

LinkedIn: A New Frontier for Gender Bias in Hiring Practices

Honors Thesis by
Marwa Houalla

In Partial Fulfillment of the Requirements for the
Honors Bachelors Degree in Computer Science



UNIVERSITY OF MICHIGAN
Ann Arbor, Michigan

2023

Presented April 7, 2023

© 2023

Marwa Houalla
UMID: 60076175

All rights reserved

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to Professor Fabiana Silva for her unwavering support, invaluable guidance, and expertise in overseeing my research. As a Computer Scientist with a limited background in sociology, her feedback and mentorship were instrumental in shaping the direction of my honors thesis. I am deeply grateful for her patience, encouragement, and for always making herself available to answer my questions. I would also like to extend my thanks to Professor Hector Garcia Ramirez, my co-supervisor, for his additional support and for providing valuable insights into the subject matter. His advice and recommendations were instrumental in refining my research methodology and in ensuring the rigor and validity of my findings. Lastly, I would like to acknowledge the countless friends who have provided feedback and support throughout my academic journey. In particular, I would like to thank Ethan Graber for his unwavering support, willingness to listen, and for always providing thoughtful and constructive feedback during the editing process.

ABSTRACT

In the realm of employment, the importance of recommendations is well-established. Yet, the impact of gender-based differences in recommendation content remains underexplored. Further, prior research has largely neglected the use of LinkedIn for professional recommendations. To address these gaps in literature, I conducted an original study with 1,000 recent university graduates to investigate gender differences in the quality and quantity of recommendations given and received by men and women. I used natural language processing methods and textual analysis to extract and classify recommendation content. My findings reveal mixed results with regards to gender homophily in recommendations. While women were similarly likely to receive recommendations from men and women, men were more likely to receive recommendations from men than women. In terms of content, the key difference I find is that men are described as more agentic and meaningfully managerial than women, with fairly minimal differences in other categories. For instance, both men and women received similar amounts of praise and positive valence in their recommendations. Together, these findings may reinforce gender stereotypes and have implications for organizations, hiring managers, and LinkedIn users evaluating job candidates.

Chapter 1

INTRODUCTION

1.1 Introduction

Letters of recommendation have been the subject of extensive scholarly inquiry in recent years. These letters supplement traditional hiring methods, notably resumes and interviews, and have been found to exert a significant influence on interview opportunities and ranking within applicant pools (Bureau of National Affairs, 1988; Friedman and Williams, 1982; Levy-Leboyer, 1994). A resume audit study demonstrated that the inclusion of a reference letter can increase employer call-backs by up to 60% (Abel, 2017). Furthermore, research indicates that firms often rely on recommendations from current employees as a key factor in their recruitment processes, with informal networks accounting for approximately 50% of all job placements in the United States, and around 70% of organizations implementing referral-based hiring initiatives (Burks, 2015). Given the demonstrated importance of recommendations in the hiring landscape, there is a pressing need for deeper research on their role and impact.

1.2 Theorizing Gender Differences in Recommendations

Extensive research in social networks has revealed the prevalence of *gender homophily*¹ across various labor market levels, including research collaborations and high-status² occupations (McPherson et al., 2001; Turrentine 2018; Madera, 2018; Kwiek, 2020; Williams, 2021). The persistence of gender homophily in senior hiring processes suggests the potential for gender bias to be more pervasive than previously thought. Thus, this study aims to examine the potential for gender differences in the recommendation process, which is crucial to understanding the mechanisms underlying gender inequality in hiring processes.

Are women and men equally likely to give and receive recommendations? Given the prevalence of gender homophily in social networks, it is necessary to scrutinize gender-based differences in the frequency of giving and receiving recommendations. Drawing from literature on gender homophily and social perception biases, I propose two hypotheses. Hypothesis 1a posits that men and women are more likely to *give* recommendations to individuals of the same gender. This hypothesis is supported by the well-established tendency of individuals to form social connections with

those who share similar characteristics, such as gender, as well as the empirical evidence that gender homophily exists across social networks (McPherson, Smith-Lovin, and Cook, 2001; Zhang et al, 2021). Similarly, Hypothesis 1b suggests that women and men are more likely to *receive* recommendations from individuals of the same gender, which may be attributed to gender homophily and the perception of same-gender individuals as more competent and trustworthy.

Another critical aspect of recommendations that merits examination is gender-based differences in the content of recommendations. Prior research on gender bias has suggested that women may encounter disadvantages when receiving recommendations. For example, Moss-Racusin (2012) found that letters of recommendation for female applicants emphasized communal qualities, while those for male applicants emphasized agentic qualities³. These findings suggest that gender bias may affect not only the quantity of recommendations received but also the quality and content of those recommendations.

Hypothesis 2 proposes that men are described as more agentic and less communal compared to women in their recommendations, based on pervasive gender stereotypes in western societies that associate men with assertiveness, competitiveness, and leadership, and women with warmth, empathy, and nurturing. These stereotypes may shape how recommenders evaluate and describe the skills and qualities of male and female candidates, resulting in differential emphasis on agentic vs. communal traits.

This study attempts to provide insights into gender differences in the recommendation process on LinkedIn. By examining the specific ways in which gender operates in the context of recommendations, this study aims to expand and clarify existing knowledge on the topic and contribute to the development of more equitable hiring practices.

