

Zihan Qiu, Li Zhang and Lei Lei\*

# Big Five Personalities Across 200 Years: A Large-Scale Study on the Description of Male- and Female-Dominated Occupations

<https://doi.org/10.1515/dsll-2025-0010>

Received June 15, 2025; accepted July 13, 2025; published online August 29, 2025

**Abstract:** The present study aims to investigate the issue of Big Five personalities from the perspective of gendered-segregated occupations. Descriptions of major male- and female-dominated occupations across 200 years were examined in a large-scale syntactically annotated corpus. Some results of interest were found. First, for both male- and female-dominated occupations, Agreeableness was found the most frequently occurring trait category, while Neuroticism was the least frequently occurring one. Second, in terms of the trends of the use of the Big Five traits, the past two centuries witnessed an overall drop in Agreeableness traits and a growth in Neuroticism traits for both male- and female dominated occupations. For Extraversion, Openness, and Conscientiousness, their trends differed between the two gendered occupational groups. Possible reasons for our findings, such as those pertaining to historical background, sociocultural environments, and implicit public perceptions, are offered. Significance of our work is also discussed.

**Keywords:** Big Five; personalities; gender-dominated occupations; gender differences; diachronic changes

## 1 Introduction

Personality research has been an area of interest in psychology and sociology over the past decades (Hamilton et al. 2006; Lee and Chin 2019), and many theories have been developed, of which the Big Five personality model is arguably the most widely examined one (Ye et al. 2018). The Big Five model is based on the lexical hypothesis that trait terms in everyday language “encode what ordinary people observe about the people in their social milieu,” and is therefore a “model of social perception” (Srivastava 2010, 69). The use of personality related words in social interactions not

---

\*Corresponding author: **Lei Lei**, Institute of Language Sciences, Shanghai International Studies University, Shanghai, China, E-mail: leileicn@126.com

**Zihan Qiu and Li Zhang**, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China

 Open Access. © 2025 the author(s), published by De Gruyter on behalf of Chongqing University, China

 This work is licensed under the Creative Commons Attribution 4.0 International License.

only indicates an individual's judgement of the human being, but also reflects standard, mainstream views in the society. The Big Five model has been widely employed to explore various social issues such as that of gender (e.g., Roivainen 2020; Schulz and Bahník 2019; Ye et al. 2018), mental health (Liu et al. 2013; Lyon et al. 2021), and education (e.g., Poropat 2009; Vedel 2014).

In this study, we focus on gender differences in the workplace and examine the Big Five personality adjectives that have been used to describe gendered-segregated or – dominated occupations across the past 200 years.

## 1.1 The Big Five Model

The Big Five model, also known as the Five-Factor Model, is a framework for research on individual differences and personality traits. It suggests that most personality traits can be classified into five broad dimensions: Extraversion (e.g., extroverted, talkative, and aggressive), Agreeableness (e.g., sympathetic, kind, and warm), Conscientiousness (e.g., organized, responsible, and efficient), Neuroticism (e.g., temperamental, irritable, and jealous), and Openness (e.g., intelligent, unconventional, and inquisitive) (Anglim and O'Connor 2019; Coker et al. 2002; Goldberg 1990, 1992).

The development of the Big Five model can be traced back to Allport and Odbert (1936), who collected 17,953 personality trait terms based on dictionary consultation and identified a lexicon of 4,505 stable traits. Follow-up studies filtered out the difficult and ambiguous terms of Allport and Odbert's (1936) lexicon, and obtained several hundreds of representative and high-frequency personality descriptive adjectives (e.g., Goldberg 1990; Norman 1967). Based on the adjectives, further research identified several general personality traits and proposed different models, such as the three-factor model Dark Tetrad and the six-factor model HEXACO (Ashton et al. 2004; Paulhus and Williams 2002). The Big Five model, however, stood out as the most compelling personality structure in personality research for its robustness and comprehensiveness (De Raad 1998; Saucier and Goldberg 1996; Ye et al. 2018), which thereby offers a theoretical basis for the present study.

Two methods have been used to measure the Big Five model, i.e., the questionnaire survey approach, where a series of descriptive statements (e.g., "I often feel inferior to others") are offered for respondents to recommend their acceptance of the statements (e.g., Hogan 1986), and the lexical approach, which collects and analyzes the use of personality trait terms of a lexicon in corpus texts (Roivainen 2013). Generally, the questionnaire survey approach is often used for personality assessment, whereas the lexical approach works for personality perception in a large dataset or for a long span. Given the aim of the present study, we followed the lexical

approach to delve into the descriptions of male- and female-dominated occupations over the past two centuries.

The Big Five model has been applied to a wide range of areas, including political behavior (Gerber et al. 2011), job performance (Barrick and Mount 1991; Van Aarde et al. 2017), academic performance (Poropat 2009; Vedel 2014), neurosciences (Liu et al. 2013; Lyon et al. 2021), social media (Bleidorn and Hopwood 2019; Liu et al. 2021), and alcohol abuse (Malouff et al. 2007). It functions as a predictor of certain behaviors and emotions (e.g., Barrick and Mount 1991; Correa et al. 2010) and a measure of people's personalities and psychological characteristics (e.g., Fischer et al. 2020; Liu et al. 2021; Qiu et al. 2012). For example, in an investigation of the Big Five personality and academic performance, Poropat (2009) found that students of high Conscientiousness tended to perform better in school and college. The finding was confirmed by many follow-up studies such as Vedel (2014) that reported on a positive correlation between Conscientiousness and students' academic achievement in terms of GPA with their majors as a moderator. In addition, in workplace contexts, conscientiousness and Neuroticism (or known as emotional stability) play critical roles (Van Aarde et al. 2017). Therefore, highly conscientious and emotionally stable applicants are preferred by employers.

## 1.2 The Big Five, Gender, and Occupations

Many studies have examined the relationship between gender and personality traits from the big five perspective (e.g., Iimura and Taku 2018; Mac Giolla and Kajonius 2019; Weisberg et al. 2011). For example, Weisberg et al. (2011) found that women scored higher than men in Neuroticism and Agreeableness, and lower in Extraversion. For another example, Iimura and Taku (2018) investigated gender differences between resilience and Big Five personality traits of Japanese teenagers. They found that Neuroticism was the most effective predictor of resilience for female teenagers and Extraversion for male teenagers.

Recently, personality psychologists began to explore gender differences based on linguistic features in the texts. For example, Roivainen (2015) examined the use of the bigrams of *adjective + person* and *adjective + women* on Twitter and in the Google Books Ngrams database and found that the most frequent personality adjectives did not necessarily have high factor loadings. The method was adopted by many follow-up studies. For instance, Ye et al. (2018) examined the use of personality trait adjectives as modifiers for *woman/women* and *man/men* in Google English Books Ngrams to investigate gender similarities and differences over the past two hundred years. A finding of interest is that personality adjectives concerning the Big Five factors, except Openness, are more frequently used to modify men than women.

Schulz and Bahník (2019) was much similar to Ye et al. (2018) but considered more modified nouns (i.e., *man*, *woman*, *boy*, *girl*, *person*, and *child*). They found that *men* tend to be described more positively than *women* across the past century. To be specific, *men* were more likely to be perceived as *honest* and *good* and less likely as *foolish* and *unhappy*.

No research has investigated if there exists any gender difference in the description of occupations, particularly those gender-segregated ones, from the Big Five perspective. Gender segregation refers to the phenomenon that workers concentrate in distinct occupations due to their gender (Reskin 1993). Gender segregation in the workplace has received much attention and long been a key focus in the pursuit of gender equality, particularly for women's rights and female labor force participation (Kelly and Grant 2012; Renzulli et al. 2013). For example, women are underrepresented in high-paying occupations such as programmers, firefighters, and engineers, but predominate in such professions as nursing and K-12 teaching that offer lower salaries, benefits, or promotion opportunities. Against this backdrop, occupational gender stereotypes are easily produced for gender imbalance. Motschenbacher and Roivainen's (2020) interdisciplinary research, for instance, argued that nurse is often subconsciously considered as female. For such gender-stereotyped occupations, how they have been described is still open for exploration.

