(pp, by = congruence, bs = "fs") + s(log(s2rt * 1000)) +
s(Rname, bs = "re") + s(pp, bs = "re")
Note: RName refers to Item and PP refers to Participant.

Firstly, there was a significant effect of Congruence: Congruent items took significantly less time to read than Incongruent ones (p = 1.54e-14). Secondly, there was a significant two-way interaction between Congruence and R Gender: Incongruent Male items took significantly more time to read than Incongruent Female ones (p = 0.00538). In other words, items describing masculine roles being fulfilled by female agents took longer times to read than those describing feminine roles being fulfilled by male agents. This effect appears to be contradictory to the two-way interaction between Congruence and R Gender found in the key press analysis above, where Incongruent Female items received significantly fewer "Yes" responses than Incongruent Male items; however, it must be understood that key press responses gather information about metalinguistic judgements of sensibility while reading times gather information about the unconscious time taken to comprehend the same statement. In other words,

key presses and reading times do not measure the same processes and hence cannot be expected to show identical effects (Hammond-Thrasher & Järvikivi, 2023; Redl et al., 2021). The implications of this difference in measurement will be discussed in detail shortly. Importantly, the difference in reading times between Congruent versus Incongruent Female items was more significantly pronounced than that of Congruent versus Incongruent Male items ($p < 0.001$).

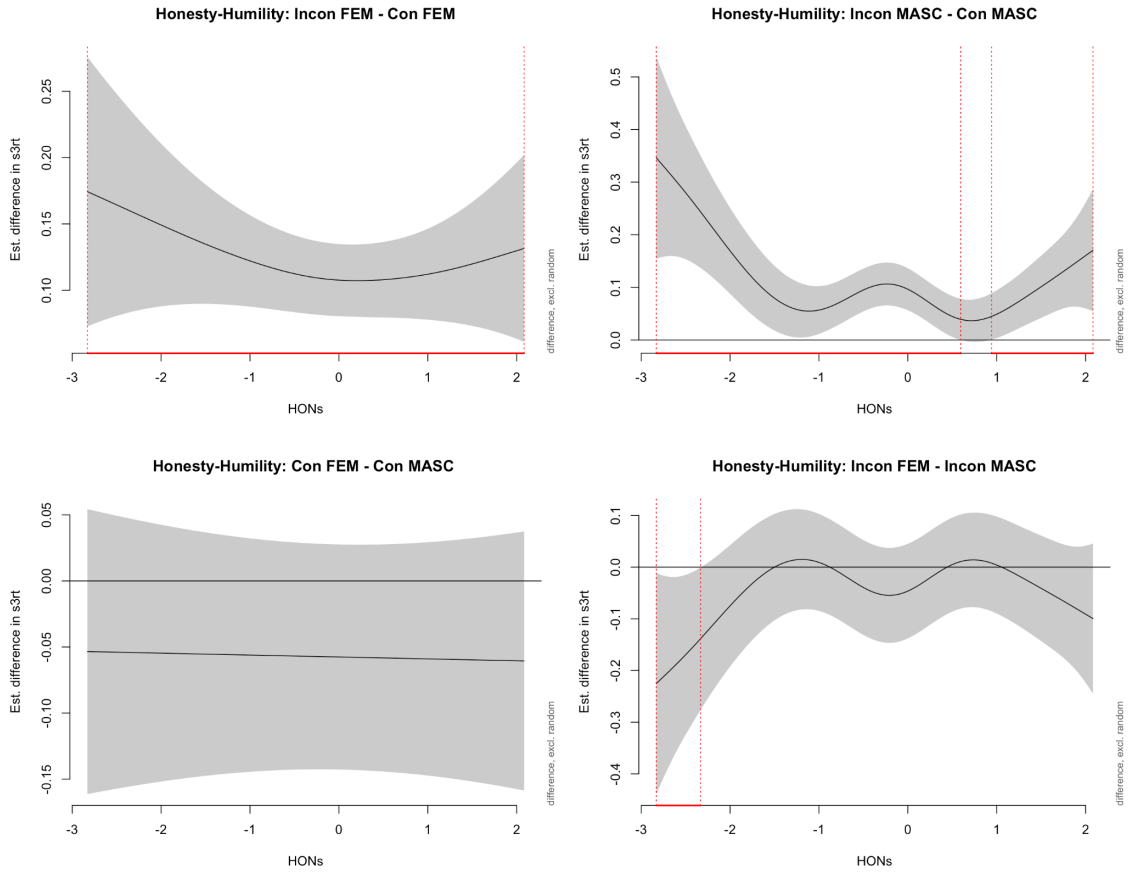
Finally, there was a significant two-way interaction between R Gender and A Gender: items pairing feminine A Genders with masculine R Genders predicted significantly longer reading times compared to those pairing both masculine A Genders and R Genders ($p = 0.0018$). This interaction contributed to the two-way interaction between R Gender and Congruence. Unlike the primary analysis of sensibility check response data, there was no significant three-way interaction between R Gender, A Gender, and Congruence in this model; however, the interaction between R Gender and A Gender indicates that A Gender did significantly contribute to reading times. Additionally, given that these reading time data were by-sentence rather than by-segment (i.e., less precise than word-by-word reaction times), it was expected that reading time data would be less indicative of gender stereotype effects during this analysis (Kennison & Trofe, 2003; Garnham et al., 2002). Importantly, third-sentence reading times measure both the time taken to read the given sentence and the active decision to proceed to the sensibility check question (Racine, 2014). Here, it is impossible to claim that these reading times only pertain to solely the time taken to read a sentence: this data also encodes any decisional hesitation and attentional difficulties. Hence, reading time data may be indicative of the process of language comprehension, but cannot be taken at face value to solely encode reading time itself.

2.2.2.1 Individual Differences Effects on Sentence Reading Times

In order to inspect the effects of the individual difference measures on the sentence reading times, I first formed a four-level predictor, CxR, by merging Congruence and R Gender due to the strongly significant two-way interaction between the two variables. All individual difference predictors were tested separately to avoid the inflation of the chance of false positives. Again, GAMMs were employed with sentence reading time as the dependent variable and CxR as a parametric fixed predictor. Each model included a smooth term for the interaction between CxR and the given individual difference predictor (i.e., each of the seven HEXACO PI-R traits along with PIQ and GQI scores), a smooth for second sentence log reading time, random smooths for participant and item, as well as by-subject factor smooths for CxR. There was a single significant effect indicating that Honesty-Humility scores predicted sentence reading times based on the identity of CxR ($p < 0.001$) which can be found in Appendix H; no other significant effects were found.

Figure 4 illustrates the differences in third-sentence reading times (labeled s3rt) between CxR and Honesty-Humility scores (i.e., the differences between the regression lines for each pair of CxR levels). Again, for each plot, gray shading marks 99% confidence internals and the red dotted lines indicate points on the x-axis between which the regression lines are significantly different from one another. Across all plots, increasing values on the x-axis represent a stronger presence of Honesty-Humility traits (labeled HONs).

Figure 4. Difference Plots Depicting the Interaction Between Honesty-Humility Scores (X-Axis) and CxR (Panel Title) and Their Effects on Sentence Reading Times (logged, ms; Y-Axis).



Note: Increasing scores on the x-axis represent higher degrees of Honesty-Humility

The upper left plot shows the interaction between Honesty-Humility scores and reading times for Incongruent versus Congruent Female items (i.e., those pairing masculine A Genders and pronouns with feminine R Genders versus those pairing all three feminine components, respectively). Here, participants with higher Honesty-Humility scores showed a significant decrease in their time taken to read Incongruent Female items compared to their less Honest-Humble counterparts; this is indicated by the decreased positive values on the x-axis for Incongruent versus Congruent Female items. The upper right plot shows the interaction between Honesty-Humility scores and reading times for Congruent versus Incongruent Male items (i.e.,

those pairing feminine A Genders and pronouns with masculine R Genders versus those pairing all three masculine components, respectively). Here, participants with higher Honesty-Humility scores showed a significant decrease in their time taken to read Incongruent Male items compared to their less Honest-Humble counterparts. The lower left plot shows the interaction between Honesty-Humility scores and reading times for Congruent Female versus Male items. Here, there was no significant difference found. Finally, the lower right plot shows the interaction between Honesty-Humility scores and reading times for Incongruent Female versus Male items. Here, participants with lower Honesty-Humility scores showed a significant increase in their time taken to read Incongruent Male items compared to their more Honest-Humble counterparts; this is indicated by the increased negative difference depicted on the y-axis. Conversely, higher Honesty-Humility scores predicted less of a reading time difference between Incongruent Female versus Male items.