1.3 Methodological Limitations of Previous Empirical Research

Despite an extensive literature on referrals and recommendations in the labor market, we still have much to learn about gender differences in recommendations. Social network literature⁴ has made significant contributions to understanding gender homophily in job referrals (see Zeltzer, 2020; Neugart, 2019; Farzana, 2022; Beugnot, 2020). However, subsequent studies have been limited in their external validity⁵ and generalizability to actual hiring practices. Consider the laboratory finding that women exhibit gender homophily in referral practices by favoring female candi-

dates, while men have no preference (Beugnot, 2020). The relevance of this finding to practical hiring contexts remains uncertain due to the complexities of real-world hiring contexts, which can be influenced by job requirements and company culture.

Moreover, two studies found that gender homophilous networks play a crucial role in the gender segregation of jobs (Fernandez and Sosa, 2005; Brown et al., 2016). Fernandez and Sosa (2005) found that female employees in a U.S. call center referred 75% women, while male employees referred only 44% men. Brown et al. (2016) discovered that 63.5% of job referrals in a mid-sized US corporation were between individuals of the same gender. Nonetheless, the extent to which these findings are applicable across different industries and contexts is unknown.

In addition to examining the structure of social networks, it is also important to consider the linguistic content of recommendations and how it may affect gender differences in hiring processes. While much of the existing literature on social networks has focused on gender homophily, research specifically examining the use of gendered language⁶ in recommendations has shed light on the systematic differences in the way female and male applicants are described (see Aamodt, 1993; Moss-Racusin et al., 2012). Numerous studies have consistently found that letters of recommendation for female applicants differ from those for male applicants in terms of length, the use of negative language, as well as the description of women as more communal and less agentic⁷, and the use of fewer ability words and more grindstone words⁸ (Trix and Psenka, 2003; Madera et al., 2009; Schmader et al., 2007); recommenders also tend to use more standout adjectives and "authentic" words⁹ to describe male candidates (Lin et al., 2019).

Yet, the scholarship on gendered language in recommendations suffers from limited sample sizes that are not representative of broader populations. All studies mentioned in the preceding literature have exclusively drawn participants from medical or academic settings, with the largest study involving only 886 participants. One study that broadened its scope analyzed an international dataset of 1,224 recommendation letters for postdoctoral fellowships in the geosciences and found that female applicants were only half as likely to receive excellent letters¹⁰ compared to male applicants (Kuheli Dutt et al., 2016). Although widely acknowledged as a significant step towards understanding gender differences in recommendation letters, the study's exclusive focus on geosciences highlights the need for future research that includes a more diverse range of industries and job positions.

To address the methodological limitations of previous studies, this study employs a

novel approach that analyzes recommendations¹¹ sourced from LinkedIn, a popular online platform widely used for professional networking. This study broadens the scope of previous research, which has traditionally concentrated on a single industry, by examining recent university graduates across diverse industries. I provide an accurate reflection of the evolving social landscape and subtle nuances of gendered language used in contemporary recommendation practices, effectively bridging gaps in the existing literature on gender homophily and gendered language.

Chapter 2

THE STUDY

2.1 The Study

I conducted a two-stage study to examine gender homophily and gendered language in recommendation letters for graduates with bachelor's degrees from the University of Michigan's class of 2017. In the first stage, I employed a sampling strategy to obtain a representative sample of graduates, with targeted sampling of graduates from popular majors. In the second stage, I collected recommendations from the selected graduates' LinkedIn profiles using a Python script, and conducted textual analysis to identify average length as well as gender-associated words and phrases.

A limitation of this two-stage approach may be that I do not observe a complete picture of recommendation practices. Nevertheless, by focusing on LinkedIn as a data source, I was able to capture a broad range of recommendations given to recent university graduates. First, LinkedIn is a widely used platform for professional networking and career development, making it a relevant site for studying contemporary recommendation practices. Second, I randomly selected a sample of graduates and analyzed their LinkedIn profiles, rather than relying solely on profiles recommended by LinkedIn's algorithm, which could have introduced bias towards more popular profiles. As an example, we could imagine that—on average—there are gender differences in the number of recommendations received by men and women University of Michigan graduates on LinkedIn, but that there are no gender differences in the number of recommendations received by the most popular University of Michigan profiles (perhaps because they all receive recommendations). Thus, if we restricted our sample to those that were recommended via LinkedIn's algorithm, we would not be able to make broader claims about gender differences in the number of recommendations received by the full population of University of Michigan graduates on LinkedIn. Third, my study focuses on a sample of University of Michigan graduates, who are a significant target of recruitment efforts from various industries. While this study does not capture the recommendation practices of all demographic groups or industries, my findings provide insights into the referral practices of a significant population of early-career professionals.

2.2 Sampling Frame: University of Michigan Alumni

The Office of University Development (OUD) at the University of Michigan provided me with a complete list of University of Michigan alumni who graduated in 2017 and earned a bachelor's degree from any of the three campuses: Ann Arbor, Flint, Dearborn. The dataset includes information on the graduates' full names, first majors, second majors (if applicable), gender identity, consistent aggregate degree, and race, among others. I ensured the privacy of the graduates by limiting my analysis to publicly available information, as part of my efforts to safeguard their confidentiality.

To minimize selection bias, I utilized the complete list of alumni instead of relying exclusively on LinkedIn for my sample selection. Depending solely on LinkedIn profiles could introduce unintended bias towards those who have created profiles or maintain active profiles, which could result in a non-representative sample.

My final analytic sample comprised 1,000 graduates randomly selected from the full list of 9,593, 2017 graduates (4,598 men and 4,995 women), with equal representation of 500 men and 500 women. Graduates were chosen early in their professional careers so I could focus on the role of recommendations in the hiring process, rather than on the candidate's broader professional background. I also chose recent graduates to ensure a comprehensive sampling process, given that women tend to change their last names after marriage, which would make it more difficult to locate them in online searches. Only graduates' first majors were considered when selecting samples, and for those whose LinkedIn profiles could not be found, I documented this and expanded the scope of the search¹².

