Investigating Gender Bias in STEM Job Advertisements

Malika Dikshit
Carnegie Mellon Univeristy Qatar
mdikshit@andrew.cmu.edu

Houda Bouamor
Carnegie Mellon Univeristy Qatar
hbouamor@andrew.cmu.edu

Nizar Habash
New York University, Abu Dhabi
nizar.habash@nyu.edu

Abstract

Gender inequality has been historically prevalent in academia, especially within the fields of Science, Technology, Engineering, and Mathematics (STEM). In this study, we propose to examine gender bias in academic job descriptions in the STEM fields. We go a step further than previous studies that merely identify individual words as masculine-coded and feminine-coded and delve into the contextual language used in academic job advertisements. We design a novel approach to detect gender biases in job descriptions using Natural Language Processing (NLP) techniques. Going beyond binary masculine-feminine stereotypes, we propose three big groups *types* to understand *gender* bias in the language of job descriptions, namely *agentic*, *balanced*, and *communal*. We cluster similar information in job descriptions into these three groups using contrastive learning and various clustering techniques. This research contributes to the field of gender bias detection by providing a novel approach and methodology for categorizing gender bias in job descriptions, which can aid more effective and targeted job advertisements that will be equally appealing across all genders.

1 Introduction

Academic institutions in recent decades have strived to launch several initiatives addressing diversity, equity, and inclusion in the hopes of making academic gender representation more equal. However, the problem of gender bias still persists in academia and widens particularly among STEM fields. Casad et al. (2020) report that at the top 50 research universities in the U.S., women hold only 31% of the tenured or tenure-track faculty positions. Cech et al. (2011) present that gender disparity in STEM persists because of two reasons: women leave STEM careers because they feel that their family plans will be hindered because of their professional lives, and due to low self-assessment of their skills in STEM's intellectual tasks. They introduce the concept of 'professional role confidence', and argue that women's lack of this confidence, compared to their male counterparts, reduces their likelihood of pursuing careers in engineering.

Men and women have been found to identify with different goals as core human motivations. Men often relate more with agentic goals, while women relate with communal goals (Bakan, 1996). Agentic goals display an affinity to one's status, achievement, and independence, along with speaking assertively, influencing others, and initiating tasks whereas communal goals showcase a drive to contribute to the community, connect, and share with others.

Gaucher et al. (2011) suggest that women may experience intimidation and barriers with job descriptions that are formulated using agentic language. They present a list of masculine-coded and feminine-coded words that represent agentic and communal traits, respectively, and show that women judge jobs with a lot of agentic language as less appealing than jobs containing communal language. Building on Gaucher et al. (2011)'s work, Matfield (2016) created Gender Decoder, a freely available tool that quantifies gender bias in job descriptions by counting the number of masculine and feminine-coded words in them. This simple dictionary look-up approach relies only on the frequency of gender-coded words in the job advertisement and does not consider the contextual meaning of the word or how it is used in the sentence.

Current studies that deal with gender bias in job descriptions also rely on this method of labeling job descriptions as gender-biased or gender-neutral. Often, singular words are simply labeled as masculine-coded or feminine-coded. This practice can reinforce traditional gender stereotypes. Instead, employing terms like *agentic*, *balanced*, and *communal* offers a more nuanced and inclusive approach to understanding language biases.

Proceedings of the 5th Workshop on Gender Bias in Natural Language Processing (GeBNLP), pages 179–189
August 16, 2024 ©2024 Association for Computational Linguistics

By categorizing job descriptions based on these dimensions, we move away from reinforcing gender norms and acknowledge the diverse ways in which individuals can express themselves and their abilities. The research question that we address in this work is centered around how we can employ more comprehensive criteria to determine gender bias in job advertisements, moving beyond simply identifying specific masculine-coded words to label an advertisement as biased towards masculinity?

This paper makes the following contributions to understanding gender bias in job descriptions:

1. A novel dataset of 6,031 academic STEM job descriptions compiled semi-automatically using a combination of manual and web-scraping techniques.
2. A novel methodology to label job descriptions as agentic, communal, or balanced based on their dense numerical vector representations (embeddings) obtained from sentence-level transformer models fine-tuned with contrastive learning techniques.
3. An in-depth analysis of the anatomy of job advertisements focusing on the distribution and positioning of agentic *vs.* communal language within the body of the ad.
4. A departure from the conventional practice of categorizing job descriptions as masculine-coded or feminine-coded, which may inadvertently perpetuate gender stereotypes. Instead, we adopt a more nuanced approach employing the neutral terminology of agentic, balanced, and communal. This shift aims to challenge traditional gender norms within the discourse surrounding gender bias in job descriptions.

2 Related Work

Several research studies have addressed the gender disparity prevalent in academic faculty positions. According to a study by Llorens et al. (2021), citing data from the Society for Neuroscience, there has been a notable increase in the proportion of women applicants to PhD programs in recent years. Specifically, the percentage of female applicants increased from 38% in 2000-2001 to 57% in 2016-2017, with a corresponding matriculation rate of 48% for women in the latter year. However, despite these gains in representation among applicants and matriculants, women only accounted for 30% of all faculty positions in PhD programs, indicating a significant disparity in gender representation at the faculty level. In STEM fields, although there has been a steady rise in the number of female candidates obtaining postgraduate degrees in recent years, the representation of women in faculty positions has remained largely unchanged (Casad et al., 2020).