Chapter 3

3. Discussion

This thesis introduced the Stereotype Effect, which denotes the increase in cognitive effort needed to integrate violations of gender stereotyped expectations during language comprehension. This Effect has been shown to be character gender-specific: sentences describing male agents fulfilling stereotypically feminine roles have predicted larger Stereotype Effects than those describing female agents fulfilling stereotypically masculine roles. In the present research, participants read short stories describing characters using gender stereotyped adjective-noun pairings followed by pronoun resolutions that either aligned or violated these stereotyped expectations; participants' reading times and confirmations of story sensibility were recorded. The current study sought to replicate the Stereotype Effect in a discourse context and determine whether the character gender-specific Effect is indeed replicable across multiple sentences. Furthermore, the current study introduced a paradigm that not only included the traditionally studied occupational role names but also gender stereotyped descriptive adjectives. I asked whether these adjectives elicited a similar cognitive disruption to that of the Stereotype Effect as well as whether there was an interaction between this Effect and gender stereotyped adjective-noun pairings. Finally, the current study introduced individual differences measures following a similar design to previous studies in an effort to determine whether a discourse-based Stereotype Effect would be predicted, to an extent, by individual differences in personality traits, political ideology, and gender identity.

Above, I outlined results from the current study, namely those indicated by the collection of binary key-press responses and third-sentence reading times. Before discussing the

implications of these findings, I will provide an interim summary of the results detailed above. Firstly, the study replicated the Stereotype Effect in a multi-sentence (i.e., discourse) setting. Vignettes containing gender stereotype violations at both the adjective- and noun- level received fewer sensibility confirmations and longer reading times than those without any gender stereotype violations. Secondly, a character gender-specific Stereotype Effect was replicated. Short stories describing a male character fulfilling a stereotypically feminine role received fewer sensibility confirmations compared to vignettes describing a female character fulfilling a stereotypically masculine role. Interestingly, reading time data indicated that vignettes describing a female character fulfilling a stereotypically masculine role took longer to read; however, the difference in reading times between short stories containing feminine stereotype violations and those without violations of these same stereotypes was greater. Here, again, it must be understood that third-sentence reading times encode both the time taken to read a given sentence and the active decision to proceed to the next associated task (Racine, 2014). For this reason, binary sensibility check responses are the most reliable for measuring the strength and directionality of this replicated Stereotype Effect.

Of the experimental results, perhaps the most interesting is that binary key-press response data indicated a noteworthy three-way interaction between adjective-level gender stereotypes, noun-level gender stereotypes, and the presence of violations of these gender stereotyped expectations. Vignettes combining feminine adjectives and nouns with male pronouns and those combining masculine adjectives and nouns with female pronouns were responded to with the least number of sensibility confirmations, indicating that double gender stereotype violations (i.e., violations of aligned adjective-noun couplings) elicit the strongest Stereotype Effect. Importantly, this effect itself was character gender-specific: those vignettes combining feminine

adjectives and nouns with male pronouns were responded to with the least number of sensibility confirmations overall. If these same vignettes were constructed such that the gender stereotype of the adjective matched the explicit gender of the pronoun while maintaining the noun-level gender stereotype violation, the proportion of sensibility confirmations increased, indicating that, to an extent, adjective-level gender stereotype alignment facilitates language comprehension. Conversely, if vignettes contained gender stereotype alignment between the noun and the pronoun while violating this gender stereotyped expectation with an incongruent adjective, the proportion of sensibility confirmations was lower than that of fully congruent short stories yet higher than that of only adjective-level gender stereotype alignment. Here, noun-level gender stereotype alignment facilitated language comprehension more than adjective-level gender stereotype alignment. Notably, reading time analysis did not reproduce this three-way interaction.

Regarding results from the analysis of individual differences, binary sensibility confirmations indicated that political ideology predicts the language comprehension of gender stereotype violations for certain conditions. More Conservative participants were more sensitive than their Liberal counterparts to items pairing feminine adjectives, masculine role names, and masculine pronouns. While reading time data did not indicate these effects, they did indicate that Honesty-Humility predicts the language comprehension of gender stereotype violations. Highly Honest-Humble participants were hyposensitive to gender stereotype violations overall: these participants read vignettes containing both male and female gender stereotype violations significantly faster than their less Honest-Humble counterparts. Interestingly, less Honest-Humble participants were hypersensitive to male gender stereotype violations: these

participants took longer to read vignettes describing a female agent fulfilling a stereotypically masculine role than their more Honest-Humble counterparts.

3.1 Stereotype Effect Replication

The successful replication of the Stereotype Effect found through binary sensibility check analysis confirms several facts. Firstly, the Stereotype Effect is present during offline language processing: conscious decisions about the relationship between vignette contents and participants' beliefs about a story's ability to make sense were significantly predicted by the presence of gender stereotype violations, successfully verifying the offline replication of the Stereotype Effect (Hammond-Thrasher & Järvikivi, 2023; Garnham et al., 2012). This result reflects the finding that information that contradicts sociocultural world knowledge, specifically gender stereotyped expectations, is particularly noteworthy during language processing. This confirms the fact that gender stereotype violations are perceived saliently (Hoyle et al., 2019; Rosette et al., 2015; Gaucher et al., 2011; Diekman et al., 2004; Heilman et al., 2004) and that readers make inferences about upcoming information using both the content of linguistic input and its associated sociocultural meaning (Guerra et al., 2021; Marrville, 2017; Molinaro et al., 2016; Oakhill et al., 2005; Osterhout et al., 1997; Banaji & Hardin, 1996). Importantly, failure to provide sensitivity confirmations should not be interpreted as indicating that the sentences are not plausible or that participants believe the situations described were not suitable in the real world. Rather, fewer sensibility confirmations are consistently interpreted as indicating that participants were attentive to some element of contradictory language (i.e., that of gender stereotype violations) during language comprehension (Kennison & Trofe, 2003; Garnham et al., 2002). Secondly, these findings provide evidence for the presence of the Stereotype Effect within a multi-sentence (i.e., discourse) setting. While the Stereotype Effect has previously been

produced in a sentential context, the present study is among the first (see Garnham et al., 2012) to replicate it across sentences (i.e., in a setting where stereotype violations are confirmed in a sentence additional to the presentation of gender stereotype knowledge). This result confirms previous work asserting that gender stereotype information is incorporated into inferential processing immediately and automatically and that this information is incredibly important during language processing (Guerra et al., 2021; Molinaro et al., 2016; Osterhout et al., 1997): even when separated from pronoun resolutions, the gender stereotyped components within the short stories produced a significant Stereotype Effect. That is, gender stereotypes are salient throughout discourse and can predict a reader's ability to comprehend this discourse beyond the context of a single sentence.

Furthermore, binary sensibility check analysis in the present study indicated that the degree to which participants experienced the Stereotype Effect was predicted by the typical gender of the stereotype itself: stories describing male agents fulfilling stereotypically feminine roles correlated with lower proportions of sensibility confirmations than those describing female agents fulfilling stereotypically masculine roles. This result is in line with previous research (see Hammond-Thrasher & Järvikivi, 2023, and Grant et al., 2019), suggesting that male individuals are more strongly believed to perform according to sociocultural expectations of gender roles compared to women (Gygax et al., 2016). This character gender-specific Stereotype Effect can be explained by the overall increased stability of male gender roles over time as contrasted by the historical migration of women away from traditional female gender roles (Haines et al., 2016). Here, Social Role and Dominance Theories predict that the performance of masculine roles by male individuals would be more favored than the performance of feminine roles by female individuals (Eagly, 1997; Sidanius & Pratto, 1999). In line with this, a female pronoun following

a stereotypically male occupational role name would be perceived less saliently than a male pronoun following a stereotypically female occupational role name. It must be noted that another factor may contribute to this character gender-specific Stereotype Effect: research has shown that stereotypically masculine role names are preferred as subjects and easier to integrate during language comprehension compared to their female counterparts, which may perhaps mask the Stereotype Effect experienced when a woman is described as fulfilling a stereotypically masculine role (Redl et al., 2021, and Marrville, 2017).