I randomly selected 500 graduates from the ten most popular majors and 500 from the remaining majors, with equal representation of men and women. Specifically, I selected 25 men and women from each of the ten most popular majors, and 250 men and 250 women from across the broader population of majors. This approach allowed me to compare the recommendations received by men and women in the same major, which is important given the significant gender segregation across fields of study (Werfhorst, 2017) and the potential variation in the use of LinkedIn recommendations across fields. The descriptive statistics for the ten most popular majors are presented in Table 1, listed in order from most to least popular.

2.3 Profile Collection

This study involved the collection of public LinkedIn profiles using a custom Python script. The script utilized web scraping techniques, including the requests library and the regex module. The script's primary function was to accept a full name and search for that name on Google, after which it extracted the public profile links of individuals matching that name on LinkedIn. To evade detection as an automated bot, the script incorporated a random user agent selection to vary the headers of HTTP requests¹³ and a sleep timer between requests to mimic human browsing behavior, which prevents sending too many requests too quickly.

To reduce the risk of being blocked or challenged¹⁴ and minimize the number of requests made to LinkedIn, I limited the number of search results returned per query. These measures were implemented to overcome the limitations associated with relying on Google's search engine, which can be subject to IP blocking or CAPTCHA challenges.

Upon completion of the search, I manually reviewed the collected profiles to ensure that they accurately matched the individuals I was searching for. The script returned a list of profile links, which I parsed using regular expressions to extract clean links for further analysis.

2.4 Recommendation Data Extraction

A secondary Python script was developed to gather public information on the given and received recommendations from LinkedIn profiles. The script implements web scraping methods, aided by established Python libraries, such as BeautifulSoup and requests, to navigate to the designated profile page and extract the necessary data from the recommendation segment. The script utilizes the requests library to send HTTP requests to LinkedIn, and extracts profiles of the individuals from the response. Afterward, the script goes through each profile and calls the singleScan subroutine to extract data, which is then saved in a record. The data retrieved by the script includes the names and genders of the recommender and receiver, their respective positions, and the content of the recommendation.

2.5 Textual Analysis

The present study employed textual analysis to examine language variations in recommendations based on gender. My research methodology aligns with established protocols proposed by prior scholars in the field.

To evaluate the degree of consideration and thoroughness evident in recommendations, I employed a word-counting approach introduced by Trix and Psenka (2003) to determine the average length of recommendations. This analysis was facilitated using Linguistic Inquiry Word Count (LIWC) software, a validated text analysis program. Given that longer recommendations are often considered more influential and meaningful in decision-making, assessing the average length of recommendations offers insights into the quality and depth of the relationship between the recommender and the recommended individual.

To identify gendered language in recommendations, gender bias scores were computed for each recommendation given and received using a default list of gendered words (see Appendix A) (Rudman 2001, Schmader 2008, Judith et. al, 2019). I opted to use a default list of gendered words to account for the possibility that dictionaries created within the dimensions of prior research may not be generalizable across various fields, industries, and populations.

I applied natural language processing techniques, specifically a Naive Bayes classifier¹⁵ and a Binary Bag-of-Words Model¹⁶, to categorize recommendation letters. Drawing on prior research by Ross (2017) and Correll (2020), I developed a set of categories, which included compassion, standout vs. doubt-raising language, communal and group-oriented language, agentic and managerial language, neutral valence, and positive valence (see Appendix B for more information). Each category was associated with a specific set of words and phrases, and I used these features to classify the recommendations using the binary bag-of-words representation. To execute my methodology, I constructed categories with corresponding word and phrase dictionaries, and then transformed the data into a binary bag-of-words representation. I trained a Naive Bayes classifier on this representation using the scikit-learn Python framework to classify the LinkedIn recommendations.

RESULTS

3.1 Gender-based Differences in Sources of Recommendations

Are women and men equally likely to give and receive recommendations? I investigate the potential gender differences in the frequency and nature of recommendations given and received by recent university graduates across diverse industries. Table 2 presents the distribution of recommendations among 1,000 participants, stratified by gender. Of the total participants, 74 were recommended, with a slightly (but statistically insignificant) higher proportion of women (8.4%) receiving recommendations compared to men (6.4%) ($t=1.2$, $p=.23$). Additionally, on average, women received more recommendations ($M=1.76$) than men ($M=1.69$). Further, while women were similarly likely to receive recommendations from women and men (53% of their recommendations came from other women, which indicates a statistically insignificant difference in recommender gender: $t=0.58$, $p=.57$), men were much more likely to receive recommendations from other men (81% of their recommendations came from other men; $t=5.9$, $p<.000$). Thus, I find partial support for Hypothesis 1b: while men were significantly more likely to receive same-gender recommendations, women were not.

A small proportion of participants gave recommendations, with men being slightly (but statistically insignificantly) more likely to give recommendations than women (7.2% versus 5.4%; $t=1.17$, $p=.24$). Further, among those participants that gave recommendations, both women and men were somewhat (but statistically insignificantly) more likely to give recommendations to people of their same gender (53% for women and 59% for men; $t=.52$ and $p=.6$ for women, and $t=1.39$ and $p=.17$ for men). Thus, I do not find robust evidence for Hypothesis 1a: while descriptive women and men do provide more recommendations to people of the same gender, the difference is not statistically significant.

3.2 Gender-based Differences in Content of Recommendations

To examine gender-based differences in recommendations, I conduct a content analysis of recommendation letters. I begin by identifying and categorizing the presence of female- and male-associated words and performance-related traits. Additionally, I measure the length of recommendations to determine gender-based differences in

level of detail provided. Taken together, I describe levels of qualitative difference but do not establish statistical differences.