### 1.3 The Present Study

The purpose of the present study is to explore how the gender-segregated occupations were described across the past 200 years from the Big Five perspective. The following two questions guided our research:

- 1) How are the occupation-related Big Five traits distributed in the past two centuries?
- 2) What are the changing patterns of the occupation-related Big Five traits in the past two centuries?

## 2 Methods

In this section, we describe the research methods used in the study.

## 2.1 Dataset

The dataset used in the study is the Syntactic Ngrams Corpus (Eng-1M collection), a part of the Google Books Ngrams Dataset (Goldberg and Orwant 2013; Roth et al. 2018). We used the dataset for the following reasons. First, the dataset of ngrams was extracted from one million English books published from 1520 to 2008. The large size and long span of the dataset offered us a chance for a diachronic exploration of the description of gender-segregated occupations (see Table 1 for the descriptive statistics of the dataset). Second, as a mirror of society, books revealed to us the social realities, public perceptions, and dominant cultures in a certain period (Mayer 2007). The choice of books as the data source would thus help detect public perceptions of gender in the workplace context. Third and more importantly, the texts in the dataset have been syntactically parsed, with every ngram annotated with part-of-speech tags and dependency relations (Goldberg and Orwant 2013). The syntactic annotations provide us a chance to retrieve useful information for the present study (see Section 2.3 for the details).

## 2.2 List of Gender-Related Occupation Words

To identify the typical gender-segregated occupations, we used the latest report of the U.S. Bureau of Labor Statistics (2021) (available at <https://www.bls.gov/opub/reports/womens-databook/2020/home.htm>), which reports on the sex ratios of more than 550 occupations. Hence, based on the statistics of the report, we classified an occupation as a female-dominated one if the proportion of its female employees was larger than 70 % and as a male-dominated one if the proportion of its female employees was below 30 %.

Occupations that met the following criteria were selected for our analyses. First, an occupation should be specific in meaning and clearly distinguished from other

**Table 1:** Description of the dataset.

| Corpus characteristics   | Syntactic ngrams corpus (Eng-1M collection) |
|--------------------------|---------------------------------------------|
| Number of books          | 1 million                                   |
| Tokens/Words             | 101.3 billion                               |
| Time span                | 1520–2008                                   |
| Language                 | English                                     |
| Syntactic representation | Dependency grammar                          |

occupations. For example, *demonstrators* and *painters* are terms that refer to more than one occupation (i.e., salespersons versus marchers and artists versus decorators). Hence, they were filtered out from the follow-up analyses. Second, we included the occupations whose titles are one word, and excluded those whose titles consist of more than one word (e.g., social workers). The reason is that the dataset used in our study contains ngrams or word pairs of dependency relations that are composed of a Big Five adjectival modifier and a one-word occupation title (see below). In addition, we combined occupation titles of similar meaning into one title. For example, *licensed practical and licensed vocational nurses* (90.8 % female), *nurse practitioners* (87.8 % female), and *nursing aides* (88.3 % female) were combined into *nurses*. A list of 38 occupations was obtained, with 19 for male-dominated occupations and 19 for female-dominated occupations (see Table 2). We finally referred to Koch et al.'s (2015) meta-analysis of gendered occupations and confirmed the representativeness of the 38 occupations. The list of removed occupations is presented in Appendix A, with the final merged occupation terms in Appendix B in the Supplementary Material.

## 2.3 Data Processing and Statistical Analyses

The procedure of data processing and statistical analyses are described as follows. First, we downloaded the texts of the Google Books Syntactic Ngrams Corpus (<http://commondatastorage.googleapis.com/books/syntactic-ngrams/index.html>). The texts contain ngrams of dependency relations and their frequencies in certain years.

Second, we extracted, from the downloaded texts, all ngrams or word pairs in the dependency relation *amod* (i.e., *adjectival modifier*, or an adjective as a modifier to modify a noun). Then, we retrieved two types of bigrams, i.e., *Big Five adjective + occupation* (e.g., *humble auditor*, *shrewd detective*; hereinafter Big Five ngrams) and *any adjective + occupation* (hereinafter any-adjective ngrams). On the one hand, in the Big Five ngrams, the *occupation* should be one of the 38 gender-segregated occupations, in either singular or plural form. Its *adjectives* should be on the list of 435 personality adjectives developed by Saucier and Goldberg (1996). It should be noted that three adjectives, i.e., *practical*, *social*, and *independent*, were not used in our analyses, since they were often used to specify an occupation (e.g., *practical nurse*) rather than for descriptive personality purposes. On the other hand, in any-adjective ngrams, the adjectives may be included or not included in Saucier and Goldberg's (1996) list of personality adjectives (e.g., *beautiful* and *strong*), as long as its occupations are on our list 38 occupations. We extracted the *any-adjective ngrams* to facilitate the follow-up statistical analysis, i.e., data normalization.

**Table 2:** Frequencies of 38 male- and female-dominated occupations modified by *any adjective* (the any-adjective ngrams) in the Google Books Syntactic Ngrams Corpus.

| Male-dominated occupations |               |                            | Female-dominated occupations |               |                            |
|----------------------------|---------------|----------------------------|------------------------------|---------------|----------------------------|
| Occupation                 | Raw frequency | Total normalized frequency | Occupation                   | Raw frequency | Total normalized frequency |
| 1 Clergy/priest            | 1,201,546     | 147,523                    | Teacher                      | 1,713,742     | 210,409                    |
| 2 Engineer                 | 535,160       | 65,705                     | Secretary                    | 435,419       | 53,460                     |
| 3 Guard                    | 503,554       | 61,825                     | Nurse                        | 353,007       | 43,341                     |
| 4 Laborer                  | 467,749       | 57,429                     | Maid/<br>housekeeper         | 248,674       | 30,532                     |
| 5 Police/<br>policeman     | 332,132       | 40,778                     | Assistant                    | 226,712       | 27,835                     |
| 6 Operator                 | 252,435       | 30,993                     | Psychologist                 | 183,427       | 22,521                     |
| 7 Architect                | 221,521       | 27,198                     | Counselor                    | 163,703       | 20,099                     |
| 8 Mechanic                 | 215,288       | 26,432                     | Therapist                    | 88,724        | 10,893                     |
| 9 Courier/<br>messenger    | 183,952       | 22,585                     | Librarian                    | 58,842        | 7,224                      |
| 10 Driver                  | 132,073       | 16,216                     | Accountant                   | 50,164        | 6,159                      |
| 11 Mover/porter            | 126,256       | 15,501                     | Auditor                      | 41,257        | 5,065                      |
| 12 Chef/cook               | 125,489       | 15,407                     | Hostess                      | 38,640        | 4,744                      |
| 13 Detective               | 42,858        | 5,262                      | Waiter/<br>waitress          | 36,938        | 4,535                      |
| 14 Carpenter               | 39,090        | 4,799                      | Tailor                       | 32,310        | 3,967                      |
| 15 Butcher                 | 23,316        | 2,863                      | Interviewer                  | 15,333        | 1,883                      |
| 16 Barber                  | 14,273        | 1,752                      | Teller/cashier               | 13,961        | 1,714                      |
| 17 Programmer              | 13,128        | 1,612                      | Dietitian                    | 2,173         | 267                        |
| 18 Electrician             | 7,814         | 959                        | Receptionist                 | 725           | 89                         |
| 19 Plumber                 | 3,273         | 402                        | Babysitter                   | 173           | 21                         |
| Total                      | 4,440,907     | 545,242                    | Total                        | 1,713,742     | 454,758                    |

The occupations are listed in the descending order of their normalized frequencies in the Google Books Syntactic Ngrams Corpus.