3.2 Adjective- and Noun- Based Stereotype Processing

Previous knowledge of occupational role name stereotype violations concretely outlined the expected results of noun-based stereotype processing (i.e., the presence of the Stereotype Effect; Hammond-Thrasher & Järvikivi, 2023; Marrville, 2017; Oakhill et al., 2015; Banaji & Hardin, 1996). As mentioned above, this Effect was successfully replicated. However, despite the overwhelming evidence identifying the prevalence of gender stereotyped adjectives (see Hoyle et al., 2019), there was no evidence outlining the effect of adjective-level gender stereotype clashes (i.e., *compassionate* + *she* or *dominant* + *he*) during language comprehension. That is, the perception of violations of gender stereotyped adjectives has been shown to be notable (Gaucher et al., 2011), but this has yet to be examined in language processing literature. Hence, the current study's binary sensibility check analysis provides the first evidence for the differential effects of adjective- and noun- based gender stereotype processing. The smallest proportion of sensibility confirmations was given to short stories containing both adjective- and noun- level gender stereotype violations (i.e., those pairing feminine adjectives and role names or masculine adjectives and role names with male versus female pronouns, respectively), indicating that these double violations predict the largest Stereotype Effect of the possible conditions. This confirms

language processing research indicating that sociocultural world knowledge, regardless of its form, is integrated during language comprehension automatically and immediately (Guerra et al., 2021; Molinaro et al., 2016; Osterhout et al., 1997): both pieces of stereotyped language predicted this large Effect. Stories that paired adjectives congruent with the pronoun resolution but incongruent with the occupational role name's stereotypical gender (i.e., those items only containing noun-level stereotype violations) received a larger proportion of sensibility confirmations than the condition previously mentioned, indicating that adjective-level gender stereotype alignment facilitates language processing to a certain extent. In other words, an adjective that is in alignment with a pronoun confirmed lower in discourse facilitates language comprehension, even when there is a clashing noun present. This is in line with previous literature examining adjective-based facilitatory processing, where an adjective that confirms a property of an item or agent described in a sentence facilitates language comprehension (see Sedivy et al., 1999).

Interestingly, stories pairing nouns congruent with the pronoun resolution but incongruent with the descriptive adjective's stereotypical gender (i.e., those items only containing adjective-level stereotype violations) received an even higher proportion of sensibility confirmations than the condition previously mentioned, indicating that noun-level gender stereotype alignment facilitates language processing to a greater extent than adjective-level gender stereotype alignment; however, adjective-level stereotype violations still predicted lower proportions of sensibility confirmations than fully congruent items. In other words, a noun that is in alignment with a pronoun confirmed lower in discourse facilitates language comprehension, even when there is a clashing adjective present, and this facilitation is greater than adjective-level alignment. This result is explained by the fact that the subject of a sentence (i.e.,

the agent of an action) is emphasized during language processing (Redl et al., 2021; Guerra et al., 2021). In the context of these experimental vignettes, the role name serves as the agent of the action described in the second sentence: hence, the sociocultural information linked to this noun is facilitatory than the information linked to the descriptive adjective. Overall, the current study confirms that both adjective- and noun-level gender stereotype violations predict the degree to which the Stereotype Effect is experienced: double violations elicit the largest Effect, noun-level violations elicit a smaller Effect, and adjective-level violations elicit the smallest Effect.

3.3 Individual Differences Profiles

The results of the present study indicate that some individual differences in personality traits and political ideology, but *not* gender identity, significantly modulate the language comprehension of discourse containing gender stereotype clashes. In other words, similarly to previous studies (Hammond-Thrasher & Järvikivi, 2023; Puhacheuskaya & Järvikivi, 2022; Marrville, 2017; Van den Brink et al., 2012), the results suggest that a reader's personality and political profile influences how noteworthy contradictions to sociocultural world knowledge are during language comprehension. Firstly, binary sensibility check responses from the current study suggest that readers who are more conservative in their political beliefs (i.e., individuals who demonstrate wishes for traditional orders, structures, dogmatism, and discipline; Everett, 2013) are hypersensitive to items pairing feminine adjectives, masculine role names, and male pronouns compared to readers who are more liberal in their political beliefs (i.e., individuals who hold socially progressive beliefs supporting civil liberties, openness to new ideas, and fair democracy; Everett, 2013). In other words, more conservative readers provided significantly fewer sensibility confirmations for stories describing male agents fulfilling masculine gender roles who were described to be feminine by gender stereotyped adjectives. It is interesting that the Stereotype

Effect did not appear to be generalized to the same pattern of gender-specificity as the replicated character gender-specific Effect found above (i.e., increased sensitivity to descriptions of male agents fulfilling stereotypically female roles). Previous research indicates that certain political profiles predict susceptibility to this character gender-specific Stereotype Effect (Hammond-Thrasher & Järvikivi, 2023). A possible explanation for the current result is that more conservative participants are more sensitive to adjective-level gender stereotype violations such that describing male agents as possessing feminine traits elicits fewer sensibility confirmations. Here, the gender-specificity of this Effect is apparent yet contained at the adjective-level. However, this is a very weak interaction between PIQ scores and binary sensibility responses that necessitates further investigation: perhaps conservative participants are less likely to accept feminine-masculine adjective-noun pairings overall, but the evidence here is inconclusive.

Results from the current study that examined third-sentence reading times indicate that readers with high scores on the Honesty-Humility dimension of the HEXACO PI-R were hyposensitive to gender stereotype violations overall. Specifically, these participants read vignettes containing both male and female gender stereotype violations significantly faster than their less Honest-Humble counterparts. Importantly, this effect is novel in the realm of gender stereotype language processing. To understand this effect, it is important to understand what the dimension describes in terms of personality traits. Individuals who score highly on the Honesty-Humility dimension of the HEXACO PI-R feel little temptation to break rules, avoid manipulating others for personal gain, and feel no wish to elevate their social status (Ashton & Lee, 2007). Importantly, these individuals are sincere, fair, modest, and avoid greed: they are unwilling to cheat, genuine and authentic, and overall uninterested in superiority (Ashton & Lee,

2014). Research suggests that individuals who are more Honest-Humble approach language processing with increased justice and fairness concerns, meaning that they incorporate their sociocultural value of prosocial behavior (i.e., actions that help, comfort, and cooperate with others) during language comprehension (Ścigała et al., 2022). Furthermore, individuals who are more Honest-Humble avoid dishonesty and are quick to accept interpersonal differences during romantic communications (Reinhardt & Reinhard, 2023). For these reasons, it is possible that the increased proportion of sensibility confirmations granted to stories containing gender stereotype violations reflects this increased willingness to accept differences and promote fairness during human interactions. This novel result is particularly interesting, as there is a distinct lack of literature outlining the role of individual differences in Honesty-Humility on sociocultural language processing. Certainly, this effect must be investigated further to confirm its validity: determining a personality profile that facilitates the language comprehension of gender stereotype violations would pose implications for increased gender diversity overall.

3.4 Limitations and Future Endeavors

There are some limitations present in the current study that should be discussed. Firstly, despite previous literature that implicates the roles of individual differences in other traits during language processing, such as Openness to Experience, Extraversion, and Conscientiousness, no additional effects were found in the current study. The lack of an effect of gender identity can be explained by previous studies which found no effect of participant gender identity on sociocultural language processing (Hubert Lyall, 2019; Van den Brink et al., 2012). The motivation for including GQI scores as a predictor of sensibility check responses and third-sentence reading times was exploratory: including a non-dichotomous measure of individual differences in gender identity is an important methodological step towards gender

equality. Explanations for the lack of effects of other individual differences in personality traits included in the current study are difficult to determine. Firstly, third-sentence reading times, as previously mentioned, are imprecise measures of the time taken to read a target sentence (i.e., the sentence containing the pronoun resolution which captures the moment of stereotype violation). Given a more precise outline of reading times, perhaps individual differences results would be more definitive (Puhacheuskaya & Järvikivi, 2022; Hubert Lyall, 2019; Marrville, 2017). Secondly, because this is a discourse processing study, perhaps the fact that sensibility check questions took place after reading three sentences meant that the strength of the Stereotype Effect diminished slightly over time, meaning that individual difference effects were more difficult to measure in relation to direct, real-time language processing. Due to these factors, it is important to replicate this study following an online language processing paradigm. Notably, the majority of individual difference measures were uniformly distributed (Appendix I): this means that despite the convenience sampling performed, there was representation of the nuanced distributions of individual differences. Gender identity lacked normal distribution, which could be due to the limited variability of the scale used (McGuire et al., 2018).