Tables 3 and 4 present the gender bias scores associated with recommendations received and given, respectively. The scores are based on the frequency of female and male-associated words in the content of the recommendations. A negative score indicates a greater male bias, a positive score indicates a greater female bias, and a score of zero denotes a neutral bias.

To begin with, I observed that recommendations between men and women were generally neutral with regards to language. For women receiving recommendations, the average bias score was -0.6, while for men it was 1.3158. This suggests a slightly higher (but insignificant) male bias in recommendations for women and a slightly higher female bias for men.

Interestingly, when it came to giving recommendations, both men and women showed no gender bias, evidenced by a median bias score of zero. However, when women recommended men and the reverse was also true, the median bias scores were 20 and 17 respectively, indicating a moderate female bias.

Tables 5 and 6 display descriptive statistics for the average length of received and given recommendations, respectively. It appears that women tend to receive slightly longer recommendations than men. Specifically, the median length of recommendations received by women is slightly higher than men (83 vs. 73 words). Further, the median length of recommendations given by women is 98 words, while men receive a median of 83 words. Although these differences are small, they indicate a slight preference for long recommendations among women.

Finally, Tables 7 and 8 examine gendered-language use in performance-trait categories of recommendations. Together, these findings illustrate minimal differences between recommendations given to men and women regarding general performance. For instance, women received praise in a slightly higher percentage of recommendations than men (54.7% vs. 45.6%). Further, men were evaluated based on their maturity in the field in a slightly higher percentage of recommendations than women (21.1% vs. 18.7%). Indeed, the differences were small, and both men and women were generally described similarly in their recommendations.

Inconsistent with theorization, men and women were described using communal language in a similar percentage of recommendations (8.8% vs. 10.7%). Perhaps more surprisingly, though also consistent with prior research, men were described

using agentic language in a significantly higher percentage of recommendations than women (61.4% vs. 4%). Thus, I find partial support for Hypothesis 2: while men were described as more agentic and meaningfully managerial than women in their recommendations, both were described as communal and group-oriented at similar numbers of recommendations.

LIMITATIONS AND FUTURE RESEARCH

This study has several limitations that future research should examine. First, while LinkedIn offers a convenient means of collecting recommendations, it is essential to recognize that this platform deviates from traditional hiring processes that rely on paper-based job applications and in-person interviews. This departure from established norms may introduce a bias in the data, as LinkedIn users may not be representative of the broader population of graduates seeking employment. Additionally, this study was restricted to graduates who had created a LinkedIn profile, which excludes those who have not adopted this platform as part of their professional networking and job-seeking strategies.

Second, I employed a Naive Bayes algorithm for binary classification of recommendation content. While this approach has been widely used in text classification tasks, it is not without limitations. One of the main assumptions of Naive Bayes is that the features are conditionally independent, which may not hold true in all cases. This could lead to overestimation or underestimation of the probabilities and ultimately impact the accuracy of the classifier. Another limitation of the Naive Bayes algorithm is its susceptibility to outliers or noisy data, which could further affect the quality of the classification.

Third, this study has limitations with regards to the sample size and geographic scope of data. Indeed, drawing from university graduates as a sample population provides insight into a wide range of industries. However, I focused exclusively on graduates from a single university, which may limit the generalizability of our findings to other contexts. As an example, the University of Michigan is a highly-ranked institution, and this may have influenced the occupational outcomes of its graduates. Specifically, graduates from this university may be more likely to enter high-status occupations, which could restrict the generalizability of this study's findings to graduates from a university with different rankings or reputations.

Despite these limitations, there is much potential for future research in this area. One promising avenue for future investigation would be to compare the efficacy of LinkedIn-based recommendations and referrals to those obtained through traditional hiring processes. This could involve a large-scale, multi-site study that compares

the outcomes of job candidates who are recommended through LinkedIn to those who are selected through more traditional channels.

Another area for future research is to explore the potential for using natural language processing and machine learning techniques to analyze the content of LinkedIn recommendations and referrals. This could help to identify patterns and trends in the way that endorsers and referrers communicate their opinions, and could shed light on the factors that are most influential in driving hiring decisions.

Finally, it would be interesting to expand the scope of our research beyond the geographic regions that I studied. By examining LinkedIn data from a broader geographic scope, we could gain a more comprehensive understanding of the factors that drive hiring decisions and the role that social networking platforms like LinkedIn play in this process.

DISCUSSION AND CONCLUSION

The present study examined gender differences in giving and receiving recommendations on LinkedIn, a widely-used social networking site. I analyzed recommendation data from a sample of 1,000 university graduates using textual analysis and classification techniques. The results indicate several important gendered patterns in recommendation behavior and content, with implications for gender-based networks and unconscious biases.

First, while women were similarly likely to receive recommendations from men and women, men were more likely to receive recommendations from men than women. This suggests that gender homophily may exist in recommendation behavior and that men may benefit from gender-based networks that women may not have access to.

Second, women were found to receive and give slightly longer recommendations than men, which may reflect the gendered expectation of women to be nurturing and communicative. However, the differences in recommendation length were observably small, which may suggest insignificant implications for the difference in level of detail in recommendations between men and women.

Even so, gender may still frame recommendation content, affecting how men and women are viewed, valued, or both. Men were described as more agentic and meaningfully managerial than women, with minimal differences in other categories. This finding is consistent with prior research on gendered language in recommendations, which have characterized men as more assertive and in control than women.