Third, the raw frequencies of the target ngrams in each year were retrieved. Since the data before 1800 were scarce (i.e., the total frequency of all ngrams per year is lower than 100), we only examined the data from 1800 to 2008. Then, we calculated the total frequency of each ngram by decade, i.e., 21 decades divided the past 209 years to facilitate the subsequent data processing. All the foregoing data processing were accomplished with home-made Python scripts.

Last, the statistical analyses were carried out to address the research questions. For the first research question, we calculated the frequencies of personality adjectives on each Big Five factor. To make the frequencies comparable across the occupations and decades, the frequencies were normalized by dividing the raw

frequency of the Big Five adjectives for male- or female-dominated occupations on each factor by the total frequency of all Big Five adjectives for the gender-segregated occupations, and then multiplying one million (Shi and Lei 2021).

For the second research question, we normalized the frequency of the Big Five ngrams by decade and factor by dividing the raw frequencies of the Big Five adjectives for male- or female-dominated occupations in a decade by the total frequency of all Big-Five adjectives in that period, and then multiplied by one million (Lin and Lei 2020; Shi and Lei 2021; Ye et al. 2018). The normalized frequencies were plotted and smoothed by locally weighted regression (loess) curves to facilitate the observation of the diachronic changes (Huang et al. 2017; Schulz and Bahník 2019). Then, we used the Bayesian Information Criterion (BIC, with the R package *segmented*) (Muggeo 2008; Schulz and Bahník 2019) to identify the breakpoints (also known as change points or inflection points) of the curves. The breakpoints that were measured in years were manually adjusted to be measured in decades so that they could be in line with the present study. For example, the breakpoint of 1845 was included in the decade of the 1840s. With the appropriate breakpoints chosen, we divided the examined period into several time segments, in each of which simple linear regressions were performed to check whether the upward or downward trends were statically significant (Lei and Zhang 2018; Wen and Lei 2022). For those curves with no significant breakpoints, simple linear regressions were fitted. All the foregoing statistical analyses were accomplished with home-made R scripts.

### 3 Results

In this section, the results of the study are reported.

#### 3.1 Overall Distributions of the Occupation-Related Big Five Traits

The distributions of Big Five ngrams concerning gender-segregated occupations are reported in Table 3. Male-dominated occupations were, in terms of normalized frequency, most frequently described as Conscientiousness, followed by Openness and Agreeableness. In contrast, female-dominated occupations were most frequently portrayed as Agreeableness, with Conscientiousness and Openness ranking second and third. For both gender-dominated occupation groups, Neuroticism was the least frequently occurring factor. Furthermore, female-dominated occupations were more likely to be described in our dataset concerning personality traits than male-dominated occupations.

**Table 3:** Overall distributions of the occupation-related Big Five traits.

| Big Five factor/<br>category | Male-dominated<br>occupations |                         | Female-dominated<br>occupations |                         | Total normalized<br>frequency |
|------------------------------|-------------------------------|-------------------------|---------------------------------|-------------------------|-------------------------------|
|                              | Raw<br>frequency              | Normalized<br>frequency | Raw<br>frequency                | Normalized<br>frequency |                               |
| I (Extraversion)             | 22,496                        | 1,417,492               | 21,331                          | 1,305,195               | 2,722,687                     |
| II (Agreeableness)           | 30,944                        | 2,014,511               | 81,216                          | 4,785,489               | 6,800,000                     |
| III (Conscientiousness)      | 44,274                        | 3,018,412               | 50,659                          | 2,779,180               | 5,797,592                     |
| IV (Neuroticism)             | 13,485                        | 783,286                 | 4,866                           | 279,946                 | 1,063,232                     |
| V (Openness)                 | 40,099                        | 2,623,103               | 34,829                          | 1,993,386               | 4,616,489                     |
| Total                        | 151,298                       | 9,856,804               | 192,901                         | 11,143,196              | 21,000,000                    |

Concerning the specific Big Five adjectives, Table 4 shows the top 40 most frequently used Big Five adjectives. The total frequency of these 40 adjectives accounted for more than half of the frequency of all Big Five terms. As shown in the table, most of the terms were positive. To be specific, male-dominated occupations were more likely to be described as *simple*, *casual*, *efficient*, and *honest*, while female-dominated occupations as *religious*, *wise*, *efficient*, and *moral*. One point worth noting here is that none of those highly frequent adjectives for female-dominated occupations belonged to Neuroticism. This finding seems contradictory to those in previous studies that Neuroticism had long been claimed to be more closely associated with women (Iimura and Taku 2018; Jorm 1987; Weisberg et al. 2011). We will get back to this issue in Section 4.

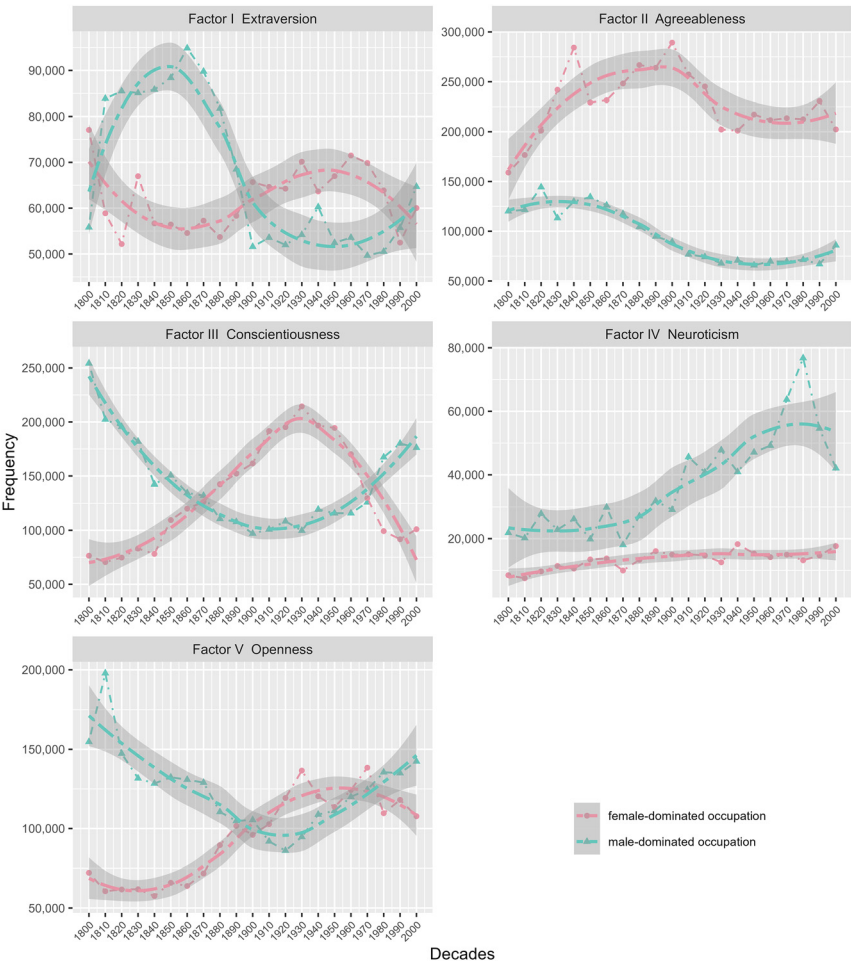
**Table 4:** The most frequent occupation-related Big Five adjectives.