As previously noted, binary sensibility confirmations measure directly participants' meta-linguistic judgements of vignette sensibility (i.e., conscious decisions about the relationship between vignette contents and participants' beliefs of a story's ability to make sense; Tjuka et al., 2022). Contrastively, third-sentence reading times measure both the time taken to read the given sentence and the active decision to proceed to the sensibility check question (Racine, 2014). Here, it is impossible to claim that these reading times only pertain to solely the time taken to read a sentence: this data also encodes any decisional hesitation and attentional difficulties. Hence, reading time data may be indicative of the process of language comprehension, but

cannot be taken at face value to solely encode reading time itself. For this reason, binary sensibility check responses served as the primary data source for this study. It is vital to perform a replication of the current study with precise measures of reading times: currently, a self-paced reading time version of this study is in progress to determine the actual time course of language processing throughout these same short stories.

3.5 Conclusion

The current research highlights several implications for studies of language comprehension. Firstly, in line with previous research, sociocultural world knowledge played an important role in the processing of these social pragmatic stories, where vignettes containing violations of common gender stereotypes were perceived as less sensible than stories without these violations. Importantly, this Stereotype Effect was more pronounced for violations describing male agents fulfilling feminine occupations, indicating that statements that contradict world knowledge about female gender roles have a strong influence on language comprehension. Secondly, the role of this sociocultural knowledge in the comprehension of gender stereotyped language differed based on the conditions of the stereotype violations. Double stereotype violations were perceived as the least sensible, followed by noun-level violations and then adjective-level violations. This indicates that the syntactic role of a piece of gender stereotyped language predicts the degree to which a violation of its gender influences language processing. Finally, individual differences in political ideology and Honesty-Humility predicted the processing of these stories, indicating that people possessing certain political and personality profiles may allocate more or less resources to sociocultural world knowledge during language comprehension. Notably, high degrees of conservatism predicted more sensitivity to adjective-level violations of male gender roles, while

high degrees of Honesty-Humility predicted less sensitivity to gender stereotype violations overall.

References

AAUW Report. (2023). The Simple Truth About the Gender Pay Gap. https://www.aauw.org/resources/research/simple-truth/

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness—Communion with Warmth and Morality. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01810

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*(1), 50–68. https://doi.org/10.1037/a0014411

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*(4), 340–345. https://doi.org/10.1080/00223890902935878

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, agreeableness, and emotionality factors: A review of research and theory. *Personality and Social Psychology Review, 18*, 139–152. https://doi.org/10.1177/1088868314523838

Banaji, M. R., & Hardin, C. D. (1996). Automatic Stereotyping. *Psychological Science, 7*(3), 136–141. https://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162. https://doi.org/10.1037/h0036215

Block, J. (19950601). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*(2), 187. https://doi.org/10.1037/0033-2909.117.2.187

Bouchard Jr, T., & Mcgue, M. (2003). Genetic and environmental influences on human

psychological differences. *Journal of Neurobiology, 54*, 4–45. https://doi.org/10.1002/neu.10160

Bradford, N. J., & Catalpa, J. M. (2019). Social and psychological heterogeneity among binary transgender, non-binary transgender and cisgender individuals. *Psychology & Sexuality, 10,* 69–82. https://doi.org/10.1080/19419899.2018.1552185

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*(1), 103–128. https://doi.org/10.1016/j.jml.2006.07.005

Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond Gender Stereotypes in Language Comprehension: Self Sex-Role Descriptions Affect the Brain's Potentials Associated with Agreement Processing. *Frontiers in Psychology, 6.* https://doi.org/10.3389/fpsyg.2015.01953

Caplan, D. (1996). *Language: Structure, Processing, and Disorders*. Bradford.

Castillo-Mayén, R., & Montes-Berges, B. (2014). Analysis of current gender stereotypes. *Anales de Psicología, 30*(3), 1044–1060.

Coco, N. A. (2021). Gender Identity, Discrimination, and Adjustment Among College Students. *Psychology Commons*.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition, 102*(3), 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

de Graaf, N. M., Huisman, B., Cohen-Kettenis, P. T., Twist, J., Hage, K., Carmichael, P., Kreukels, B. P. C., & Steensma, T. D. (2021). Psychological functioning in non-binary identifying adolescents and adults. *Journal of Sex & Marital Therapy*.

https://doi.org/10.1080/0092623X.2021.1950087

de Hoop, H., & Lamers, M. (2006). Incremental distinguishability of subject and object. In L. Kulikov, A. L. Malchukov, & P. de Swart (Eds.), *Case, Valency and Transitivity* (pp. 269–287). John Benjamins Publishing Company. https://doi.org/10.1075/slcs.77.17hoo

Díaz-Andreu, M. (2005). Gender identity. In *The Archaeology of Identity: Approaches to Gender, Age, Status, Ethnicity and Religion*. London, Routledge: 13-42.

Diekman, A., Goodfriend, W., & Goodwin, S. (2004). Dynamic Stereotypes of Power: Perceived Change and Stability in Gender Hierarchies. *Sex Roles, 50*, 201–215. https://doi.org/10.1023/B:SERS.0000015552.22775.44

Diessel, H. (2019). *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. The Cambridge University Press. https://doi.org/10.1017/9781108671040

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. New York, Psychology Press.

Eagly, A. H. (1997). Sex differences in social behavior: Comparing social role theory and evolutionary psychology. *American Psychologist, 52*, 1380–1383. https://doi.org/10.1037/0003-066X.52.12.1380.b

Everett, J. A. C. (2013). The 12 Item Social and Economic Conservatism Scale (SECS). *PLOS ONE, 8*(12), e82131. https://doi.org/10.1371/journal.pone.0082131

Fiawornu, V. Y. (2022, Fall). Language and Worldview in the Processing of Lies. *Education Research Archives*. https://doi.org/10.7939/r3-kdy9-mr88

Friedman, H. S., & Schustack, M. W. (2016). *Personality: Classic Theories and Modern Research*. New York, Pearson.

Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P., & Oakhill, J. (2012). Gender Representation

in Different Languages and Grammatical Marking on Pronouns: When Beauticians, Musicians, and Mechanics Remain Men. *Discourse Processes*. https://doi.org/10.1080/0163853X.2012.688184

Garnham, A., Oakhill, J., & Reynolds, D. (2002). Are inferences from stereotyped role names to characters' gender made elaboratively? *Memory & Cognition, 30*(3), 439–446. https://doi.org/10.3758/BF03194944

Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology, 101*(1), 109–128. https://doi.org/10.1037/a0022530

Glenberg, A., Becker, R., Klötzer, S., Kolanko, L., Müller, S., & Rinck, M. (2009). Episodic Affordances Contribute to Language Comprehension. *Language and Cognition, 1*, 113–135. https://doi.org/10.1515/LANGCOG.2009.006

Grant, A., Grey, S., & van Hell, J. G. (2020). Male fashionistas and female football fans: Gender stereotypes affect neurophysiological correlates of semantic processing during speech comprehension. *Journal of Neurolinguistics, 53*, 100876. https://doi.org/10.1016/j.jneuroling.2019.100876

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Grenier, C. (2006). *Political Ideology Questionnaire*.