Finally, the findings of this study highlight the potential existence of gender differences in recommendation patterns, including the likelihood, content, and length of recommendations. As such, this research has significant implications for both individuals using LinkedIn and organizations and hiring managers evaluating job candidates. It is crucial to increase awareness of gendered networking patterns and the presence of unconscious biases that can impact individuals' language choices and recommendations. Further research is needed to investigate the impact of gender on other aspects of professional networking and career advancement.

Chapter 6

TABLES

Table 1: **Descriptive Statistics of Top 10 Most Popular Majors**

Major	Women	Men	Total	% Women	% Men
Psychology	441	128	569	77.5%	22.5%
Computer Science	101	437	538	18.8%	81.2%
Business Administration	200	282	482	41.5%	58.5%
Mechanical Engineering	92	338	480	21.2%	78.8%
Political Science	158	175	333	47.4%	52.6%
Economics	70	229	299	23.4%	76.6%
Industrial and Operations Engineering	86	133	219	39.3%	60.7%
Mathematics	104	128	232	44.8%	55.2%
Neuroscience	125	84	209	59.8%	40.2%
Biopsychology, Cognition, and Neuroscience	140	61	201	69.7%	30.3%

Table 2: **Participant Gender Distribution and Recommendation Status**

	Women	Men
# participants	500	500
% recommended	8.40%	6.40%
Avg. recommendations received (among those who receive recommendations)	1.79	1.69
% recommendations from same gender	53%	81%
% who gave recommendations	5.40%	7.20%
Avg. recommendations given (among those who gave recommendations)	2.15	1.75
% recommendations given to same gender	53%	59%

Table 3: **Gender Bias Scores for Recommendations Received**

	Median	Mean	Standard Deviation
Women	0	-0.6	53.72
Men	0	1.3158	55.7722

Table 4: **Gender Bias Scores for Recommendations Given**

	Median	Mean	Standard Deviation
Women	0	8.5455	53.3037
Women to men	20	13.3103	54.9026
Women to women	0	3.5862	54.1799
Men	0	16.0476	51.2711
Men to women	17	15.6786	54.4168
Men to men	0	16.3429	49.4148

Table 5: Average length of recommendations received

	Median	Mean	Standard Deviation
Women	83	94.9867	53.3118
Men	73	76.4211	41.6211

Table 6: Average length of recommendations given

	Median	Mean	Standard Deviation
Women	98	113.1552	56.6946
Women to men	93	100.9310	44.2831
Women to women	109	125.3793	65.3668
Men	83	99.6191	81.4918
Men to women	85	118.5714	110.0429
Men to men	83	84.4571	44.2847

Table 7: Gendered Language use in Performance-trait Categories

Category	Men		Women	
	Count	Percentage	Count	Percentage
General performance: positive valence				
<i>Praise</i>	31	45.6%	41	54.7%
Vague praise	3	5.3%	3	4%
<i>Hardworker/dedicated/grindstone language</i>	11	19.3%	20	26.7%
<i>Promotion</i>	1	1.8%	6	8%
<i>Mentorship</i>	5	8.8%	6	8%
<i>Experience</i>	12	21.1%	14	18.7%
Maturity/experience in field	2	3.5%	3	4%
General performance: neutral valence				
<i>Feedback from someone other than reviewer</i>	18	31.5%	17	22.7%
<i>Evaluator's tone</i>	0	0%	8	10.7%
<i>Shout-out/thanking someone</i>	7	12.3%	10	13.3%
<i>Service work</i>	1	1.8%	11	14.7%
Networking	4	7%	16	21.3%
<i>Description of the work</i>	32	56.1%	31	41.3%
State of the project	2	3.5%	3	4%
Status/visibility of project	0	0%	10	13.3%
Agentic and managerial language				
<i>Agentic language</i>	35	61.4%	3	4%
<i>Managerial skills</i>	13	22.8%	9	12%
Leader of/in teams	19	33.3%	16	21.3%
Attributes of others' accomplishments to their leadership	4	7%	2	2.7%
Directly mentions promotion opportunities	0	0%	2	2.7%
Communal and group-oriented language				
<i>Communal</i>	5	8.8%	8	10.7%
Nuturing or communal language	1	1.8%	4	5.3%
Team player/collaborator	21	36.8%	22	29.3%
Team accomplishments	4	7%	3	4%
Leader in organizational skill	2	3.5%	8	10.7%
Standout language vs. doubt-raising language				
Visionary	2	3.5%	2	2.7%
<i>Genius/game-changer language</i>	1	1.8%	3	4%
<i>Ideal worker</i>	5	8.8%	12	16%
<i>Improvement</i>	15	26.3%	14	18.7%
Future changes	1	1.8%	7	9.3%
Non-technical improvement	3	5.3%	4	5.3%
Vague improvement	2	3.5%	4	5.3%
Vague call for future improvement	1	1.8%	2	2.7%
Specific suggestion for future improvement	1	1.8%	2	2.7%
Needs to develop a technical skill	2	3.51%	3	4.0%
Compassion				
<i>Generous</i>	2	3.51%	0	0%
<i>Helpful</i>	2	3.51%	6	8%
<i>Kind</i>	3	5.26%	4	5.33%
<i>Patient</i>	3	5.26%	0	0%
<i>Respectful</i>	2	3.51%	1	1.33%
Compassionate behavior	3	5.26%	3	4%
Demonstrates compassion towards colleagues	6	10.53%	3	4%
Shows empathy towards others	8	14.04%	7	9.33%