| Gender occupational<br>group    | Big Five adjectives                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Male-dominated<br>occupations   | <i>simple</i> (V), <i>casual</i> (IV), <i>efficient</i> (III), <i>honest</i> (II), <i>logical</i> (III), <i>intelligent</i> (V), <i>strict</i> (III), <i>ingenious</i> (V), <i>careful</i> (III), <i>ignorant</i> (V), <i>zealous</i> (I), <i>humble</i> (II), <i>active</i> (I), <i>clever</i> (V), <i>industrious</i> (III), <i>friendly</i> (I), <i>careless</i> (III), <i>dignified</i> (III), <i>wise</i> (III), <i>brave</i> (IV)             |
| Female-dominated<br>occupations | <i>religious</i> (II), <i>wise</i> (III), <i>efficient</i> (III), <i>moral</i> (II), <i>intelligent</i> (V), <i>conscientious</i> (II), <i>progressive</i> (V), <i>active</i> (I), <i>alert</i> (III), <i>enthusiastic</i> (I), <i>careful</i> (III), <i>thoughtful</i> (II), <i>sympathetic</i> (II), <i>ethical</i> (V), <i>creative</i> (V), <i>earnest</i> (II), <i>honest</i> (II), <i>clever</i> (V), <i>simple</i> (V), <i>critical</i> (II) |

The adjectives are listed in descending order of their normalized frequencies.

### 3.2 Trends of the Occupation Related Big Five Traits Across the Past 200 Years

In this section, we report on the diachronic trajectories of the use of personality traits in terms of the Big Five factors and gender-dominated occupations (see Figure 1).



**Figure 1:** Trends of the use of the Big Five traits concerning male- and female-dominated occupations from 1800 to 2008. The smooth curves were fitted with loess regression, and the shaded areas represent 95 % confidence intervals.

### 3.2.1 Factor I Extraversion

For male-dominated occupations, the breakpoint of 1865 (95 % CI = 1855–1874) was detected; the results of simple linear regression showed a significant increase of the use of Extraversion adjectives with a medium effect size from 1800 to 1860 ( $F(1, 5) = 8.02$ ,  $p = 0.04$ , multiple  $R^2 = 0.62$ , adjusted  $R^2 = 0.54$ ), and a significant decrease with a medium effect size afterwards ( $F(1, 13) = 11.15$ ,  $p < 0.001$ , multiple  $R^2 = 0.46$ , adjusted  $R^2 = 0.42$ ).

For female-dominated occupations, although there was no statistically significant change in the use of Extraversion adjectives across time ( $F(1, 19) = 0.45$ ,  $p = 0.51$ , multiple  $R^2 = 0.02$ , adjusted  $R^2 = 0.03$ ), a slightly upward trend was observed.

### 3.2.2 Factor II Agreeableness

For male-dominated occupations, the results of simple linear regression showed a significant drop of the use of Agreeableness adjectives with a medium effect size ( $F(1, 19) = 60.64$ ,  $p < 0.001$ , multiple  $R^2 = 0.76$ , adjusted  $R^2 = 0.75$ ).

For female-dominated occupations, the breakpoint was located in the year 1906 (95 % CI = 1887–1925); the results of simple linear regression showed a significant increase in the use of Agreeableness adjectives with a medium effect size over the nineteenth century ( $F(1, 9) = 19.44$ ,  $p < 0.001$ , multiple  $R^2 = 0.68$ , adjusted  $R^2 = 0.65$ ), and a significant fall with a medium effect size across the twentieth century ( $F(1, 9) = 6.828$ ,  $p = 0.03$ , multiple  $R^2 = 0.43$ , adjusted  $R^2 = 0.36$ ). It should be noted that female-dominated occupations are more likely to be described with Agreeableness adjectives than male-dominated occupations across the examined 200 years, which supported the point of women's higher Agreeableness found in previous studies (Weisberg et al. 2011).

### 3.2.3 Factor III Conscientiousness

For male-dominated occupations, the breakpoint of 1902 (95 % CI = 1891–1914) was detected; a significant increase of the use of Conscientiousness adjectives with a large effect size was found in the first half of the examined 200 years ( $F(1, 9) = 97.75$ ,  $p < 0.001$ , multiple  $R^2 = 0.92$ , adjusted  $R^2 = 0.91$ ), followed by a significant decline in the second half of the period ( $F(1, 9) = 39.41$ ,  $p < 0.001$ , multiple  $R^2 = 0.81$ , adjusted  $R^2 = 0.79$ ).

For female-dominated occupations, the breakpoint was the year 1935 (95 % CI = 1928–1944); the use of Conscientiousness adjectives increased significantly with a large effect size before the 1930s ( $F(1, 12) = 247.30$ ,  $p < 0.001$ , multiple  $R^2 = 0.95$ , adjusted  $R^2 = 0.95$ ) and decreased significantly with a large effect size after that decade ( $F(1, 6) = 69.42$ ,  $p < 0.001$ , multiple  $R^2 = 0.92$ , adjusted  $R^2 = 0.91$ ).

### 3.2.4 Factor IV Neuroticism

The past two hundred years witnessed significant rises in the use of Neuroticism adjectives for both male- and female-dominated occupations with medium effect sizes, but the effect size for the male-dominated ones was larger than that of the female-dominated ones (male-dominated occupations:  $F(1, 19) = 46.73$ ,  $p < 0.001$ , multiple  $R^2 = 0.70$ , adjusted  $R^2 = 0.71$ ; female-dominated occupations:  $F(1, 19) = 25.86$ ,  $p < 0.001$ , multiple  $R^2 = 0.57$ , adjusted  $R^2 = 0.55$ ). During this period, Neuroticism adjectives were more commonly used to describe male-dominated occupations than to describe the female-dominated ones, which was consistent with our earlier-mentioned observation (see Section 3.1) that no Neuroticism adjectives was highly frequently used for female-dominated occupations.

### 3.2.5 Factor V Openness

For male-dominated occupations, with the breakpoint of 1917 (95 % CI = 1902–1933), the use of Openness adjectives increased significantly with a medium effect size until the 1910s ( $F(1, 10) = 30.77$ ,  $p < 0.001$ , multiple  $R^2 = 0.75$ , adjusted  $R^2 = 0.73$ ), followed by a significant decrease with a large effect size ( $F(1, 8) = 196.00$ ,  $p < 0.001$ , multiple  $R^2 = 0.96$ , adjusted  $R^2 = 0.96$ ).

For female-dominated occupations, the calculation of simple linear regression exhibited a significant increase with a medium effect size across the past 200 years ( $F(1, 19) = 55.84$ ,  $p < 0.001$ , multiple  $R^2 = 0.75$ , adjusted  $R^2 = 0.73$ ).

In general, the use of Neuroticism adjectives increased for both male- and female-dominated occupations over the past 200 years, while the use of Agreeableness adjectives decreased over the twentieth century. For Extraversion, Conscientiousness, and Openness, despite their more complicated diachronic changes, they all showed an overall decreasing use for male-dominated occupations and an overall increasing use for female-dominated occupations across the examined period (or at least before the mid-twentieth century). The major trends of the occupation-related Big Five traits will be discussed in detail in the next section.

## 4 Discussion

In this section, we discuss major findings of interest in our study.

## 4.1 Analysis of the Distribution of the Big Five Traits

The first research question in this study focuses on the Big Five personality adjectives used to describe the 38 occupations. Our results supported the works of previous studies that, of the Big Five factor categories, Agreeableness and Neuroticism were the most and least frequently ones to describe both male- and female-dominated occupations across the past 200 years (Peabody and Goldberg 1989; Roivainen 2013, 2020; Ye et al. 2018).