Current studies related to gender in NLP have looked at gender bias in the context of large language models (Haim et al., 2024; del Arco et al., 2024) and presented that LLMs are biased unfavorable for females. In this work, we use NLP to assess gender bias at the beginning of the hiring cycle - in job descriptions. A significant contributing factor to the gender disparity in academia is the significant lack of gender diversity within applicant pools. The initial point of contact between academic employers and job seekers typically occurs through job postings. Research indicates that the content and language used in job postings play a crucial role in influencing an applicant's decision to apply for a particular position (Feldman et al., 2006). Gaucher et al. (2011) found that job descriptions in male-dominated fields tend to contain words associated with masculine stereotypes more frequently than those in female-dominated fields. They demonstrated that job advertisements featuring more agentic language were perceived to be more suitable for men, making these positions less appealing to women candidates.

Wan et al. (2023) draw inspiration from social science findings and propose *Language Agency* as a metric for gender bias evaluation in LLM-generated professional documents. They present,"Bias in language agency states that women are more likely to be described using communal adjectives in professional documents, such as delightful and compassionate, while men are more likely to be described using "agentic" adjectives, such as leader or exceptional". Through their findings, they reveal that ChatGPT generates reference letters with biased levels of language agency for male and female candidates. When describing female candidates, ChatGPT uses communal phrases such as "great to work with", "communicates well", and "kind". On the other hand, the model describes male candidates as being more agentic, using phrases such as "a standout in the industry" and "a true original". Through their study, Wan et al. (2023) demonstrate that there is a distinct difference in the way males and females are described in terms of language agency.

Different studies have referred to the concept of language agency to evaluate job descriptions as masculine or feminine coded. Oldford and Fiset (2021) have followed Gaucher et al. (2011)'s method of annotating job descriptions using a dictionary look-up approach. They focused on classifying finance internship job postings based on masculine and feminine words, as well as evaluating the text based on the percentage of adjectives and verbs that are either agentic (e.g., overcomes, confident, etc.) or communal (e.g., aided, loyal, etc.). Their finding revealed that women exhibit greater goal congruity, leading to enhanced motivation and a greater sense of fit when job postings are high in communal language and low in agentic language.

Tang et al. (2017) adopt the approaches used by Textio[1] and Unitive[2], two recruitment assistant services dedicated to promoting inclusivity in job advertisements, to detect gender bias in job descriptions. They observe and adapt the techniques of both these services and introduce two metrics: Gender Tone and Gender Target to assess gender bias in the advertisement. Gender Target follows Textio's method and calculates the occurrences of gendered words in the ad, with masculine and feminine terms offsetting each other. Meanwhile, Gender Tone assigns a weight to gendered words based on their specificity, with the cumulative weights reflecting the overall gender tone of the ad.

Most of the works that evaluate gender bias in job descriptions (Bernstein et al., 2022; Born and Taris, 2010; O'Brien et al., 2022; Frissen et al., 2022; Oldford and Fiset, 2021; Read et al., 2023; Sella et al., 2023; Zhu et al., 2021), rely on the frequency of individual gender-coded words to assess gender bias in job advertisements, neglecting to explore the contextual positioning of these words within the advertisements. This limitation is noteworthy as it overlooks the nuanced interplay between language and context in conveying bias or its absence.

For instance, consider the following two sentences, which illustrate the importance of analyzing the contextual meaning of gender-coded words to detect gender bias: "We want a *competitive* member to join our team" and "We offer *competitive* remuneration." In both sentences, the word *"competitive"* is used. However, the first sentence implies a competitive environment or culture, potentially favoring traits typically associated with agentic language. In contrast, the second sentence simply indicates that the compensation provided is competitive in the market, without implying any specific gender-related traits or preferences. These examples show that examining the context in which these words are employed allows for a more accurate assessment of whether gender bias is present and facilitates a clearer understanding of the intended message.

3 Methodology

This section provides an overview of the methodology employed to address the research questions outlined in Section 1. Figure 1 presents a summary of these steps.

3.1 Data Collection

In this work, we build a novel dataset of academic job descriptions centered on STEM subjects, addressing a notable gap in the current literature. Our dataset comprises 6,031 meticulously curated academic STEM job advertisements. The job advertisements were collected semi-automatically[3] from several academic job databases, including HigherEdJobs,[4] TimesHigherEducation,[5] Jobs.ac.uk,[6] AcademicJobsOnline,[7] and The Chronicle of Higher Education.[8] The job advertisements were collected from regions spanning the globe, encompassing the United States, Europe, Asia, and Australia. Figure 2 shows an example of an academic job advertisement in our dataset.

3.2 Data Cleaning and Preprocessing

3.2.1 Standardization

We wanted to avoid clustering job advertisements based on the universities or titles advertised and focus on possible gender bias in the text, so we decided to replace the names of universities and academic positions with standard tokens, making these uniform across all the job advertisements. To ensure uniformity in job descriptions, we replaced named entities (universities, organizations) and academic job titles with standardized tokens <ORG> and <TITLE> respectively.