Guerra, E., Bernotat, J., Carvacho, H., & Bohner, G. (2021). Ladies First: Gender Stereotypes Drive Anticipatory Eye-Movements During Incremental Sentence Interpretation. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.589429

Gygax, P. M., Garnham, A., & Doehren, S. (2016). What Do True Gender Ratios and Stereotype Norms Really Tell Us? *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01036

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The Times They Are a-Changing … or Are They Not? A Comparison of Gender Stereotypes, 1983–2014. *Psychology of Women Quarterly, 40*(3), 353–363. https://doi.org/10.1177/0361684316634081

Hammond-Thrasher, S., & Jarvikivi, J. (2023). Rating gender stereotype violations: The effects of personality and politics. *Frontiers in Communication, 8*. https://doi.org/10.3389/fcomm.2023.1050662

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The Multiple Dimensions of Gender Stereotypes: A Current Look at Men's and Women's Characterizations of Others and Themselves. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.00011

Hirsh, J. B., DeYoung, C. G., Xiaowen X., & Peterson, J. B. (2010). Compassionate liberals and polite conservatives: Associations of agreeableness with political ideology and moral values. *Personality & Social Psychology Bulletin, 36*(5), 655–664. https://doi.org/10.1177/0146167210366854

Hoyle, A., Wolf-Sonkin, L., Wallach, H., Agustein, I., & Cotterell, R. (2019). Unsupervised Discovery of Gendered Language through Latent-Variable Modeling. *CopeNLU*. https://copenlu.github.io/publication/2019_acl_hoyle/

Hubert Lyall, I., & Järvikivi, J. (2021). Individual Differences in Political Ideology and Disgust Sensitivity Affect Real-Time Spoken Language Comprehension. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.699071

Human Rights Campaign. (2022). Fatal Violence Against the Transgender and Gender

Non-Conforming Community in 2022. https://www.hrc.org/resources/fatal-violence-against-the-transgender-and-gender-non-conforming-community-in-2022

Kallitsounaki, A., & Williams, D. M. (2022). Implicit and explicit gender-related cognition, gender dysphoria, autistic-like traits, and mentalizing: Differences between autistic and non-autistic cisgender and transgender adults. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-022-02386-5

Kaup, B., Vega, I. de la, & Dudschig, J. S. and C. (2015). The role of sensorimotor processes in meaning composition. In *Conceptual and Interactive Embodiment*. Routledge.

Kennison, S. M., & Trofe, J. L. (2003). Comprehending Pronouns: A Role for Word-Specific Gender Stereotype Information. *Journal of Psycholinguistic Research, 32*(3), 355–378. https://doi.org/10.1023/A:1023599719948

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology, 107*(3), 371–392. https://doi.org/10.1037/a0037215

Krosnick, J. A. (1991). The Stability of Political Preferences: Comparisons of Symbolic and Nonsymbolic Attitudes. *American Journal of Political Science, 35*(3), 547–576. https://doi.org/10.2307/2111553

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition, 40*, 812–826. https://doi.org/10.3758/s13421-011-0179-8

Lechevalier, B., Petit, M. C., Eustache, F., Lambert, J., Chapon, F., & Viader, F. (1989). Regional cerebral blood flow during comprehension and speech (in cerebrally healthy subjects).

Brain and Language, 37(1), 1–11. https://doi.org/10.1016/0093-934x(89)90097-7

Lin, X., & Zhang, D. (1999). Inference in Generalized Additive Mixed Models by Using Smoothing Splines. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 61*(2), 381–400.

Lippa, R. A., Preston, K., & Penner, J. (2014). Women's Representation in 60 Occupations from 1972 to 2010: More Women in High-Status Jobs, Few Women in Things-Oriented Jobs. *PLOS ONE, 9*(5), e95960. https://doi.org/10.1371/journal.pone.0095960

Marrville, C. (2017, Fall). Gender and dominance in action: World view and emotional affect in language processing and use. *Education and Research Archive*. https://doi.org/10.7939/R3Z31P280

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McGuire, J. K., Beek, T. F., Catalpa, J. M., & Steensma, T. D. (2018). The Genderqueer Identity (GQI) Scale: Measurement and validation of four distinct subscales with trans and LGBQ clinical and community samples in two countries. *The International Journal of Transgenderism, 20*(2–3), 289–304. https://doi.org/10.1080/15532739.2018.1460735

McLellan, D. (1995). *Ideology*. U of Minnesota Press.

Moran, T., Hughes, S., Van Dessel, P., & De Houwer, J. (2022). The Role of Trait Inferences in Evaluative Conditioning. *Collabra: Psychology, 8*(1), 31738. https://doi.org/10.1525/collabra.31738

Moshagen, M., Thielmann, I., Hilbig, B. E., & Zettler, I. (2019). Meta-analytic investigations of the HEXACO Personality Inventory(-Revised): Reliability generalization, self–observer

agreement, intercorrelations, and relations to demographic variables. *Zeitschrift Für Psychologie, 227*, 186–194. https://doi.org/10.1027/2151-2604/a000377

Nieto, N., Peterson, V., Rufiner, H. L., Kamienkowski, J. E., & Spies, R. (2022). Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition. *Scientific Data, 9*(1), 52. https://doi.org/10.1038/s41597-022-01147-2

Ontario Human Rights Commission. (2023). Gender identity and gender expression. https://www.ohrc.on.ca/en/policy-preventing-discrimination-because-gender-identity-and-gender-expression/3-gender-identity-and-gender-expression

Osterhout, L., Bersick, M., & Mclaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25*(3), 273–285. https://doi.org/10.3758/BF03211283

Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language, 67*(4), 426–448. https://doi.org/10.1016/j.jml.2012.07.003

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Petersen, J., & Hyde, J. S. (2014). Chapter Two—Gender-Related Academic and Occupational Interests and Goals. In L. S. Liben & R. S. Bigler (Eds.), *Advances in Child Development and Behavior* (Vol. 47, pp. 43–76). JAI. https://doi.org/10.1016/bs.acdb.2014.04.004

Puhacheuskaya, V., & Järvikivi, J. (2022). I was being sarcastic!: The effect of foreign accent and political ideology on irony (mis)understanding. *Acta Psychologica, 222*, 103479. https://doi.org/10.1016/j.actpsy.2021.103479

R Core Team. (1969). *Graphics Package*. https://www.rdocumentation.org/packages/graphics/versions/3.6.2

R Core Team. (2020). *R Foundation for Statistical Computing*.

Redl, T., Frank, S. L., Swart, P. de, & Hoop, H. de. (2021). The male bias of a generically-intended masculine pronoun: Evidence from eye-tracking and sentence evaluation. *PLOS ONE, 16*(4), e0249309. https://doi.org/10.1371/journal.pone.0249309

Reinhardt, N., & Reinhard, M.-A. (2023). Honesty–humility negatively correlates with dishonesty in romantic relationships. *Journal of Personality and Social Psychology.* https://doi.org/10.1037/pspp0000456

Revelle, W., Wilt, J., & Condon, D. M. (2011). *Individual Differences and Differential Psychology: A brief history and prospect*. https://doi.org/10.1002/9781444343120.ch1

Rij, J. van, Wieling, M., Baayen, R. H., & Rijn, H. van. (2022). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs* (2.4.1). https://cran.r-project.org/web/packages/itsadug/index.html

Rosette, A. S., Mueller, J. S., & Lebel, R. D. (2015). Are male leaders penalized for seeking help? The influence of gender and asking behaviors on competence perceptions. *The Leadership Quarterly, 26*(5), 749–762. https://doi.org/10.1016/j.leaqua.2015.02.001

Sato, M. (2020). The neurobiology of sex differences during language processing in healthy adults: A systematic review and a meta-analysis. *Neuropsychologia, 140*, 107404. https://doi.org/10.1016/j.neuropsychologia.2020.107404

Ścigała, K. A., Arkoudi, I., Schild, C., Pfattheicher, S., & Zettler, I. (2022). The relation between Honesty-Humility and moral concerns as expressed in language. *PsyArXiv*. https://doi.org/10.31234/osf.io/7x6f5

Sedivy, J. C., K. Tanenhaus, M., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*(2), 109–147. https://doi.org/10.1016/S0010-0277(99)00025-6

Settle, J. E., Dawes, C. T., & Fowler, J. H. (2009). The heritability of partisan attachment. *New York University Publications*. https://nyuscholars.nyu.edu/en/publications/the-heritability-of-partisan-attachment

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Slugoski, B. R., & Turnbull, W. (1988). Cruel to be kind and kind to be cruel: Sarcasm, banter and social relations. *Journal of Language and Social Psychology, 7*(2). https://doi.org/10.1177/0261927X880070020

Stott, M. (2019). *Word Meaning and Context: A Critical Examination of Contrasting Perspectives* [MA by Research, Sheffield Hallam University]. https://doi.org/10.7190/shu-thesis-00144

Tjuka, A., Forkel, R., & List, J.-M. (2022). Linking norms, ratings, and relations of words and concepts across multiple language varieties. *Behavior Research Methods, 54*(2), 864–884. https://doi.org/10.3758/s13428-021-01650-1

Van Berkum, J. J. A., de Goede, D., van Alphen, P., Mulder, E., & Kerstholt, J. H. (2013). How robust is the language architecture? The case of mood. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00505

van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for

inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience, 7*(2), 173–183. https://doi.org/10.1093/scan/nsq094

Wood, S. (2017). *Generalized Additive Models*. Chapman and Hall.