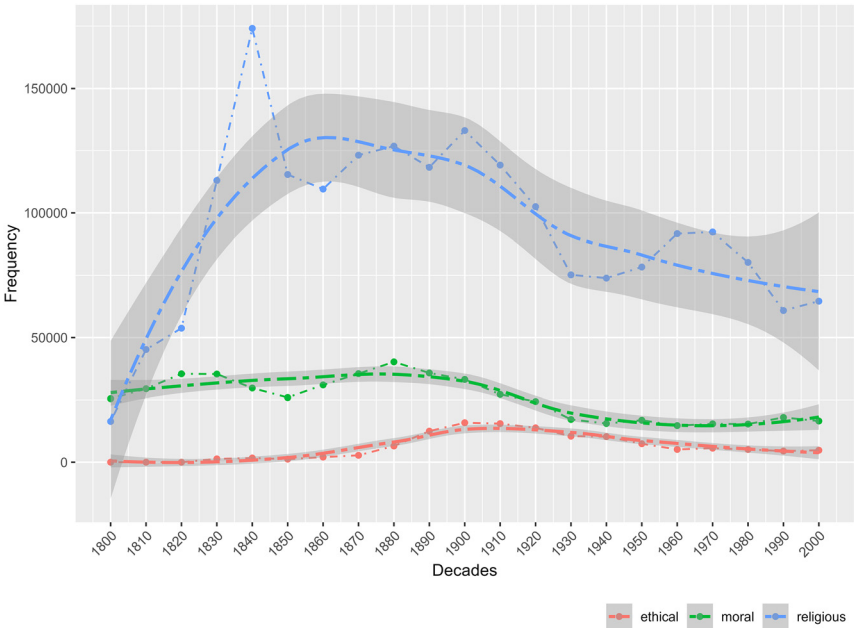
A possible reason for the prevalence of Agreeableness might be the crucial role that Agreeableness plays in collaboration in the workplace. Firstly, Agreeableness is strongly correlated with reciprocal altruism (Ashton et al. 1998), which refers to “the exchange of beneficial acts between individuals” (Seyfarth and Cheney 1984, 541). That is, one performs an act which is beneficial to another at the expense of himself/herself with the expectation that he/she will be reciprocated someday (Ashton et al. 1998; Seyfarth and Cheney 1984, 541). According to Ashton et al. (1998), reciprocal altruism is conducive to long-term collaboration, particularly when the gains of the recipient exceed the costs of the benefactor, since it is sensible that people prefer to collaborate with those who can bring them a lot and seldom bargain with them over the gains or losses. Therefore, people with a higher degree of agreeableness, who display a greater tendency towards reciprocally altruistic behavior, are more likely to sustain a cooperative relationship with others. Secondly, as a pro-social personality trait category, Agreeableness serves as a regulator of interpersonal conflict (Dijkstra et al. 2005). It has been demonstrated that people high in Agreeableness desire to develop a good relationship with others, which motivates them to avert conflicts, and vice versa (e.g., Dijkstra et al. 2005). Even when conflicts occur, those with positive Agreeableness-related traits (e.g., *sympathetic, sincere, cooperative, truthful, honest*) can handle them well (Morgeson et al. 2005). The conflict-management ability is highly valued in teamwork settings, since it may help promote more effective collaboration in teamwork. Considering the importance of collaborative interactions in the workplace, regular associations between Agreeableness adjectives and occupation-related terms are expected.

The fewer occurrence of Neuroticism traits may be explained by the linguistic positivity bias (Roivainen 2020; Wen and Lei 2022; Ye et al. 2018). The linguistic positivity bias is a phenomenon that positive words are preferred to be used in human language, which may result from the motive for more pleasant social interactions, politeness considerations, and the tendency that the brighter side of life is favored by humans over the darker one (Nezhad 2021; Wen and Lei 2022). Of the Big Five factors, Neuroticism is the only factor with more negative trait terms than the positive ones (Costa and McCrae 1980; Saucier and Goldberg 1996). Because of the

human instinctive affection for the positive words and things, Neuroticism adjectives were rarely used in the past 200 years.

Linguistic positivity bias may also be evidenced by the fact that most of the top 40 most frequent Big Five adjectives modifying the gendered occupations (see Table 4) are positive ones (35 of the 40 adjectives, to be specific). However, not all of those positive terms, if not just taken literally, are as agreeable as they seem to be. A case in point is *religious* (see Figure 2), which is usually used to describe someone loyal to religion. In our study, *religious* (18 %) was the most frequent personality trait for female-dominated occupations, but its prevalence seemingly implies the society's emphasis on female subordination. First, being *religious* for women was tantamount to obeying an androcentric belief system, since women were the victims of religion (Jakubiak and Murphy 1986). Religion has long been viewed as the embodiment of patriarchy and misogyny, where men are naturally superior to and thus justifiably able to rule over women in the name of the divine God (Jakubiak and Murphy 1986; Morgan 2002). A straightforward manifestation lay in the underrepresentation and marginalization of female workers in clerical professions, and even for women who served the church, their laborers were disparaged as “unskilled” and “simple” (Mangion 2005, 227). Moreover, women were shackled by a mass of religious dogmas that imposed great restrictions on women's dress, marriage, codes of behavior, reproduction, etc. Some regulations were particularly aimed at occupational gender segregation, preventing women from entering some traditionally male occupations (Bahramitash 2004). To this end, *religious* is in fact a trait in favor of males. Second, the high frequency of *religious* could be regarded as an attempt to curtail women's participation in the social domain at that time, which would threaten male privileges (Reskin and Irene 1988). As the past 200 years witnessed religious decline caused by secularization and women's entry into the job market (Franck and Iannaccone 2014; Roth et al. 2018), the frequent occurrence of *religious* for the feminized occupations could be some form of covert language oppression, making oblique yet repeated references to women's subordination to male dominance.

Besides *religious*, frequently used Big Five descriptive adjectives such as *moral* (4.5 %) and *ethical* (1.4 %) are also of interest. It seems that people tended to use such terms with the overtones of obedience and conservativeness, which, if not to require or indoctrinate, but at least to expect women to comply with the so-called social norms. The traditional, social expectations were thus projected on women, molding them into what a patriarchal society wanted them to be (Jakubiak and Murphy 1986). In this sense, *religious*, *moral*, and *ethical* cannot be regarded as positive terms, particularly when they were seldom used to describe male-dominated occupations in the past two centuries.



**Figure 2:** Trends of some high-frequency Big Five adjectives.

## 4.2 Diachronic Analysis of the Big Five Traits

The second research question sought to explore the diachronic trends of the gendered occupations from a Big Five perspective from 1800 to 2008.

First, an upward trend in the use of Neuroticism traits was found for both male- and female-dominated occupations. A possible explanation for these results may be the increased anxiety and stress in society, both of which are the negative emotions strongly correlated with Neuroticism (Costa and McCrae 1980; Frankenbach et al. 2021; Twenge 2000). The past centuries witnessed social changes resulting from the information revolution and industrialization, which boosted people’s standard of living and enhanced productivity in every industry (Rosen 1998; Twenge 2000). However, social progress did not bring about personal happiness. People even felt more anxious and insecure in “the age of anxiety” (Spielberger and Rickman 1990, 69), due to concerns from environmental threats such as health problems and infectious diseases, growing crime and violence, and the outbreak of war (Twenge 2000). Given the strong correlation between Neuroticism and such negative emotions as anxiety, worry, fear, insecurity, loneliness, and guilt (Costa and McCrae 1980; Frankenbach et al. 2021; Twenge 2000), the widespread feelings of anxiety and stress

among the public could cause the rise in Neuroticism scores in the whole society (Twenge 2000). Therefore, both male- and female-dominated occupations became more likely to be associated with the Neuroticism-related personality traits.