Table 8: Representative Quotations Demonstrating Key Types of Performance-traits

Type of Evaluation	Women	Men
General performance: positive valence	<p>“[Name] was one of the best managers I have ever worked with. She brings energy and accountability to any team. [Name] always picked up slack and assumed responsibility under stress. She also managed with a strong growth mindset, making her an invaluable asset to any organization.”</p>	<p>“[Name] is the most intelligent, thoughtful, and creative product leader I’ve ever met.”</p>
General performance: neutral valence	<p>“[Name] is hard-working and results-oriented. When given an assignment, she starts immediately, works diligently, and produces excellent results. She expressed an interest in tax policy and economic issues when she began her internship and so I gave her a variety of tax assignments involving the congressional tax writing committees, including covering hearings, analyzing testimony, and writing reports which were sent to clients.”</p>	<p>“For the past year I have had the pleasure of working with [Name] developing and designing multiple marketing platforms for over 25 frontend engineers and more than 170 million unique page views.”</p>
Agentic and managerial language	<p>“[Name] is one of the best managers I’ve ever worked for. She has an innate ability to work successfully with people at all levels of the organization. [Name] possesses a unique combination of executive oversight, risk assessment, analytical thinking, technical understanding and the willingness and ability to get in the trenches and work through complex projects and detailed plans.”</p>	<p>“[Name] is a leader and valued member of our product team. Not only is he a talented engineer, but he cares deeply about experience design and how the products impact the users and continually push us to make things better for them in faster, more scalable ways. He has spearheaded our component library initiatives and helps mentor the team in the craft. We are lucky to have him.”</p>
Communal and group-oriented language	<p>“I have had the opportunity to watch [Name] working in two very different dimensions. At the micro level, she has been able to untangle projects that were stuck for a long time through a deep understanding of the project objectives and a clear separation of the various underlying issues. At the macro level, she has proven time and time again to be able to adapt to changing business and product targets without losing sight of the user’s goals while keeping the team together and functional even in times of uncertainty. In both circumstances, [Name] has been able to achieve a successful outcome while keeping the team’s morale high and with a remarkable empathy and good humor.”</p>	<p>“[Name] is a great asset to The Solar Car Team. His impeccable work ethic inspires fellow Business Division members to do their very best. His drive and determination create a very accepting and productive work environment. I can honestly say that working with Pavan has not only been a pleasure but an educational experience that I will cherish for some time. [Name] is a gift to this team.”</p>

Table 8: Representative Quotations Demonstrating Key Types of Performance-traits (Continued)

Type of Evaluation	Women	Men
Standout language	“[Name]’s combination of confidence, professionalism, and ability to execute makes her a unicorn!”	“[Name] was always a standout intern. His ability to balance the need for analytical prowess with big picture thinking was particularly helpful. He is able to take a step back to look at the problem at hand and draw from his vast and diverse knowledge to enable him to come to an insightful solution. [Name] was new to our industry, but that didn’t stop him from exceeding expectations.”
Questioning language / vague improvement	“[Name] interned at CityGrid Media when I was the CEO of the company. She worked with the communications team to help develop a marketing plan for the new rewards program: CitySearch Local Rewards. [Name] was asked to create a marketing proposal for the program. She conducted thorough research, utilized analytics, asked meaningful questions, and listened carefully as she adjusted her proposal. She presented her proposal to the marketing team, and we were all impressed by her initiative, her drive and her dedication to the project. As the youngest member of the company at the time, [Name] brought in fresh, innovative ideas to help CityGrid. [Name] is a very hard worker, and I believe that she would be an asset to any organization.”	“[Name] and fellow interns did a great job for Rapid-Line during the Summer of 2015. He learned a great deal about being a practicing engineer and generated great value for us in the process.”
Compassion	“I have had the privilege of knowing [Name] for 11 years. [Name] and I were great friends since middle school. I have gained knowledge and courage from this human being. Solely with her personality and evincing compassion, she has shown me the skills and schedule I must obtain (in my private and work life) to get offered a promotion within my job”	"I was able to work with [Name] on a personal level as I was his builder for his home. I can say with all sincerity that [Name] was a pleasure to work for. He is honest, fair, personable, a problem solver, and extremely genuine. I would work for him again in a heartbeat!"

Chapter 7

NOTES

1. Gender homophily refers to the tendency for individuals to form social ties with others who share their gender.
2. In academic literature, high-status occupations are typically defined as those that require a high level of education or specialized training, involve substantial levels of decision-making authority, and offer significant financial rewards and prestige within society. Examples of high-status occupations include doctors, lawyers, executives, and other professional and managerial roles. The use of this term is intended to capture the social and economic significance of certain occupations within a given society.
3. Communal qualities refer to traits that are associated with kindness, empathy, and interpersonal skills. Examples of communal qualities include being supportive, nurturing, and helpful towards others. These qualities are often seen as more stereotypically feminine and may be valued less than agentic qualities, such as assertiveness, ambition, and leadership, which are associated with stereotypical masculinity.
4. Social network literature refers to studies that analyze social networks, which are systems of individuals or organizations that are connected through various social relationships, such as friendships, communication, or work ties.
5. The extent to which research findings can be generalized to settings beyond the study's specific context.
6. Gendered language refers to language that reflects and may reinforce gender stereotypes.
7. Agentic refers to personality traits that are typically associated with stereotypically masculine characteristics, such as assertiveness, independence, and confidence. In contrast, communal refers to personality traits that are typically associated with stereotypically feminine characteristics, such as warmth, helpfulness, and cooperation.