[1] https://textio.com/products/recruiting
[2] https://unitive.org/

[3] Using Python's beautifulsoup library
[4] https://www.higheredjobs.com/faculty/
[5] https://www.timeshighereducation.com/unijobs/
[6] http://Jobs.ac.uk
[7] https://academicjobsonline.org/
[8] https://jobs.chronicle.com/

Figure 1: Overview of the methodology to analyze gender bias in job advertisements

Assistant Professor of Physiology - Texas A&M International University - Texas A&M International University (TAMIU) is a comprehensive regional university and part of The Texas A&M University System. Poised at the Gateway to Mexico and serving as the cultural and intellectual hub of a vibrant multilingual and multicultural community, it is also a designated Hispanic-serving institution (HSI). Qualified applicants should possess a Ph.D. or equivalent terminal degree in physiology or a related field. Candidates should demonstrate a strong commitment to teaching, scholarly research, and service to the university and community. Additionally, candidates with experience in mentoring and supporting underrepresented students are highly encouraged to apply.

Figure 2: Example of a job advertisement

3.2.2 Text Preprocessing

We employed a range of text preprocessing techniques to refine the job descriptions for subsequent analysis, which included:

- Removing HTML tags
- Replacing hyperlinks with a placeholder term ("LINK")
- Converting HTML entities such as " " and "&" to their corresponding characters
- Substituting currency symbols like "$" with their corresponding term (“$” to "dollar”)
- Eliminating numerical values

Then, the preprocessed job descriptions were tokenized into individual sentences to allow a more fine-grained approach to identifying gender bias within lengthy job descriptions. Furthermore, we filtered out sentences that indicated technical skills by examining the presence of subject names and technical terms commonly associated with STEM disciplines that we observed in the dataset.[9]

[9]Programming, Physics, Bio*, Chemistry, Mechanic*, Electronic*, Volcanology, Math*, Statistics, Crystallography, Spectroscopy, Engineering, Electrochemistry, Machine, Geolog*, Robot*, Stata, Python, C++, Lab*, Software, Unix, Linux, Java, Python.

3.3 Contrastive Learning for Sentence Level Representation

We get the embeddings of the sentences extracted from the job descriptions using the SentenceTransformers framework (Reimers and Gurevych, 2019). The specific SentenceTransformer model used to obtain the embeddings was all-mpnet-base-v2, which was fine-tuned by its authors using a contrastive learning objective. The model was trained during 100k steps using a batch size of 1024. The sequence length was limited to 128 tokens, and the AdamW optimizer was used with a 2e-5 learning rate. For the fine-tuning, the cosine similarity was computed from each possible sentence pair from the batch, and then the cross entropy loss was applied by comparing with true pairs.

In our context, the SentenceTransformers model generates similar embeddings for sentences with similar meanings or contexts and dissimilar embeddings for sentences with different meanings or contexts. Figure 3 presents a simple diagram of how SentenceTransformers models fine-tune embeddings using a contrastive learning objective.

The embedding of the following sentences *“strong analytical, technical and problem-solving skills”* and *“strong drive, motivation, and ambition, with the capacity to deliver on challenging tasks and to meet deadlines individually and as part of a team”* displaying agentic language are similar with a cosine similarity score of 0.72. While the similarity between the ones for *“strong analytical, technical and problem-solving skills”* which represents agentic language, and *“strong willingness to mentor and guide undergraduate and graduate students”* reflecting communal traits is smaller with a cosine-similarity score of 0.41. This highlights the model’s capability to capture subtle linguistic cues and biases.

3.4 Sentence Level Labeling

Once the embeddings of the sentences were obtained, we applied the K-means clustering algorithm with k=3 to group them into three main clusters. Our K-means model utilizes these embed-

Figure 3: Contrastive learning used to fine-tune sparse embedding

dings as input and assigns a cluster label to each embedding, denoted as agentic, communal, or balanced. To ensure the reproducibility of results, we employed a random seed of 42. Sentences representing agentic traits, such as assertiveness, ambition, and self-reliance, are likely to end up within clusters characterized by shared linguistic patterns and thematic content. Conversely, communal sentiments, emphasizing collaboration, empathy, and inclusivity, may converge in distinct clusters reflecting their unique semantic profiles. Additionally, neutral or balanced sentences, which exhibit a combination of agentic and communal traits or lack strong alignment with either category, may also be identified and clustered accordingly.

3.5 Job Level Labeling

Our primary objective is to categorize an entire job description as either agentic, communal, or balanced, rather than focusing solely on individual sentences. To achieve this, we explore 2 distinct techniques (T1-T2).

Technique 1 (T1): Dictionary Look-up Method For every job ad, we count the number of masculine-coded and feminine-coded words as defined by Gaucher et al. (2011). If a job advertisement contains more masculine-coded words than feminine-coded ones, it is labeled as agentic. Conversely, if it contains equal masculine and feminine-coded words, it is labeled as balanced. Finally, if it has fewer masculine-coded words than feminine-coded ones, it is designated as communal.

Technique 2 (T2): Embedding-based Method In order to transition from sentence-level to job-level labeling, we compute the average embedding of each job description from the sentence-level embeddings. We then use the k-means clustering algorithm to assign each job advertisement to one of the agentic, balanced, or communal clusters.

4 Evaluation, Results, and Analysis

In this section, we present an analysis and visualization of the clustering model, evaluate its performance, and examine sentence-level label distributions within job advertisements.