In addition, Neuroticism traits were more frequently used to describe male-dominated occupations than female-dominated ones. This finding, together with the absence of Neuroticism adjectives used to modify female-dominated occupations on our list of highly frequent Big Five adjectives (see Section 3.1), seems inconsistent with the results in previous studies that argued women had a higher level of Neuroticism (Imura and Taku 2018; Jorm 1987; Weisberg et al. 2011). The public anxiety (that stemmed from the environmental threats) could still account for our finding here. According to Twenge (2000), one source of anxiety and stress for men in the past century was the increasing number of women who entered higher education and pursued their professional careers. The increased anxiety and stress in society, particularly for men, may explain the more Neuroticism-related nature of male-dominated occupations in the twentieth century. This finding is also complementary to and provides evidence to our assumption on the frequent occurrence of *religious* as an effort to defend male privileges in fear of the women rise in the previous section.

Second, a significant decrease in the use of Agreeableness adjectives for both occupational groups was found in the twentieth century (despite an increase in the nineteenth century for feminized occupations). The findings may be due to the decreasing social connectedness of the people in the society. The Agreeableness scores were proven correlated with the social connectedness or people's connection/attachment with/to others (Ashton et al. 1998). When people's social connectedness in terms of social indicators such as the divorce rate, the birth rate, and the number of citizens living alone went down over the past century (Twenge 2000), their lives became less agreeable, which was accordingly reflected in the less use of Agreeableness terms to describe the traits of the occupations in our study.

Another point of interest is the more frequent occurrences of Agreeableness for female-dominated occupations, which is consistent with the findings in previous studies that women showed higher scores on Agreeableness than men (Ashton et al. 1998; Weisberg et al. 2011). Women are considered more sympathetic and liable to be connected with others and to form social attachments (Ashton et al. 1998). Such traits of women are, as discussed earlier, positively correlated to Agreeableness, hence resulting in higher occurrences of Agreeableness for female-dominated occupations across the past 200 years.

Similar trends were found in Extraversion and Openness across the examined period. Specifically, Extraversion and Openness exhibited more or less an increasing tendency to describe female-dominated occupations. For male-dominated occupations, both the two trait categories experienced a fall before the mid-twentieth

century. The opposite trends between the two gendered occupational groups may arise from some important historical events that afforded women unprecedented opportunities to go out of the family and enter the labor markets, such as the First Wave of feminism, the Great Depression, and the World War II (Adelmann et al. 1989). For example, the tight finances during the Great Depression forced every member of the society, regardless of her/his gender, to find a paid job. It should be noted that people with Extraversion traits at work are deemed as sociable, energetic, confident, and gregarious, and those with Openness traits are considered as creative, intelligent, curious, and broad-minded (Neal et al. 2012). Hence, both of the two trait categories encourage people to communicate with the outside world and may thus not be common in the women who were isolated from the public sphere and only revolved around family affairs. As an increasing number of women entered a profession, whether by choice or necessity, they were more likely to generate energy and maintain intellectual curiosity through social interactions, i.e., have higher levels of Extraversion and Openness (Neal et al. 2012). By contrast, women's engagement in the public domain captured part of the employment market and influenced male employment to some extent, which may bring about the decreasing use of Extraversion and Openness traits for male-dominated occupations.

Conscientiousness had roughly the same trends as Extraversion and Openness before the mid-twentieth century, after which, however, the trend reversed. It may suggest that women failed to have feelings about work as positive as men at the end of the last century, considering the positive associations between Conscientiousness and job satisfaction (Topino et al. 2021). Many factors would contribute to job satisfaction, such as career progressions, fringe benefits, work pressure, and management practices, which are also often the manifestations and the results of occupational gender segregation (Kelly and Grant 2012). Of course, the trends may not suffice to conclude women's low job satisfaction, nor is that the focus of our research. There must be many other possible explanations for the trends of Conscientiousness, and we just offer a presumption from the lexical Big Five perspective that helps us understand how women may feel in the workplace at that time.

## 5 Implications and Suggestions for Future Research

The present study is significant as follows. First, as one of the major findings of our study, Agreeableness and Neuroticism were respectively the most and least frequently used trait categories, with the occurrence of Agreeableness traits

decreasing and that of Neuroticism traits increasing over the past 200 years. The trends may display the widespread negative emotions such as stress, fear, anxiety, and depression across society in that period, as people's negative aspects of personality traits got increasingly obvious.

Second, the descriptions for male- and female-dominated occupations differed from 1800 to 2008, some of which seem contradictory to our expectations of women's social roles and behavior patterns. For example, female-dominated occupations were always associated with religion-related personality trait terms (i.e., *religious*, *moral*, and *ethical*), which may reinforce women's obedience and confine them to the patriarchal society, either consciously or subconsciously. Also, the trends of the use of Extraversion, Openness, and Conscientiousness traits seem to indicate that despite more women being given access to jobs, problems still existed with female workers' job satisfaction. At this point, gender equality still has a long way to go. It is our hope that the insights gained from this study can take a step closer to gender equality.

Third, this study followed the analyses of Schulz and Bahník (2019) and Ye et al. (2018), and they both acknowledged the limitations of the result interpretation that failed to consider the historical changes. The present study is, however, embedded in the historical context, as we tried to explain the results according to the historical background and sociocultural environment of the times.

Last, from the methodological perspective, we used dependency-based method to search and extract target ngrams in the present study, which provides researchers not only efficient retrieval of the data but also accurate results. Future studies may consider the method in similar scenarios such as in lexicon-based research of other socio-psychological traits.

The dataset used in this study contains only texts from books. Hence, future studies may extend our work and retrieve the data from other types of texts such as news, social media, and journal articles, in order to paint a fuller picture of gender differences in workplace contexts. In addition, our study is also limited in that it only focuses on gender differences in personalities. Future research can examine the differences in the big five traits between other groups such as people of different races or ages, or focus on the big five personalities concerning groups such as Asians and Mexican immigrants.

**Research ethics:** Not applicable.

**Informed consent:** Not applicable.

**Author contributions:** QZ: methodology, visualization, writing – original draft, writing - review & editing. ZL: writing – review & editing, supervision. LL: methodology, visualization, writing – original draft, writing – review & editing, supervision, funding acquisition. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Use of Large Language Models, AI and Machine Learning Tools:** None declared.

**Conflict of interest:** The authors state no conflict of interest.

**Research funding:** This work is supported by the Major Research Grant of Shanghai International Studies University (Grant number: 23ZD011).

**Data availability:** The data used in the study are open-sourced data. The authors have stated the source of the data in the manuscript.