Zettler, I., Thielmann, I., Hilbig, B. E., & Moshagen, M. (2020). The nomological net of the HEXACO model of personality: A large-scale meta-analytic investigation. *Perspectives on Psychological Science, 15*, 723–760. https://doi.org/10.1177/1745691619895036

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. *Psychological Science, 13*(2). http://www.jstor.org/stable/40063860

Appendix A

Participant Demographics - Gender

Male	99
Female	110
Prefer not to say	2
TOTAL	211

Participant Demographics - Country of Residence

Canada	138
United States of America	73
TOTAL	211

Participant Demographics - Age

Age Range	**Median Age**	**Mean Age**
19-76	37	37.9

Appendix B

Adjective Ratings

Participants were asked to rate how masculine or feminine a set of adjectives appeared to be, respectively, on a five-point Likert scale. Average ratings across all items were 3.10, while median ratings were 3.06. Those highlighted as blue were accepted as masculine; those highlighted in gray were accepted as feminine; those highlighted in red were accepted as feminine.

Adjective	**Rating**
Buff	1.500000000
Aggressive	1.711538462
Dominant	1.826923077
Forceful	1.826923077
Tough	1.961538462
Boastful	2.096153846
Insensitive	2.115384615
Competitive	2.192307692
Assertive	2.230769231
Strong	2.307692308
Stern	2.326923077
Blunt	2.346153846
Uncooperative	2.346153846
Cruel	2.442307692
Daring	2.480769231
Powerful	2.480769231
Loud	2.519230769
Robust	2.538461538
Dexterous	2.615384615
Decisive	2.711538462
Peaceful	2.711538462
Average height	2.730769231
Seasoned	2.730769231
Particular	2.750000000
Analytical	2.750000000

Rude	2.769230769
Frantic	2.807692308
Trained	2.846153846
Impartial	2.903846154
Skillful	2.942307692
Proficient	2.942307692
Outspoken	2.942307692
Competent	2.980769231
Old	3.013423761
Neutral	3.038461538
Normal	3.038461538
Experienced	3.057692308
Excellent	3.057692308
Coherent	3.096153846
Accurate	3.115384615
New	3.134615385
Objective	3.153846154
Independent	3.153846154
Intelligent	3.173076923
Talented	3.211538462
Consistent	3.230769231
Young	3.250000000
Average weight	3.269230769
Whiny	3.346153846
Clever	3.384615385
Weak	3.423076923
Quiet	3.519230769
Shy	3.576923077
Fussy	3.576923077
Classy	3.615384615
Shy	3.615384615
Vibrant	3.692307692
Organized	3.750000000
Tender	3.826923077
Flexible	3.826923077
Sentimental	3.884615385
Gentle	3.942307692
Sensitive	3.942307692

Affectionate	3.961538462
Submissive	3.980769231
Compassionate	4.019230769
Emotional	4.057692308
Ditsy	4.076923077
Flirtatious	4.096153846
Soft	4.115384615
Lovely	4.365384615
Nurturing	4.423076923
Beautiful	4.500000000
Gorgeous	4.711538462

Role Noun Ratings

Participants were asked to rate how masculine or feminine a set of role nouns appeared to be, respectively, on a five-point Likert scale. Average ratings across all items were 3.01, while median ratings were 3.00. Those highlighted as blue were accepted as masculine; those highlighted in gray were accepted as feminine; those highlighted in red were accepted as feminine.

Role Noun	**Rating**
Truck driver	1.403846154
Heavy equipment operator	1.403846154
Plumber	1.538461538
Construction worker	1.557692308
Miner	1.557692308
Butcher	1.576923077
Soldier	1.576923077
Carpenter	1.634615385
Groundskeeper	1.692307692
Firefighter	1.750000000
Electrician	1.788461538
Car dealer	1.788461538
Building contractor	1.807692308
Farmer	1.846153846
Maintenance worker	1.865384615

President	1.903846154
Engineer	1.942307692
Golfer	1.980769231
Pilot	2.000000000
Stockbroker	2.019230769
Sailor	2.115384615
Mayor	2.115384615
Police officer	2.173076923
Boss	2.173076923
Politician	2.250000000
FBI agent	2.250000000
Professional athlete	2.269230769
Surgeon	2.326923077
Dentist	2.480769231
Architect	2.576923077
Junior	2.596153846
Professor	2.615384615
Geologist	2.634615385
Physicist	2.653846154
Banker	2.692307692
Seatmate	2.807692308
Delegate	2.807692308
Spectator	2.865384615
Patron	2.865384615
Finalist	2.923076923
Undergraduate	2.942307692
Pedestrian	2.961538462
Visitor	2.961538462
Adult	2.961538462
Person	2.961538462
Onlooker	2.961538462
Neighbour	2.980769231
Observer	2.980769231
Eyewitness	2.980769231
Participant	2.980769231
Guest	2.980769231
Pupil	2.980769231
Student	3.000000000

Musician	3.000000000
Child	3.000000000
Individual	3.019230769
Client	3.019230769
Consumer	3.019230769
Classmate	3.057692308
Registrant	3.057692308
Customer	3.057692308
Sightseer	3.076923077
Passenger	3.115384615
Attendee	3.173076923
Witness	3.173076923
Fan	3.173076923
Sophomore	3.192307692
Cashier	3.230769231
Judge	3.230769231
File clerk	3.230769231
Concert goer	3.230769231
Listener	3.269230769
Server	3.500000000
Artist	3.500000000
Teacher	3.538461538
Veterinarian	3.557692308
Restaurant host	3.576923077
Singer	3.596153846
Counselor	3.596153846
Dietician	3.711538462
Personal care worker	3.730769231
Cleaner	3.750000000
Administrative assistant	3.769230769
Jewelry salesperson	3.769230769
Designer	3.807692308
Dental hygienist	3.903846154
Social media influencer	3.903846154
Social worker	3.923076923
Childcare worker	4.076923077
Librarian	4.096153846
Florist	4.153846154

Dancer	4.153846154
Receptionist	4.192307692
Secretary	4.230769231
Model	4.269230769
Nurse	4.269230769
Housekeeper	4.288461538
Babysitter	4.307692308
Hair Stylist	4.365384615
Prostitute	4.519230769
Manicurist	4.519230769
Beautician	4.519230769
Nanny	4.653846154
Maid	4.692307692
Cheerleader	4.750000000