4.1 Cluster Model Analysis via Visualization

The clusters were visualized using two dimensionality reduction techniques, namely: Principal Component Analysis (Maćkiewicz and Ratajczak, 1993) and t-distributed Stochastic Neighbor Embedding (van der Maaten and Hinton, 2008). Figure 4 shows that the PCA visualization reveals a notable degree of separability between the three clusters, indicating discernible patterns or structures within each group. Similarly, the t-SNE visualization demonstrates a distinction among the three clusters. However, it's worth noting that the separations are not entirely distinct, particularly given that the sentences originate from the same domain (job descriptions).

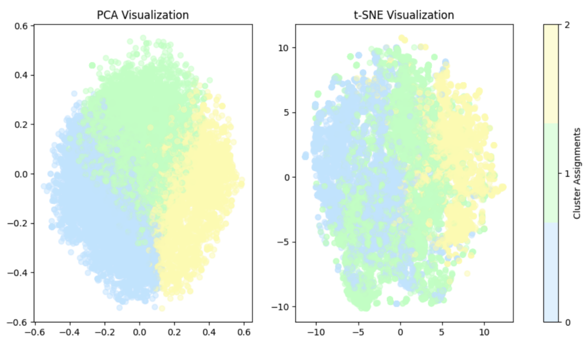

Figure 4: PCA and t-SNE visualizations of clusters

4.2 Cluster Model Evaluation

To evaluate the performance of our clustering model, we measured the cluster cohesion and separation. We used two metrics: The Davies-Bouldin Score (DBS) and the Calinski-Harabasz Index (CHI) (Pedregosa et al., 2011). We obtained a DBS of 4.03 suggesting that while there is some degree of clustering present, it is not optimal. This suggests potential overlaps or inconsistencies within the clusters. Such findings were anticipated given that the clustering was conducted on sentences originating from the same domain. On the other hand, we obtained a CHI score of 787.02, indicating strong clustering with clear separation.

Word Coding	Cluster 0	Cluster 1	Cluster 2
Masculine	**5.38%**	2.82%	1.55%
Feminine	1.26%	2.20%	**3.81%**

Table 1: Gender-coded word distributions

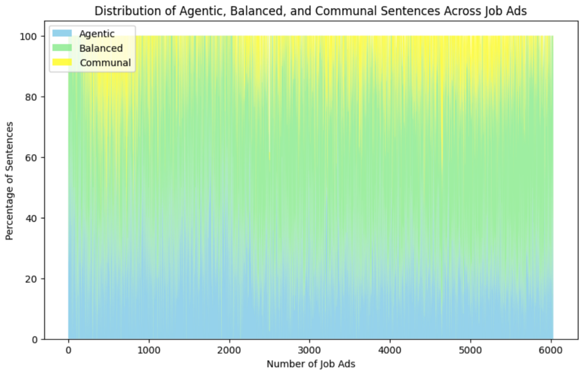

Figure 5: Distribution of agentic, balanced, and communal sentences across job ads

4.3 Cluster Naming

Upon obtaining the three clusters, we discerned a notable disparity in the distribution of Gaucher et al. (2011)'s gender-coded words across clusters. Using the statistics in Table 1, we mapped clusters 0, 1 and 2, to our definitions of *agentic*, *balanced*, and *communal* clusters, respectively. Cluster *agentic* has the highest concentration of masculine-coded words, followed by cluster *balanced* and cluster *communal*, in descending order. Cluster communal has the highest concentration of feminine-coded words suggesting a lesser emphasis on language traditionally linked with masculinity. Cluster *balanced* falls in between the other two clusters regarding the presence of gender-coded words.

4.4 Sentence-Level Label Distributions

Figure 5 shows the distribution of labeled sentences across job ads in our dataset. Communal sentences tend to make up the smallest proportion, with most job ads containing less than 20% of such sentences. Agentic sentences appear to account for roughly 20% to 40% of sentences. Balanced sentences are the most prevalent, commonly constituting over 40% to 60% of total sentences. Figure 6 provides representative examples of sentences from each of the identified clusters. Notably, the prevalence of balanced sentences within job ads suggests that a significant portion of the text is focused on conveying domain-specific responsibilities and technical requisites. Agentic sentences focus more on agentic personality traits such as independence, self-motivation, and assertiveness. Communal sentences tend to focus more on skills directed at contributing to society and the environment.

Agentic
*Are you an **exceptional** candidate?*
*ability to work in a **fast-paced** environment.*
*other desirable attributes: **self-motivated, organized, meticulous, efficient, and flexible.***

Balanced
Australian National University - Senior Lecturer in biological chemistry to further expand our capabilities, we are seeking candidates with expertise in protein chemistry, structural biology, biochemistry, biocatalysis, biophysics and/or protein engineering
We are seeking for a motivated post-doctoral fellow to work on funded research project aimed at deciphering the roles of moap- in cellular senescence and ageing-associated disorders in liver.
Skills: ph.d. degree in organization biology or other related fields is preferred.