## References

- Adelmann, Pamela K., Toni C. Antonucci, Susan E. Crohan, and Lerita M. Coleman. 1989. "Empty Nest, Cohort, and Employment in the Well-Being of Midlife Women." *Sex Roles* 20 (3–4): 173–89.
- Allport, Gordon W., and Henry S. Odbert. 1936. "Trait-names: A Psycho-Lexical Study." *Psychological Monographs* 47 (1): i–171.
- Anglim, Jeremy, and Peter O'Connor. 2019. "Measurement and Research Using the Big Five, HEXACO, and Narrow Traits: A Primer for Researchers and Practitioners." *Australian Journal of Psychology* 71 (1): 16–25.
- Ashton, Michael C., K. Lee, M. Perugini, P. Szarota, R. E. de Vries, L. Di Blas, K. Boies, and B. De Raad. 2004. "A Six-factor Structure of Personality-descriptive Adjectives: Solutions from Psycholexical Studies in Seven Languages." *Journal of Personality and Social Psychology* 86 (2): 356–66.
- Ashton, Michael C., Sampo V. Paunonen, Edward Helmes, and Douglas N. Jackson. 1998. "Kin Altruism, Reciprocal Altruism, and the Big Five Personality Factors." *Evolution and Human Behavior* 19 (4): 243–55.
- Bahramitash, Roksana. 2004. "Myths and Realities of the Impact of Political Islam on Women: Female Employment in Indonesia and Iran." *Development in Practice* 14 (4): 508–20.
- Barrick, Murray R., and Michael K. Mount. 1991. "The Big Five Personality Dimensions and Job Performance: A Meta-analysis." *Personnel Psychology* 44 (1): 1–26.
- Bleidorn, Wiebke, and Christopher James Hopwood. 2019. "Using Machine Learning to Advance Personality Assessment and Theory." *Personality and Social Psychology Review* 23 (2): 190–203.
- Coker, Linda Anne, Douglas B. Samuel, and Thomas A. Widiger. 2002. "Maladaptive Personality Functioning Within the Big Five and the Five-Factor Model." *Journal of Personality Disorders* 16 (5): 385–401.
- Correa, Teresa, Amber Willard Hinsley, and Homero Gil De Zúñiga. 2010. "Who Interacts on the Web? The Intersection of Users' Personality and Social Media Use." *Computers in Human Behavior* 26 (2): 247–53.
- Costa, Paul T., and Robert R. McCrae. 1980. "Influence of Extraversion and Neuroticism on Subjective Well-Being: Happy and Unhappy People." *Journal of Personality and Social Psychology* 38 (4): 668–78.
- De Raad, Boele. 1998. "Five Big, Big Five Issues: Rationale, Content, Structure, Status, and Crosscultural Assessment." *European Psychologist* 3 (2): 113–24.
- Dijkstra, Miranda, Dirk van Dierendonck, Arne Evers, and Carsten K. W. De Dreu. 2005. "Conflict and Well-Being at Work: The Moderating Role of Personality." *Journal of Managerial Psychology* 20 (2): 87–104.
- Fischer, Ronald, J. A. Karl, M. Luczak-Roesch, V. H. Fetvadjev, and A. Grener. 2020. "Tracing Personality Structure in Narratives: A Computational Bottom-Up Approach to Unpack Writers, Characters, and Personality in Historical Context." *European Journal of Personality* 34 (5): 917–43.

- Franck, Raphaël, and Laurence R. Iannaccone. 2014. "Religious Decline in the 20th Century West: Testing Alternative Explanations." *Public Choice* 159 (3-4): 385–414.
- Frankenbach, Julius, Tim Wildschut, Juhl Jacob, and Constantine Sedikides. 2021. "Does Neuroticism Disrupt the Psychological Benefits of Nostalgia? A Meta-Analytic Test." *European Journal of Personality* 35 (2): 249–66.
- Gerber, Alan S., G. A. Huber, D. Doherty, and C. M. Dowling. 2011. "The Big Five Personality Traits in the Political Arena." *Annual Review of Political Science* 14 (1): 265–87.
- Goldberg, Lewis R. 1990. "An Alternative 'Description of Personality': The Big-Five Factor Structure." *Journal of Personality and Social Psychology* 59 (6): 1216–29.
- Goldberg, Lewis R. 1992. "The Development of Markers for the Big-Five Factor Structure." *Psychological Assessment* 4 (1): 26–42.
- Goldberg, Yoav, and Jon Orwant. 2013. "A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books." In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics (SEM)\**, Vol. 1, 241–7. Atlanta, Georgia, USA.
- Hamilton, Mykol C., David Anderson, Michelle Broadus, and Kate Young. 2006. "Gender Stereotyping and Under-Representation of Female Characters in 200 Popular Children's Picture Books: A Twenty-First Century Update." *Sex Roles* 55 (11–12): 757–65.
- Hogan, Robert. 1986. *Hogan Personality Inventory Manual*. Minneapolis: National Computer Systems.
- Huang, Hong, Z. Wang, F. Xia, X. Shang, Y. Liu, et al. 2017. "Water Quality Trend and Change-Point Analyses Using Integration of Locally Weighted Polynomial Regression and Segmented Regression." *Environmental Science and Pollution Research* 24 (18): 827–15837.
- Iimura, Shuhei, and Kanako Taku. 2018. "Gender Differences in Relationship Between Resilience and Big Five Personality Traits in Japanese Adolescents." *Psychological Reports* 121 (5): 920–31.
- Jakubiak, M., and Sheila Murphy. 1986. "Feminism and Religion." *Counseling and Values* 31 (2): 157–64.
- Jorm, Anthony F. 1987. "Sex Differences in Neuroticism: A Quantitative Synthesis of Published Research." *Australian & New Zealand Journal of Psychiatry* 21 (4): 501–6.
- Kelly, Kimberly, and Linda Grant. 2012. "Penalties and Premiums: The Impact of Gender, Marriage, and Parenthood on Faculty Salaries in Science, Engineering and Mathematics (SEM) and Non-SEM Fields." *Social Studies of Science* 42 (6): 869–96.
- Koch, Amanda J., Susan D. D'Mello, and R. Sackett Paul. 2015. "A Meta-Analysis of Gender Stereotypes and Bias in Experimental Simulations of Employment Decision Making." *Journal of Applied Psychology* 100 (1): 128–61.
- Lee, Jackie F. K., and Andy C. O. Chin. 2019. "Are Females and Males Equitably Represented? A Study of Early Readers." *Linguistics and Education* 49: 52–61.
- Lei, Lei, and Ying Zhang. 2018. "Lack of Improvement in Scientific Integrity: An Analysis of WOS Retractions by Chinese Researchers (1997–2016)." *Science and Engineering Ethics* 24: 1409–20.
- Lin, Zhong, and Lei Lei. 2020. "The Research Trends of Multilingualism in Applied Linguistics and Education (2000–2019): A Bibliometric Analysis." *Sustainability* 12 (15): 1–14.
- Liu, Chunnian, Tian Qi, and Mengqiu Chen. 2021. "Distinguishing Personality Recognition and Quantification of Emotional Features Based on Users' Information Behavior in Social Media." *Journal of Database Management* 32 (2): 76–91.
- Liu, Wei-Yin, B. Weber, M. Reuter, S. Markett, W. C. Chu, and C. Montag. 2013. "The Big Five of Personality and Structural Imaging Revisited: A VBM–DARTEL Study." *Neuroreport* 24 (7): 375–80.
- Lyon, K. A., R. Elliott, W. Kerry Ann, J. Gabriella, and B. Lje. 2021. "Associations between Facets and Aspects of Big Five Personality and Affective Disorders: A Systematic Review and Best Evidence Synthesis." *Journal of Affective Disorders* 288: 175–88.