Appendix C

Experimental Items

Item	Type	Sentence 1	Sentence 2	Sentence 3
0	Filler	You are looking across the busy conference hall.	You see the independent attendee pick up a nametag.	You watch them pin it on a shirt.
1	Filler	You are looking out the window of your house.	You see the old neighbor examine the lawn.	You watch them start up the lawn mower.
2	Filler	You are watching the street outside your office.	You see the accurate observer read a street sign.	You watch them walk down the long street.
3	Filler	You are observing a small traffic accident.	You see the coherent eyewitness observe the cars.	You watch them write down the license plates.
4	Filler	You are running an experiment for an assignment.	You see the competent participant focus on the computer screen.	You watch them press down the space key.
5	Filler	You are hosting a dinner party in your house.	You see the consistent guest arrive at the door.	You watch them enter into the main hallway.
6	Filler	You are sitting in class at your school.	You see the outspoken pupil look at the assignment.	You watch them pick up a long pencil.
7	Filler	You are working on campus in the afternoon.	You see the experienced student work on an essay.	You watch them type on a computer keyboard.
8	Filler	You are listening to a live concert in your city.	You see the excellent musician pick up a guitar.	You watch them walk onto the bright stage.
9	Filler	You are helping out a class with homework.	You see the intelligent child look at a paper.	You watch them write down the times-tables.
10	Filler	You are walking across the center of campus.	You see the new individual enter into the quad.	You watch them walk toward the science building.
11	Filler	You are sitting inside	You see the neutral	You watch them hang

		the office in the morning.	client talk into a phone.	up a short call.
12	Filler	You are shopping for clothes at the mall.	You see the young consumer look at a shirt.	You watch them read the price tag.
13	Filler	You are inside a lecture hall on campus.	You see the proficient classmate take down some notes.	You watch them look at the projector screen.
14	Filler	You are observing a conference downtown.	You see the skillful registrant pick up a nametag.	You watch them write down a short name.
15	Filler	You are looking for winter coats at the mall.	You see the trained customer look at a price tag.	You watch them put back a big sweater.
16	Filler	You are walking to your office downtown.	You see the talented person dodge by a traffic cone.	You watch them continue down the long street.
17	Filler	You are enjoying the mountains on your holiday.	You see the normal sightseer pick up binoculars.	You watch them observe a small bird.
18	Filler	You are sitting on the train to work.	You see the impartial passenger sit by a newspaper.	You watch them drink from a coffee mug.
19	Filler	You are watching the television at the office.	You see the objective onlooker focus on the news.	You watch them turn to a newspaper crossword.
20	Congruent MASC	You are driving a car down the highway.	You see the buff truck driver turn into the next lane.	You watch him check for traffic in the side mirror.
21	Incongruent MASC Congruent ADJ	You are walking a dog down the street.	You see the aggressive heavy equipment operator turn on the machine.	You watch her drill a hole into the ground.
22	Congruent MASC Incongruent ADJ	You are entering the bathroom in your house.	You see the submissive plumber take off the sink pipe.	You watch him screw on a new pipe.
23	Incongruent MASC	You are carrying a box down the road.	You see the gentle construction worker climb up the scaffolding.	You watch her hammer in a long nail.
24	Congruent MASC	You are passing by the	You see the tough	You watch him hold

		quarry in your town.	miner enter into the cave.	onto the big toolkit.
25	Incongruent MASC Congruent ADJ	You are coming into the shop on mainstreet.	You see the boastful butcher slice into the meat.	You watch her wrap up the fresh cold cuts.
26	Congruent MASC Incongruent ADJ	You are watching a movie about the war.	You see the insensitive soldier fire off a gun.	You watch him shoot straight at the crowd.
27	Incongruent MASC	You are looking at the woodshop on your street.	You see the ditsy carpenter saw into the wood.	You watch her pick up the many pieces.
28	Congruent MASC	You are looking at the house on fire.	You see the strong fireman hold onto the water hose.	You watch him spray water at the house.
29	Incongruent MASC Congruent ADJ	You are standing in the doorway of your house.	You see the assertive electrician pick up the soldering iron.	You watch her meld together the metal wires.
30	Congruent MASC Incongruent ADJ	You are sitting inside the warehouse down the street.	You see the flirtatious car dealer point at a sports car.	You watch him gesture at the leather seats.
31	Incongruent MASC	You are sitting on a bench in your neighborhood.	You see the fussy building contractor look at the ladder.	You watch her give orders to the crew.
32	Congruent MASC	You are passing by the fields by your house.	You see the stern farmer pick through the crops.	You watch him pull out a long shovel.
33	Incongruent MASC Congruent ADJ	You are resting inside the mall in your city.	You see the uncooperative maintenance worker work on the door.	You watch her pick up a normal screwdriver.
34	Congruent MASC Incongruent ADJ	You are watching the television in your bedroom.	You see the nurturing president talk into a microphone.	You watch him look at the bright teleprompter.
35	Incongruent MASC	You are working on homework on your campus.	You see the beautiful engineer pass by your table.	You watch her talk into a cellular phone.
36	Congruent MASC	You are looking at the outdoors club by your house.	You see the powerful golfer get into the golf cart.	You watch him talk to the good driver.
37	Incongruent MASC Congruent ADJ	You are entering the airplane at the airport.	You see the robust pilot pick up the microphone.	You watch her point down the seating row.

38	Congruent MASC Incongruent ADJ	You are eating at the restaurant in the office building.	You see the classy stockbroker speak into a phone.	You watch him nod at a nearby colleague.
39	Incongruent MASC	You are walking through the field outside your house.	You see the shy groundskeeper pick up the shears.	You watch her trim off a tree branch.
40	Congruent FEM	You are sitting in the waiting room in the clinic.	You see the soft dental hygienist look at the clock.	You watch her pick up the water pick.
41	Incongruent FEM Congruent ADJ	You are walking through the mall in your city.	You see the flexible social media influencer look at the flowers.	You watch him pose for a quick selfie.
42	Congruent FEM Incongruent ADJ	You are passing by the school down the street.	You see the forceful social worker pick up a teddy bear.	You watch her walk towards the big class.
43	Incongruent FEM	You are walking by the daycare in your town.	You see the cruel childcare worker play with the crayons.	You watch him draw on a piece of paper.
44	Congruent FEM	You are sitting in the big hall on campus.	You see the compassionate librarian pick up a book.	You watch her put the book on a shelf.
45	Incongruent FEM Congruent ADJ	You are walking into the shop on mainstreet.	You see the emotional florist wrap up the flowers.	You watch him tape on a curled ribbon.
46	Congruent FEM Incongruent ADJ	You are passing by the studio on campus.	You see the competitive dancer tie up the shoes.	You watch her stretch out on the barre.
47	Incongruent FEM	You are sitting inside the office in your building.	You see the blunt secretary turn on the computer.	You watch him sign in for the day.
48	Congruent FEM	You are passing through the mall in the city.	You see the gorgeous model walk down the hallway.	You watch her turn around and walk back.
49	Incongruent FEM Congruent ADJ	You are sitting in the hospital waiting room.	You see the lovely nurse pick up a clipboard.	You watch him write down a short name.
50	Congruent FEM Incongruent ADJ	You are eating in your house in the morning.	You see the daring housekeeper pick up the dirty dishes.	You watch her place them in the dishwasher.

51	Incongruent FEM	You are walking through the park by your house.	You see the insensitive babysitter play with a sand bucket.	You watch him scoop up the dry sand.
52	Congruent FEM	You are going to an appointment at the studio.	You see the affectionate hairstylist clean up the shelf.	You watch her sweep up the dirty floor.
53	Incongruent FEM Congruent ADJ	You are walking down the street at night.	You see the vibrant prostitute get into a car.	You watch him drive off into the distance.
54	Congruent FEM Incongruent ADJ	You are arriving at an appointment at the salon.	You see the loud manicurist clean up the workstation.	You watch her wipe down the dirty surface.
55	Incongruent FEM	You are leaving the salon in the afternoon.	You see the dominant beautician wipe down the workstation.	You watch him pick up the used products.
56	Congruent FEM	You are passing by the playground down the street.	You see the organized nanny wave down the children.	You watch her walk towards a big house.
57	Incongruent FEM Congruent ADJ	You are reading in the living room of your house.	You see the tender maid wipe down the table.	You watch him turn towards the main hallway.
58	Congruent FEM Incongruent ADJ	You are watching the tournament on campus.	You see the decisive cheerleader lace up the shoes.	You watch her walk onto the bright stage.
59	Incongruent FEM	You are sitting inside the art gallery in your city.	You see the dexterous receptionist sketch on a notepad.	You watch him add on some bright color.

Appendix D

HEXACO PI-R 60-Item Version

Each item is read and ratings of agreement are elicited across a five-point Likert scale.