Communal
*The school of computing, engineering & organization holds a silver athena swan award in recognition of our **commitment to advancing gender equality.***
*We are **committed** to building and maintaining a **fair and inclusive** working environment and we would be happy to discuss arrangements for **flexible and/or blended working.***
***Ability to mentor** undergraduate, master's and PhD students*

Figure 6: Examples of sentences from each cluster according to sentence level labeling

Method	Agentic	Balanced	Communal
T1	59.08%	16.1%	24.82%
T2	20.94%	79.02%	0.04%

Table 2: Distribution of Job Ads by Method

4.5 Job-Level Labeling: Analysis and Results

In this section, we explore the results of applying our two distinct techniques aimed at labeling job advertisements (T1 and T2).

4.5.1 T1: Dictionary Look-up Method

Table 2 presents the outcomes of using Technique 1 (T1) to assign a single label to each job advertisement. Agentic-labeled job ads constitute the most prevalent category, followed by communal and balanced-labeled advertisements. It is important to note that this approach, while straightforward, primarily relies on counting words and may not provide the most accurate or nuanced understanding of gender coding in job ads.

Agentic
Australian National University - Senior Lecturer in Biological Chemistry. To further expand our capabilities, we are seeking candidates with expertise in protein chemistry, structural biology, biochemistry, biocatalysis, biophysics and/or protein engineering. ***Are you an exceptional candidate?*** *Can you demonstrate that, relative to your career stage:* ***you are, or have the potential to become, a world-class researcher*** *in biological chemistry, with strong, independent research programs funded by external grants; your research and teaching reflect the latest advances in their fields, with a* ***clear commitment to teaching excellence****; you are interested in dimensions beyond research and teaching; for example, public outreach, engaging with industry, science communication or tertiary science pedagogy; you are collaborative and collegial, and will be accessible to colleagues, research students and undergraduates; and* ***you have a high-level understanding of and commitment to the principles of inclusion, diversity, equity*** *and access in a University context.*

Balanced
We are one of the most ***diverse and vibrant*** *universities in the global capital. Our pioneering and forward-thinking vision is making a positive and significant impact to the* ***communities we serve, inspiring both our staff and students to reach their full potential****. We are seeking new colleagues to join in the Department of Bioscience lecturing in Pharmaceutical Science (BSc and MSc) and Pharmacology (BSc). Working as part of a dynamic team, you will teach and develop our modules, contribute to the design and delivery of our existing and new undergraduate and postgraduate programmes. You will be encouraged and supported to either join one of the ongoing research programmes or to initiate your own and to embrace our ethos of research-informed teaching.You will have BSc, MSc and PhD qualification in the appropriate discipline, as well as experience of teaching and/or student supervision in higher education as well as a* ***strong commitment to the student experience****.*

Communal
Assistant Professor of Physiology - Texas A&M International University - Texas A&M International University (TAMIU) is a comprehensive regional university and part of The Texas A&M University System. Poised at the Gateway to Mexico and serving as the cultural and intellectual hub of a ***vibrant multilingual and multicultural community****, it is also a designated Hispanic-serving institution (HSI). Qualified applicants should possess a Ph.D. or equivalent terminal degree in physiology or a related field. Candidates should demonstrate a* ***strong commitment to teaching, scholarly research, and service to the university and community. Additionally, candidates with experience in mentoring and supporting underrepresented students are highly encouraged to apply.***

Figure 7: Job ad examples from each T2 cluster

4.5.2 T2: Embedding-based Method

The results of assigning single labels to the jobs using T2 are presented in Table 2. The predominance of balanced labels highlights a prevalent use of language that combines agentic and communal traits or lacks strong alignment with either category. Figure 7 displays examples of job advertisements that were labeled using T2 from each cluster. The language in the agentic-labeled job description does not highlight why the university might be an appealing employer for potential candidates, and uses superlative language, seeking *'exceptional'* candidates who have the *'potential to become world-class researchers,'*. In contrast, the job description labeled as balanced includes text that promotes the university. Additionally, this job advertisement provides specific details about the roles, focusing on areas such as Pharmaceutical Science and Pharmacology. The job description labeled as communal emphasizes mentorship and community engagement, using phrases like *"service to the university and community."*

4.6 Nuanced Analysis of Bias within Job Advertisements

We conducted two distinct analyses to explore gender bias in job advertisements more thoroughly with the primary goal of pinpointing the sections of an advertisement where bias is most prevalent. This comprehensive approach allows us to identify specific segments of the ads where gender bias may be most pronounced. In the first analysis, we divided the advertisements into two main segments (Section 4.6.1), and in the second, we divided them into three parts (Section 4.6.2). This segmentation is informed by the typical structure of job ads in STEM.

4.6.1 Analyzing Clusters in Halved Job Advertisements

We divided each job advertisement in our dataset into two distinct parts: the top and bottom sections. The top part typically introduces the university and outlines the position's title. Meanwhile, the bottom part typically details the specific skills and qualifications required for the position. Subsequently, we examined how the three categories were distributed across these sections to gain insights into how gender bias may manifest differently across various sections of the ad. Results are reported in Figure 8.