8. Ability words describe a person's skills, such as "competent," "skilled," or "talented." Grindstone words, on the other hand, refer to traits associated with hard work, such as "diligent," "hardworking," or "persistent."
9. Standout adjectives refer to descriptive words that are unusual or particularly impressive, such as "brilliant," "exceptional," or "outstanding." Authentic words are those that reflect a genuine and honest assessment of an applicant's abilities, rather than generic or overly positive language that may be perceived as insincere.
10. Letters of recommendation that are deemed outstanding and provide strong endorsements of the candidate's qualifications and potential for success. These letters are typically written by individuals who have worked closely with the candidate and can provide detailed and positive assessment of their skills, abilities, and character.
11. Referrals and recommendations are often used interchangeably, but there is a subtle difference between the two terms. A referral typically refers to the act of referring a candidate to a job opening or an opportunity. In contrast, a recommendation is a formal assessment or evaluation of a candidate's skills, qualifications, and suitability for a specific job or position. While a referral can lead to a recommendation, not all referrals result in recommendations.
12. In total; of the original 500 men and 500 women searched, I could not locate 24% of the LinkedIn profiles. Among the profiles I could not find, 57% belonged to women and 43% to men. To meet our target of 500 men and 500 women, I expanded the scope of the search by randomly selecting additional individuals from the original sampling frame. I continued to randomly select individuals from the same major and gender categories until I reached the desired number for each major.
13. HTTP is the underlying protocol used by the web. When a web client, such as a web browser, requests a web page from a web server, it sends an HTTP request message to the server. The server responds with an HTTP response message, which contains the requested content, such as an HTML page or an image.
14. In web terminology, being "challenged" refers to the practice of websites presenting users with a CAPTCHA or other test to verify that they are not an

automated bot. This is often done to prevent spam, fraud, and other malicious activities. CAPTCHA is a type of challenge-response test used to determine whether or not a user is human.

15. The Naive Bayes classifier is a probabilistic classifier that assumes independence between features and is based on Bayes' theorem. Despite its naive assumptions, it has been shown to be effective in many natural language processing tasks.
16. The Binary Bag-of-Words Model is a widely used text representation that disregards word order and structure and presents the document as a vector of word frequencies, capturing the presence or absence of particular words or phrases. This model is particularly useful for sentiment analysis and text classification, especially with large datasets. I decided to use this model to efficiently classify a large volume of recommendation letters.

Chapter 8

REFERENCES

- Aamodt, M. G., Bryan, D. A., & Whitcomb, A. J. (1993). Predicting Performance with Letters of Recommendation. *Public Personnel Management*, **22**(1), 81–90. <https://doi.org/10.1177/009102609302200106>
- Abel, M. D., Burger, R., & Piraino, P. (2020). The Value of Reference Letters: Experimental Evidence from South Africa. *American Economic Journal: Applied Economics*, **12**(3), 40–71. <https://doi.org/10.1257/app.20180666>
- Afridi, Farzana & Dhillon, Amrita. (2022). Social Networks and the Labour Market. *IZA Discussion Papers*, **15774**. <https://ideas.repec.org/p/iza/izadps/dp15774.html>
- Beugnot, J., & Peterle, E. (2020). Gender bias in job referrals: An experimental test. *Journal of Economic Psychology*, **76**, 102209. <https://doi.org/10.1016/j.joep.2019.102209>
- Bureau of National Affairs. (1988). Recruiting and selection procedures (Personnel Policies Forum, Survey No. 146). Washington, DC: Author.
- Brown, M., Setren, E., & Topa, G. (2016). Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System. *Journal of Labor Economics*, **34**(1), 161–209. <https://doi.org/10.1086/682338>
- Burks, S. V., Cowgill, B., Hoffman, M., & Housman, M. G. (2015). The Value of Hiring through Employee Referrals*. *Quarterly Journal of Economics*, **130**(2), 805–839. <https://doi.org/10.1093/qje/qjv010>
- Correll, S. J., Weisshaar, K., Wynn, A. T., & Wehner, J. D. (2020). Inside the Black Box of Organizational Life: The Gendered Language of Performance Assessment. *American Sociological Review*, **85**(6), 1022–1050. <https://doi.org/10.1177/0003122420962080>
- De Werfhorst, V., & Herman, G. (2017). Gender Segregation across Fields of Study in Post-Secondary Education: Trends and Social Differentials. *European Sociological Review*, **33**(3), 449–464. <https://doi.org/10.1093/esr/jcx040>

- Dutt, K., Pfaff, D. L., Bernstein, A. F., Dillard, J. P., & Block, C. J. (2016). Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nature Geoscience*, *9*(11), 805
- Fernandez, R. M., & Sosa, M. L. (2005). Gendering the Job: Networks and Recruitment at a Call Center. *American Journal of Sociology*, *111*(3), 859–904. <https://doi.org/10.1086/497257>
- French, J. C., Zolin, S. J., Lampert, E. J., Aiello, A., Bencsath, K. P., Ritter, K. A., Strong, A. W., Lipman, J. M., Valente, M., & Prabhu, A. S. (2019). Gender and Letters of Recommendation: A Linguistic Comparison of the Impact of Gender on General Surgery Residency Applicants. *Journal of Surgical Education*, *76*(4), 899–905. <https://doi.org/10.1016/j.jsurg.2018.12.007>
- Friedman, T., & Williams, E. (1982). Current use of tests for employment. In A.K. Wigdor and W.R. Garner (Eds.), *Ability testing: Uses, consequences and controversies* (pp. 99–169). Washington, D.C.: National Academy Press.
- Kwiek, M. (2021). The prestige economy of higher education journals: a quantitative approach. *Higher Education*, *81*(3), 493–519. <https://doi.org/10.1007/s10734-020-00553-y>
- Levy-Leboyer, C. (1994). Selection and assessment in Europe. In H. C. Triandis, M. D. Dunnette, and L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 4, pp. 173–190). Palo Alto, CA: Consulting Psychologists.
- Lin, F., Oh, S. J., Gordon, L. K., Pineles, S. L., Rosenberg, J. B., & Tsui, I. (2019). Gender-based differences in letters of recommendation written for ophthalmology residency applicants. *BMC Medical Education*, *19*(1). <https://doi.org/10.1186/s12909-019-1910-6>
- Loeppky, C., Babenko, O., & Ross, S. (2017). Examining gender bias in the feedback shared with family medicine residents. *Education for Primary Care*, *28*(6), 319–324. <https://doi.org/10.1080/14739879.2017.1362665>
- Madera, J. M., Hebl, M. R., Dial, H., Martin, R. C., & Valian, V. (2019a). Raising Doubt in Letters of Recommendation for Academia: Gender Differences and Their Impact. *Journal of Business and Psychology*, *34*(3), 287–303. <https://doi.org/10.1007/s10869-018-9541-1>

- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41), 16474–16479. <https://doi.org/10.1073/pnas.1211286109>
- Neugart, M., & Zaharieva, A. (2018). Social Networks, Promotions, and the Glass-Ceiling Effect. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3249419>
- Rudman, L. A., & Glick, P. (2001). Prescriptive Gender Stereotypes and Backlash Toward Agentic Women. *Journal of Social Issues*, 57(4), 743–762. <https://doi.org/10.1111/0022-4537.00239>
- Schmader, T., Whitehead, J., & Wysocki, V. H. (2007). A Linguistic Comparison of Letters of Recommendation for Male and Female Chemistry and Biochemistry Job Applicants. *Sex Roles*, 57(7–8), 509–514. <https://doi.org/10.1007/s11199-007-9291-4>
- Stewart-Williams, S., & Halsey, L. G. (2021). Men, women and STEM: Why the differences and what should be done? *European Journal of Personality*, 35(1), 3–39. <https://doi.org/10.1177/0890207020962326>
- Trix, F., & Psenka, C. E. (2003). Exploring the Color of Glass: Letters of Recommendation for Female and Male Medical Faculty. *Discourse & Society*, 14(2), 191–220. <https://doi.org/10.1177/0957926503014002277>
- Turrentine, F. E., Dreisbach, C., St Ivany, A., Hanks, J. B., & Schroen, A. T. (2019). Influence of Gender on Surgical Residency Applicants' Recommendation Letters. *Journal of the American College of Surgeons*, 228(4), 356-365e3. <https://doi.org/10.1016/j.jamcollsurg.2018.12.020>
- Zeltzer, D. (2020). Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap. *American Economic Journal: Applied Economics*, 12(2), 169–197. <https://doi.org/10.1257/app.20180201>
- Zhang, N., Blissett, S., Anderson, D. E., O'Sullivan, P. S., & Qasim, A. (2021). Race and Gender Bias in Internal Medicine Program Director Letters of Recommendation.

Journal of Graduate Medical Education, 13(3), 335–344. <https://doi.org/10.4300/jgme-d-20-00929.1>

APPENDIX

9.1 Appendix A

Female-associated words	Male-associated words
Hardworking	Excellent
Conscientious	Superb
Depend	Outstanding
Meticulous	Unique
Thorough	Exceptional
diligent	Unparalleled
dedicate	Best
careful	Most
reliable	Wonderful
effort	Terrific
assiduous	Fabulous
trust	magnificent
responsible	remarkable
methodical	extraodrinary
industrious	amazing
busy	supreme
workpersist	unmatched
organize	talent
disciplined	intellect
teach	smart
instruction	skill
educate	ability
train	genius
mentor	brilliant
supervise	bright
adviser	brain
counselor	aptitude
syllabus	gift
course	propensity
classs	innate
service	capacity
colleague	flair

citizen	knack
communicate	clever
lecture	expert
student	proficient
present	capable
rapport	adept
	able
	competent
	natural
	inherent
	instinct
	adroit
	creative
	insight
	analytical
	research
	data
	study
	studies
	experiment
	scholarship
	result
	test
	finding
	publication
	publish
	vital
	method
	science
	grant
	fund
	manuscript
	project
	journal
	theory
	discover
	contribution

9.2 Appendix B

Table 2: Performance Trait Descriptions and Examples

Performance Trait	Description	Example Language
General Performance: Positive Valence	Positive connotation and emphasizes the candidate's strengths and accomplishments.	outstanding performance, exceeded expectations, exceptional work, demonstrated outstanding skills, produced excellent results, very competent, exceptionally well-suited, impressive
General Performance: Neutral Valence	Does not have positive or negative connotations.	did the job competently, performed satisfactorily, did what was expected, met the requirements, fulfilled their responsibilities, was adequate, performed to standard
Standout vs. Doubt-raising language	Either emphasizes the candidate's strengths and accomplishments (standout language) or raises doubts about their abilities or potential (doubt-raising language).	Standout: outstanding, excellent, impressive, exceptional, truly remarkable. Doubt-raising: performed satisfactorily, somewhat competent, did not meet expectations
Communal and group-oriented language	Emphasizes the candidate's ability to work well with others and contribute to a team.	collaborated well with others, demonstrated teamwork skills, is a team player, was an asset to the group, demonstrated the ability to work well in a team environment
Agentic and Managerial language	Emphasizes the candidate's assertiveness, leadership, and ability to take charge.	demonstrated leadership skills, took charge of projects, excelled in managing teams, displayed strong decision-making skills, has a natural ability to lead, demonstrated the ability to manage complex tasks
Compassion	The extent to which the recommender portrays the candidate as empathetic, caring, and understanding towards others.	demonstrated empathy and understanding towards colleagues, showed a caring attitude toward clients, provided emotional support to team members, displayed a genuine concern for others, was very supportive and understanding