- Mac Giolla, Erik, and Petri J. Kajonius. 2019. "Sex Differences in Personality Are Larger in Gender Equal Countries: Replicating and Extending a Surprising Finding." *International Journal of Psychology* 54 (6): 705–11.
- Malouff, John M., E. B. Thorsteinsson, S. E. Rooke, and N. S. Schutte. 2007. "Alcohol Involvement and the Five-Factor Model of Personality: A Meta-Analysis." *Journal of Drug Education* 37 (3): 277–294.
- Mangion, Carmen. 2005. "'Good Teacher' or 'Good Religious'? The Professional Identity of Catholic Women Religious in Nineteenth-Century England and Wales." *Women's History Review* 14 (2): 223–42.
- Mayer, Gregg. 2007. "Prosecutors in Books: Examining a Literary Disconnect from the Prosecution Function." *Law & Literature* 19 (1): 77–101.
- Morgan, Sue, ed. 2002. *Women, Religion, and Feminism in Britain, 1750–1900*. 1750–1900. Palgrave Macmillan.
- Morgeson, Frederick P., Matthew H. Reider, and Michael A. Campion. 2005. "Selecting Individuals in Team Settings: The Importance of Social Skills, Personality Characteristics, and Teamwork Knowledge." *Personnel Psychology* 58 (3): 583–611.
- Motschenbacher, Heiko, and Eka Roivainen. 2020. "Personality Traits, Adjectives and Gender." *Journal of Language and Discrimination* 4 (1): 16–50.
- Muggeo, Vito M. R. 2008. "Segmented: An R Package to Fit Regression Models with Broken-Line Relationships." *R News* 8 (1): 20–5.
- Neal, Andrew, G. Yeo, A. Koy, and T. Xiao. 2012. "Predicting the Form and Direction of Work Role Performance from the Big 5 Model of Personality Traits." *Journal of Organizational Behavior* 33 (2): 175–92.
- Nezhad, Masih Rahimi, and M. Rahimi Nezhad. 2021. "Investigation on Productivity of Synonym Words with Different Semantic Prosody in English." *International Journal of Linguistics and Translation Studies* 2 (3): 65–75.
- Norman, Warren T. 1967. *2800 Personality Trait Descriptors: Normative Operating Characteristics for a University Population*. Ann Arbor: University of Michigan, Department of Psychology.
- Paulhus, Delroy L., and Kevin M. Williams. 2002. "The Dark Triad of Personality: Narcissism, Machiavellianism, and Psychopathy." *Journal of Research in Personality* 36 (6): 556–63.
- Peabody, Dean, and Lewis R. Goldberg. 1989. "Some Determinants of Factor Structures from Personality-Trait Descriptors." *Journal of Personality and Social Psychology* 57 (3): 552–67.
- Poropat, Arthur E. 2009. "A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance." *Psychological Bulletin* 135 (2): 322–38.
- Qiu, Lin, H. Lin, J. Ramsay, and F. Yang. 2012. "You Are What You Tweet: Personality Expression and Perception on Twitter." *Journal of Research in Personality* 46 (6): 710–18.
- Renzulli, Linda A., J. Reynolds, K. Kelly, and L. Grant. 2013. "Pathways to Gender Inequality in Faculty Pay: The Impact of Institution, Academic Division, and Rank." *Research in Social Stratification and Mobility* 34: 58–72.
- Reskin, Barbara. 1993. "Sex Segregation in the Workplace." *Department of Sociology* 19: 241–70.
- Reskin, Barbara F., and Padavic Irene. 1988. "Supervisors as Gatekeepers: Male Supervisors' Response to Women's Integration in Plant Jobs." *Social Problems* 35 (5): 536–50.
- Roivainen, Eka. 2013. "Frequency of the Use of English Personality Adjectives: Implications for Personality Theory." *Journal of Research in Personality* 47 (4): 417–20.
- Roivainen, Eka. 2015. "Personality Adjectives in Twitter Tweets and in the Google Books Corpus. An Analysis of the Facet Structure of the Openness Factor of Personality." *Current Psychology* 34 (4): 621–25.

- Roivainen, Eka. 2020. "Generational Changes in Personality: The Evidence from Corpus Linguistics." *Psychological reports* 123 (2): 325–40.
- Rosen, Bernard Carl. 1998. *Winners and Losers of the Information Revolution: Psychosocial Change and Its Discontents*. Greenwood Publishing Group.
- Roth, Steffen, V. Valentinov, A. Augustinaitis, A. Mkrtchyan, and J. Kaivo-oja. 2018. "Was That Capitalism? A Future-Oriented Big Data Analysis of the English Language Area in the 19th and 20th Century." *Futures* 98: 41–8.
- Saucier, Gerard, and Lewis R. Goldberg. 1996. "Evidence for the Big Five in Analyses of Familiar English Personality Adjectives." *European Journal of Personality* 10 (1): 61–77.
- Schulz, Daniel, and Štěpán Bahník. 2019. "Gender Associations in the Twentieth-Century English-Language Literature." *Journal of Research in Personality* 81: 88–97.
- Seyfarth, Robert M., and Dorothy L. Cheney. 1984. "Grooming, Alliances and Reciprocal Altruism in Vervet Monkeys." *Nature* 308: 541–543.
- Shi, Yaqian, and Lei Lei. 2021. "Lexical Use and Social Class: A Study on Lexical Richness, Word Length, and Word Class in Spoken English." *Lingua* 262: 103155.
- Spielberger, Charles D., and Richard L. Rickman. 1990. "Assessment of State and Trait Anxiety." In *Anxiety: Psychobiological and Clinical Perspectives*, edited by Nicolò Sartorius, Vittorio Andreoli, Giovanni Cassano, Leon Eisenberg, Patricia Kielholz, Paolo Pancheri, and Giulio Racagni, 69–83. New York: Hemisphere Publishing.
- Srivastava, Sanjay. 2010. "The Five-Factor Model Describes the Structure of Social Perceptions." *Psychological Inquiry* 21 (1): 69–75.
- Topino, Eleonora, A. Di Fabio, L. Palazzeschi, and A. Gori. 2021. "Personality Traits, Workers' Age, and Job Satisfaction: The Moderated Effect of Conscientiousness." *PLoS ONE* 16 (7): e0252275.
- Twenge, Jean M. 2000. "The Age of Anxiety? The Birth Cohort Change in Anxiety and Neuroticism, 1952–1993." *Journal of Personality and Social Psychology* 79 (6): 1007–21.
- U.S. Bureau of Labor Statistics. 2021. *Women in Labor Force: A Databook*. Retrieved from: <https://www.bls.gov/opub/reports/womens-databook/2020/home.htm>.
- Van Aarde, Ninette, Deon Meiring, and Brenton M. Wiernik. 2017. "The Validity of the Big Five Personality Traits for Job Performance: Meta-Analyses of South African Studies." *International Journal of Selection and Assessment* 25 (3): 223–39.
- Vedel, Anna. 2014. "The Big Five and Tertiary Academic Performance: A Systematic Review and Meta-Analysis." *Personality and Individual Differences* 71: 66–76.
- Weisberg, Yanna J., Colin G. DeYoung, and B. Hirsh. Jacob. 2011. "Gender Differences in Personality across the Ten Aspects of the Big Five." *Frontiers in Psychology* 2: 178.
- Wen, Ju, and Lei Lei. 2022. "Linguistic Positivity Bias in Academic Writing: A Large-Scale Diachronic Study in Life Sciences across 50 Years." *Applied Linguistics* 43 (2): 340–64.
- Ye, Shenglu, S. Cai, C. Chen, Q. Wan, and X. Qian. 2018. "How Have Males and Females Been Described over the Past Two Centuries? An Analysis of Big-Five Personality-Related Adjectives in the Google English Books." *Journal of Research in Personality* 76: 6–16.

---

**Supplementary Material:** This article contains supplementary material (<https://doi.org/10.1515/dsll-2025-0010>).

## Bionotes

**Zihan Qiu**

School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China

Zihan Qiu is an MA graduate in applied linguistics at the School of Foreign Languages, Shanghai Jiao Tong University. Her research interests include corpus linguistics, quantitative linguistics, and academic English.

**Li Zhang**

School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China

Li Zhang is Professor of Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University. Her research interests include second language acquisition, AI-enhanced language teaching and learning, and academic writing. She currently leads a key project under the National Social Science Fund of China, titled “The construction and application of an adaptive writing learning model based on generative AI”.

**Lei Lei**

Institute of Language Sciences, Shanghai International Studies University, Shanghai, China

**leileicn@126.com**

Lei Lei is Professor of Applied Linguistics at the Institute of Corpus Studies and Applications, Shanghai International Studies University. His research interests are applied linguistics, second language writing, and academic English. He has published extensively in journals such as *Applied Linguistics*, *Language Teaching*, *International Journal of Corpus Linguistics*, *Journal of English for Academic Purposes*, and *System*.