Figure 8 indicates that agentic language is more prevalent in the bottom part of job descriptions compared to the top part. The frequency of communal language follows a similar pattern, although with significantly lower occurrences overall. These findings suggest a shift in language use from the beginning to the end of job descriptions, with the latter sections exhibiting a higher prevalence of

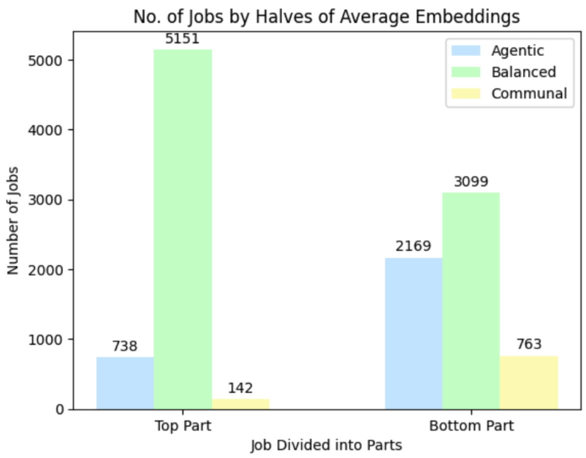

Figure 8: Distribution of agentic, balanced, and communal labels across the top and bottom parts of the job ads

(a) Top Half: Balanced

Details of the post: applicants must have completed a degree before the appointment in a data science field, which may include computer science, applied mathematics, organization, operations research, or a related field, they must demonstrate capabilities for writing code (python or r), basic knowledge in mathematical modelling, and prior experience using libraries in statistics, machine learning, or operations research. The position is funded by a research grant.

(a) Bottom Half: Communal

We will offer a competitive salary depending on qualifications and full access to the lbs environment. The ***candidate will benefit*** *from the resources of the mso community, which includes interactions with faculty and phd students,* ***mentoring*** *opportunities, access to research seminars, etc. We are an* ***equal opportunities employer****, and as such, we welcome applications from women, black and other ethnic minority candidates who are* ***under-represented*** *in our faculty.*

(b) Top Half: Agentic

About you: you will possess (or be near to completing) a relevant phd or equivalent qualification/experience in a related field of study, which may include (but is not restricted to) mathematics, physics, computer science, biophysics or engineering, you will be a ***nationally recognised authority*** *in mathematical modelling or computer simulation, you are* ***required to be motivated and demonstrate excellent knowledge*** *of the topic, possess* ***excellent problem solving****, interpersonal and communication skills and a collaborative spirit, combined with an ability to think carefully about your research.*

(b) Bottom Half: Agentic

In addition, you will: organization sufficient specialist knowledge in the discipline to develop/follow research programmes and methodologies, have a record of research output in high quality publications, have excellent written and verbal communication skills, have a ***record of active participation*** *of a member of a research team, 'be able to communicate complex and conceptual ideas to a range of groups, provide evidence of the ability to collaborate actively both internally and externally to complete research projects and advance thinking, be able to participate in and develop internal and external research networks,* ***be able to balance the pressures of research, administrative demands and competing deadlines, be willing to work flexibly to achieve project demands.***

Figure 9: Job ads with differently labelled halves

both agentic and communal language.

Examples of labels assigned to halves of the same job advertisement are provided in Figure 9(a). The top part of the job description predominantly contains balanced language, focusing on aspects such as technical skills and domain-specific information about a position in the data science field. In contrast, the bottom part of the job description exhibits a shift towards communal language, as evidenced by statements emphasizing community engagement, mentoring opportunities, and diversity initiatives.

In Figure 9(b), both parts of the job advertisement have been labeled as agentic. Both parts of this job advertisement focus less on domain-specific information about the position and more on the qualities and personality traits desired in the candidate. They do not convey much information about the university/organization offering the post but describe in superlative terms the qualities they seek in potential candidates.

4.6.2 Analyzing Clusters in Thirds of Job Advertisements

Each job description was segmented into three equal parts, with a label assigned to each section. As shown in Figure 10, the initial part predominantly exhibits balanced language. However, as the description progresses, there is a gradual transition from balanced language to increasingly pronounced agentic language towards the conclusion. Likewise, the frequency of jobs labeled as communal also increases in the latter third of the description but is always less than the agentic class.

Figure 11(a) shows examples of three parts of a job advertisement annotated separately. In this table, the first two parts of the job description describe domain-specific knowledge, while the last part describes agentic personality traits such as goal-oriented performance and motivation.

Figure 11(b) highlights the segmentation of the job advertisement into three parts. The initial two sections primarily outline technical skills and qualifications. The final part, labeled as communal, focuses on how candidates can contribute to and benefit from the academic community within the organization.

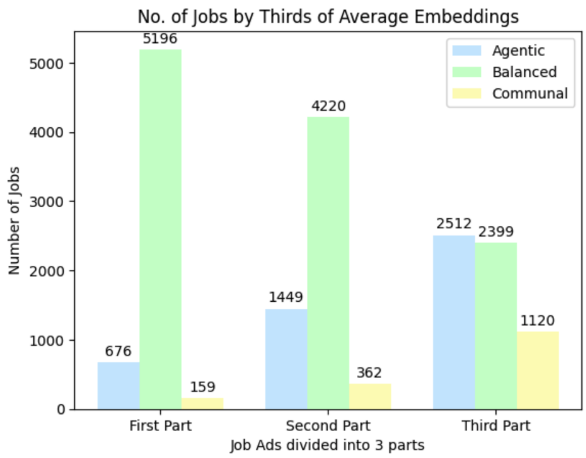

Figure 10: Distribution of agentic, balanced, and communal labels across job ad divided into 3 parts

5 Conclusion and Outlook

In this work, we not only identify the presence of gendered language but also highlight the sections of job ads where this language is most prevalent, offering targeted opportunities for intervention. We also carried out an analysis of our labeled dataset which revealed that agentic language is more frequently used than communal language, potentially perpetuating gender stereotypes that favor male applicants. However, balanced sentences that combine traits or lack strong alignment with gendered categories are the most common, suggesting an evolving landscape of job descriptions that attempt to be more inclusive.