Item	Statement
1	I would be quite bored by a visit to an art gallery.
2	I plan ahead and organize things, to avoid scrambling at the last minute.
3	I rarely hold a grudge, even against people who have badly wronged me.
4	I feel reasonably satisfied with myself overall.
5	I would feel afraid if I had to travel in bad weather conditions.
6	I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed.
7	I'm interested in learning about the history and politics of other countries.
8	I often push myself very hard when trying to achieve a goal.
9	People sometimes tell me that I am too critical of others.
10	I rarely express my opinions in group meetings.
11	I sometimes can't help worrying about little things.
12	If I knew that I could never get caught, I would be willing to steal a million dollars.
13	I would enjoy creating a work of art, such as a novel, a song, or a painting.
14	When working on something, I don't pay much attention to small details.
15	People sometimes tell me that I'm too stubborn.
16	I prefer jobs that involve active social interaction to those that involve working alone.
17	When I suffer from a painful experience, I need someone to make me feel comfortable.
18	Having a lot of money is not especially important to me.
19	I think that paying attention to radical ideas is a waste of time.
20	I make decisions based on the feeling of the moment rather than on careful thought.
21	People think of me as someone who has a quick temper.
22	On most days, I feel cheerful and optimistic.
23	I feel like crying when I see other people crying.
24	I think that I am entitled to more respect than the average person is.
25	If I had the opportunity, I would like to attend a classical music concert.
26	When working, I sometimes have difficulties due to being disorganized.
27	My attitude toward people who have treated me badly is "forgive and forget".
28	I feel that I am an unpopular person.
29	When it comes to physical danger, I am very fearful.
30	If I want something from someone, I will laugh at that person's worst jokes.

31	I've never really enjoyed looking through an encyclopedia.
32	I do only the minimum amount of work needed to get by.
33	I tend to be lenient in judging other people.
34	In social situations, I'm usually the one who makes the first move.
35	I worry a lot less than most people do.
36	I would never accept a bribe, even if it were very large.
37	People have often told me that I have a good imagination.
38	I always try to be accurate in my work, even at the expense of time.
39	I am usually quite flexible in my opinions when people disagree with me.
40	The first thing that I always do in a new place is to make friends.
41	I can handle difficult situations without needing emotional support from anyone else.
42	I would get a lot of pleasure from owning expensive luxury goods.
43	I like people who have unconventional views.
44	I make a lot of mistakes because I don't think before I act.
45	Most people tend to get angry more quickly than I do.
46	Most people are more upbeat and dynamic than I generally am.
47	I feel strong emotions when someone close to me is going away for a long time.
48	I want people to know that I am an important person of high status.
49	I don't think of myself as the artistic or creative type.
50	People often call me a perfectionist.
51	Even when people make a lot of mistakes, I rarely say anything negative.
52	I sometimes feel that I am a worthless person.
53	Even in an emergency I wouldn't feel like panicking.
54	I wouldn't pretend to like someone just to get that person to do favors for me.
55	I find it boring to discuss philosophy.
56	I prefer to do whatever comes to mind, rather than stick to a plan.
57	When people tell me that I'm wrong, my first reaction is to argue with them.
58	When I'm in a group of people, I'm often the one who speaks on behalf of the group.
59	I remain unemotional even in situations where most people get very sentimental.
60	I'd be tempted to use counterfeit money, if I were sure I could get away with it.

Appendix E

Political Ideology Questionnaire

Each item is read. Items 1-15 elicit ratings of "For" or "Against" on a five-point Likert scale.

Items 16-27 elicit ratings of agreement on a five-point Likert scale.

Item	Statement
1	Prayer in schools?
2	Pro-choice (abortion)?
3	Cuts to welfare programs?
4	National healthcare?
5	Sex education in elementary schools?
6	Gun control?
7	Stronger labor unions?
8	Contraception?
9	Food stamp programs?
10	Same-sex marriage?
11	Aid/care for the homeless?
12	Minimum wages?
13	Political Correctness?
14	Racial quotas in the workplace?
15	Capital punishment?
16	It is better to keep things the way they are.
17	People are essentially selfish; they need to be controlled.
18	Individuals have free will; they are responsible for their own lives and problems.
19	The traditional family must be preserved at all costs.
20	Government regulations are needed to control monopolies.
21	A free market economy is the best way to ensure prosperity and fulfillment of individual needs.
22	Sometimes revolutions are necessary.
23	This country would be better off if most government programs were eliminated.
24	People are basically good but can be corrupted.
25	The free market economic system is basically exploitative and inherently unfair to working people.
26	Helping the poor encourages laziness.
27	If the rich continue to get richer and the poor to get poorer, I would support a violent revolution to correct the inequality.

Appendix F

Genderqueer Identity Scale

Items are read and ratings of agreement are elicited on a five-point Likert scale.

Item	Statement
1	I have done research about gender theory and gender roles.
2	I try to convince people that society should not insist on a gender binary.
3	I try to convince people that society expects people to be too gender conforming.
4	Around me, I make sure people are free to express whatever gender roles they want.
5	My gender role is important because I push society to question traditional gender roles.
6	I encourage other people to be more open minded about gender roles.
7	I expect my gender expression to be traditional.
8	My gender expression will likely not change much from day to day.
9	I expect that people will rarely question my gender.
10	I expect my gender to be fluid and change over time.
11	I will have a non-traditional gender role (be gender non-conforming).

Appendix G

Model Output for the Interaction Between Key Press Responses and PIQ Scores

	edf	Ref. df	Chi.sq	p-value
s(PIQs):conditionCongruentFEM	1.000e+00	1.000	0.242	0.6226
s(PIQs):conditionCongruentFEM IncongruentADJ	1.898e+00	2.303	1.282	0.5134
s(PIQs):conditionCongruentMASC	1.000e+00	1.000	0.002	0.9639
s(PIQs):conditionCongruentMASC IncongruentADJ	3.488e+00	4.333	9.929	0.0473 *
s(PIQs):conditionIncongruentFEM	1.974e+00	2.327	2.340	0.3110
s(PIQs):conditionIncongruentFEM CongruentADJ	1.000e+00	1.000	0.228	0.6331
s(PIQs):conditionIncongruentMASC	1.000e+00	1.000	0.376	0.5398
s(PIQs):conditionIncongruentMASC CongruentADJ	1.000e+00	1.001	0.121	0.7281
s(RName)	3.707e+01	38.000	4738.395	< 2e-16 ***
s(PP)	1.998e+02	206.000	1612.250	< 2e-16 ***

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Formula: keys ~ condition + s(PIQs, by = condition, bs = "tp") +
s(pp, by = condition, bs = "fs") + s(Rname, bs = "re") + s(pp, bs = "re")

Note: RName refers to Item and PP refers to Participant.

Appendix H

Model Output for the Interaction Between Third-Sentence Reading Times and Honesty-Humility Scores on the HEXACO PI-R

	edf	Ref. df	F	p-value
s(HONs):CxRCongruentFEM	1.000e+00	1.000	9.445	0.002124 **
s(HONs):CxRCongruentMASC	1.000e+00	1.000	9.893	0.001665 **
s(HONs)CxRIncongruentFEM	1.697e+00	2.131	4.525	0.009994 **
s(HONs)CxRIncongruentMASC	5.279e+00	6.370	4.078	0.000349 ***
s(log(s2rt * 1000))	7.326e+00	8.337	75.416	< 2e-16 ***
s(RName)	3.650e+01	38.000	28.685	< 2e-16 ***
s(PP)	1.961e+02	206.000	11.601	< 2e-16 ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Formula: s3rt ~ CxR + s(HON, by = CxR) + s(pp, by = CxR, bs = "fs", m = 1) +
s(log(s2rt * 1000)) + s(Rname, bs = "re") + s(pp, bs = "re")

Note: RName refers to Item and PP refers to Participant.

Appendix I

Distribution of Individual Differences Measures by Participant

HON denotes Honesty-Humility, EMO denotes Emotionality, EXTRA denotes Extraversion, AGREE denotes Agreeableness, CONSC denotes Conscientiousness, OPEN denotes Openness to Experience, PIQ denotes Political Ideology, and GIQ denotes Genderqueer Identity.