Despite these efforts, our analysis shows a significant representation of agentic language, particularly in part of the ads where candidates are described. This suggests that, while job descriptions are evolving, there is still a tendency to favor language that might discourage some potential female applicants. The use of communal language, while present, is significantly lower, highlighting a continued area for improvement in how job roles are communicated to attract a more diverse applicant pool. Our findings should encourage academic institutions to critically assess and revise their job advertisements.

There are several potential areas that could be explored to further develop the contributions of this research. Future work could focus on conducting a qualitative research study that builds on Gaucher et al. (2011)'s study to survey participants in academia and explore how the participants describe gender bias in job advertisements. Moreover, an important direction for future research is to analyze the impact of language transitions within job descriptions on potential candidates. Studies show that readers often lose focus towards the end of documents (Duggan and Payne, 2011), which might influence how they perceive job descriptions that transition from balanced to agentic or communal language. Investigating whether the final part of a job description using agentic language deters candidates who identify with communal traits could offer valuable insights into optimizing job ad structures to attract a diverse applicant pool.

Additionally, considering the global scope of the collected job descriptions, it would be valuable to investigate the role of cultural factors in clustering

(a) First Third: Balanced

Research Assistant in data-driven methods for energy systems modelling and optimisation - we seek to recruit a highly motivated candidate with proven intellectual and technical ability to conduct research in the area of modelling energy supply systems and energy demand.

(a) Second Third: Balanced

The candidate will contribute to an on-going research activities in organization and supply) that concern estimating temporal and spatial energy demand (e.g. electricity, heating and cooling) in buildings in the uk under selected decarbonisation scenarios, and developing physics-aware organization methods for optimising the operation of integrated energy networks.

(a) Last Third: Agentic

*An ideal candidate would be a **self-starter** and a **resourceful team-player with an appetite for working with the industry.** The successful candidate will also contribute to the overall research performance of organization, carrying out research leading to the publishing of work. Candidates should also **pursue excellence** in research and inspire others to do the same.*

(b) First Third: Balanced

Fully funded postdoctoral position and details of the post: applicants must have completed a phd before the appointment in operations research/management, computer science, applied mathematics, econometrics, or a related field. They must demonstrate strong capabilities for conducting original theoretical or applied research using tools from: algorithm design, stochastic modelling, market design, or machine learning.

(b) Second Third: Balanced

The position is funded by a research grant. The term is one or two years, extensible subject to continued satisfactory performance. We will offer a competitive salary and full access to the lbs environment.

(b) Last Third: Communal

*The **postdoctoral fellow is expected to contribute and benefit from** the resources of the mso community, which includes interactions with faculty and phd students, **mentoring opportunities**, organisation of reading groups, and training sessions. We are an **equal opportunities employer**, and as such, we welcome applications from women, black and other ethnic minority candidates who are **under-represented** in our faculty.*

Figure 11: Job ads with differently labelled thirds

the three types of languages—agentic, communal, and balanced. Understanding how cultural contexts influence the use of gendered language in job descriptions could provide deeper insights into the patterns observed.

The research also has certain limitations that should be acknowledged. Firstly, identifying and quantifying gender bias in text is an inherently complex and challenging task. The interpretation of gender bias is subjective and varies among readers of job descriptions. Secondly, the dataset used for analysis contains job descriptions from a specific point in time, which poses issues related to the representativeness of the training data. Additionally, the techniques of labeling gender bias used in this study face certain challenged relating to the the reliability of the resulting clusters. Averaging sentence embeddings, while beneficial for capturing general trends, may overlook specific contextual nuances, potentially leading to inaccuracies in job ad classification. The distinctions between agentic, communal, and balanced language are not always clear-cut, which could lead to occasional misclassifications. These limitations underscore the importance of refining and validating the methodology to enhance its accuracy and reliability in future applications.

Bias Statement

In this research, we focus on identifying and addressing gender bias in academic job descriptions, particularly within STEM fields. The bias we investigate revolves around the use of language that implicitly favors certain gendered traits (agentic or communal) over others, thereby influencing the perceived suitability of job positions for individuals of different genders.

Representational harms occur when job descriptions portray certain gender groups more favorably or even more often than others, reinforcing stereotypes and potentially dissuading individuals from underrepresented genders from applying. Through this research, we saw that most academic job descriptions make use of agentic language when describing ideal candidates. The imbalance in the distribution of agentic and communal language within job advertisements can lead to differences in how these positions are perceived by potential applicants. Women, who often identify more with communal goals, may be discouraged from applying to positions that heavily emphasize agentic language. This not only limits their opportunities for career advancement but also perpetuates gender disparities within academia. By recognizing and addressing bias at its source, we strive to create a more equitable environment that fosters diversity and empowers individuals of all genders to pursue careers in STEM fields.

Finally, we argue that the binary representation used by Gaucher et al. (2011) and its associated gender stereotypes, which are prevalent in the field, are harmful and should be strongly opposed. We acknowledge that there may be other minority dimensions of analysis, beyond agentic and communal, that are yet to be uncovered, and hope our work contributes to opening these areas of inquiry.

References

David Bakan. 1996. Origination, self-determination, and psychology. *Journal of Humanistic Psychology*, 36(1):9–20.

Robert H. Bernstein, Michael W. Macy, Wendy M. Williams, Christopher J. Cameron, Sterling Chance Williams-Ceci, and Stephen J. Ceci. 2022. Assessing gender bias in particle physics and social science recommendations for academic jobs. *Social Sciences*, 11(2):74.

Marise P Born and Toon W Taris. 2010. The impact of the wording of employment advertisements on students' inclination to apply for a job. *The Journal of social psychology*, 150(5):485–502.

Bettina Casad, Jillian E. Franks, Christina E. Garasky, Melinda M. Kittleman, Alanna Roesler, Deidre Hall, and Zachary W. Petzel. 2020. Gender inequality in academia: Problems and solutions for women faculty in stem. *Journal of Neuroscience Research*, 99(1):13–23.

Erin Cech, Brian Rubineau, Susan Silbey, and Caroll Seron. 2011. Professional role confidence and gendered persistence in engineering. *American Sociological Review*, 76(5):641–666.

Flor Miriam Plaza del Arco, Amanda Cercas Curry, Alba Curry, Gavin Abercrombie, and Dirk Hovy. 2024. Angry men, sad women: Large language models reflect gendered stereotypes in emotion attribution. *Preprint*, arXiv:2403.03121.

Geoffrey Duggan and Stephen Payne. 2011. Skim reading by satisficing: Evidence from eye tracking. pages 1141–1150.

Daniel C. Feldman, William O. Bearden, and David M. Hardesty. 2006. Varying the content of job advertisements: The effects of message specificity. *Journal of Advertising*, 35(1):123–141.

Richard Frissen, Kolawole John Adebayo, and Rohan Nanda. 2022. A machine learning approach to recognize bias and discrimination in job advertisements. *AI and Society*, 38(4):1–14.

Danielle Gaucher, Justin Friesen, and Aaron C. Kay. 2011. Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1):109–128.

Amit Haim, Alejandro Salinas, and Julian Nyarko. 2024. What's in a name? auditing large language models for race and gender bias. *Preprint*, arXiv:2402.14875.

Anaïs Llorens, Athina Tzovara, Ludovic Bellier, Ilina Bhaya-Grossman, Aurélie Bidet-Caulet, William K. Chang, Zachariah R. Cross, Rosa Dominguez-Faus, Adeen Flinker, Yvonne Fonken, Mark A. Gorenstein, Chris Holdgraf, Colin W. Hoy, Maria V. Ivanova, Richard T. Jimenez, Soyeon Jun, Julia W.Y. Kam, Celeste Kidd, Enitan Marcelle, Deborah Marciano, and Nina F. Dronkers. 2021. Gender bias in academia: A lifetime problem that needs solutions. *Neuron*, 109(13):2047–2074.

Kat Matfield. 2016. Gender decoder. `https://gender-decoder.katmatfield.com/`.

Andrzej Maćkiewicz and Waldemar Ratajczak. 1993. Principal components analysis (pca). *Computers & Geosciences*, 19(3):303–342.

Kelly O'Brien, Veronica Petra, Divya Lal, Kim Kwai, Marian McDonald, Judy Wallace, Chloe Jeanmonod, and Rebecca Jeanmonod. 2022. Gender coding in job advertisements for academic, non-academic, and leadership positions in emergency medicine. *The American journal of emergency medicine*, 55:6–10.

Erin Oldford and John Fiset. 2021. Decoding bias: Gendered language in finance internship job postings. *Journal of Behavioral and Experimental Finance*, 31(5).

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Catherine E Read, Jovanna A Tracz, Nour Mhaimeed, Rylie N Mainville, and Carrie A Elzie. 2023. Examination of residency program websites for the use of gendered language and imagery. *BMC medical education*, 23(1):697.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pages 3982–3992.

Edith C Sella, Aparna Joshi, Sowmya Balasubramanian, Dharshini Senthilkumar, Tianwen Ma, and Prachi Agarwal. 2023. Is there gender bias in radiology job postings? *Current problems in diagnostic radiology*, 52(3):187–191.

Shiliang Tang, Xinyi Zhang, Jenna Cryan, Miriam Metzger, Haitao Zheng, and Ben Y. Zhao. 2017. Gender bias in the job market: a longitudinal analysis. In *Proceedings of the ACM on Human-computer Interaction*, pages 1–19.

Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605.

Yixin Wan, George Pu, Jiao Sun, Aparna Garimella, Kai-Wei Chang, and Nanyun Peng. 2023. "kelly is a warm person, joseph is a role model": Gender biases in LLM-generated reference letters. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 3730–3748, Singapore. Association for Computational Linguistics.

Jia Zhu, Ellen Zerbe, Monique Ross, and Catherine G P Berdanier. 2021. The stated and hidden expectations: Applying natural language processing techniques to understand postdoctoral job postings. In *ASEE Annual Conference*.