

Kathleen Ahrens, Winnie Huiheng Zeng, Christian Burgers\* and Chu-Ren Huang

# Metaphor and gender: are words associated with source domains perceived in a gendered way?

## 隱喻與性別: 隱喻源域詞感知的性別化差異

<https://doi.org/10.1515/lingvan-2024-0021>

Received December 13, 2022; accepted March 4, 2024; published online July 11, 2024

**Abstract:** Conceptual metaphors are one of many linguistic devices that can potentially encode and reinforce gender stereotypes. However, little is known about how metaphors encode gender stereotypes, and in previous literature the concept of “gendered metaphor” has been mostly assumed rather than attested. We take the first step to tackle this issue by examining the gender typicality of specific metaphorical source domains. In the present paper, we conducted three rating experiments ( $N_{total} = 1,060$  English-speaking participants) to determine the genderedness of 50 keywords associated with five frequently used source domains (BUILDING, COMPETITION, JOURNEY, PLANT, and WAR). We found that keywords associated with three source domains (BUILDING, COMPETITION, and WAR) were viewed as more masculine, while keywords associated with the source domains of JOURNEY and PLANT were viewed as more feminine. These data offer empirical verification for gendered perceptions of keywords associated with some frequently used source domains. The result also provides the first evidence that metaphors could encode gender stereotypes by selection of source domains.

**Keywords:** conceptual metaphor theory; language and gender; source domain

**摘要:** 概念隱喻被認為是一種能標註刻板定型的語言機製。然而, 罕見有文獻聚焦於隱喻如何標註性別刻板定型的話題。目前的文獻定義性別化隱喻時大多基於主觀假設而非客觀驗證。本文邁出第一步, 通過研究特定隱喻源域是否與特定性別相關以及相關度去解決這個問題。本研究招聘了1,060名英語母語者參與三個評分實驗, 用以測試與「建築」、「競爭」、「旅程」、「植物」和「戰爭」這五個常用源域相關的50個關鍵隱喻詞的性別化程度。研究結果顯示, 與「建築」、「競爭」和「戰爭」源域相關的隱喻被視為更具男性特征, 而與「旅程」和「植物」源域相關的隱喻被視為更具女性特征。本文實證了五個常用隱喻源的不同性別化特征, 提供了隱喻經由源域的選擇標註刻板定型的初步驗證。

**關鍵詞:** 概念隱喻理論; 語言和性別; 源域

## 1 Introduction

Conceptual metaphors involve conventional cross-domain mappings that connect a number of elements from a source domain, such as JOURNEY, to a target domain, such as LIFE (Lakoff and Johnson 2003). For example, consider

\*Corresponding author: Christian Burgers, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, Netherlands, E-mail: c.f.burgers@uva.nl. <https://orcid.org/0000-0002-5652-9021>

Kathleen Ahrens, Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong SAR, China, E-mail: kathleen.ahrens@polyu.edu.hk. <https://orcid.org/0000-0002-7863-3655>

Winnie Huiheng Zeng, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University & Department of English Language and Literature, Hong Kong Shue Yan University, Hong Kong SAR, China, E-mail: winnie.zeng@connect.polyu.hk. <https://orcid.org/0000-0001-7565-2650>

Chu-Ren Huang, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China, E-mail: churen.huang@polyu.edu.hk. <https://orcid.org/0000-0002-8526-5520>

an expression such as *I don't know which path to take after graduation*. In this sentence, *path* is part of the JOURNEY source domain, and the idea being expressed is that, in the next part of this person's life, they have a variety of options available to them. Recent work (Ahrens and Jiang 2020) has demonstrated that it is possible to use corpus-based tools to verify whether or not a given lexical item, such as *path*, belongs in a given source domain, such as JOURNEY.

Various studies have demonstrated that some conceptual metaphors have been associated with reinforcing gender stereotypes (e.g., Friedman 1987; Hines 1999; Koller 2004a, 2004b; Murphy 2001; Philip 2009; Velasco Sacristán 2005; Zeng et al. 2020). In a series of studies using Amazon Mechanical Turk, Elmore and Luna-Lucero (2017) analyzed how men and women were perceived when talking about scientific discoveries. They demonstrated that men's discoveries were considered more important when they were described as coming to them in a eureka moment – that is, using the source domain of LIGHT BULB – but not when they were described using the source domain of SEEDLING. However, the opposite held true for women. The authors conclude that describing and discussing scientific discoveries as having a “light bulb” moment in textbooks is associated with a characteristic of men and hinders people from viewing women's discoveries as equally valuable to those of men.

In the political realm, Gidengil and Everitt (2003) demonstrated that women in Canadian politics who adopted “masculine” metaphors were criticized by the media for being too aggressive, even though they did not use as many masculine metaphors as men did. In the business realm, Koller (2004a) demonstrated that women are also criticized when they step out of their metaphorized restraints. In her study, she found that women were more often described in terms of WAR metaphors than men, as if the journalists were remarking upon the fact that women embody these “cutthroat” characteristics, ones not usually performed by “feminine” women. Koller (2004a) also noted that even though women were described more often in terms of WAR metaphors than men, they were also described more often by the CAREGIVER metaphor than by the ATHLETE metaphors, indicating mixed messages in how women have been conceptualized in the business world. These studies argue that women are simultaneously categorized and constrained by the metaphors they use. However, an important caveat with these earlier studies is that they are based on observational data. Winter et al. (2020), however, ran an experimental study which looked at metaphors of power in terms of the source domain of verticality. They found that men had stronger vertical preferences than women and attributed this finding to men preferring to think about power in terms of their own bodily experiences, particularly those related to physical strength. Studies have also shown gender-specific perceptions of power with men viewing aggression and bodily force as instrumental to gaining power while women may often view bodily force as a breakdown of self-control and a loss of power (Campbell et al. 1992; Schubert 2004).

The goal of the current study is to further understand the extent to which keywords associated with common source domains are potentially gendered. While previous work looked at gender stereotypicality across varying numbers of nouns (Crawford et al. 2004; Misersky et al. 2014; Scott et al. 2019), these studies either did not include all keywords we include here or did not examine differences between male and female participants. Our study differs from these previous works by focusing on the gender stereotypicality of keywords associated with five metaphorical source domains (WAR, COMPETITION, JOURNEY, BUILDING, and PLANT) as determined by the source domain verification procedure (Ahrens and Jiang 2020) and examining these associations for both men and women, similar to Winter et al. (2020). We selected these five source domains as they have been researched extensively over the past 40 years (Lakoff and Johnson 2003), have a large number of keywords associated with them, and are used frequently in spoken and written language (Charteris-Black 2004; Deignan 1995; Kövecses 2010; Zeng and Ahrens 2023).

The research questions (RQs) we address in this study are:

**RQ1:** To what extent are keywords associated with the source domains of (a) WAR, (b) COMPETITION, (c) JOURNEY, (d) BUILDING, and (e) PLANT perceived as masculine or feminine?

**RQ2:** To what extent do the (a) perceived masculinity and (b) perceived femininity of these keywords differ between male and female participants?

## 2 Methods

We designed three studies to measure the strength of association between keywords associated with source domains for masculinity (Studies 1 and 3) and femininity (Studies 2 and 3). Three corresponding surveys were created for collecting rating data from online participants.

### 2.1 Stimuli

When designing the stimuli, we looked to previous metaphor research (Lakoff and Johnson 2003), dictionaries such as the Collins Cobuild metaphor dictionary (Deignan 1995), and source domains identified in professional contexts (e.g., Charteris-Black 2004, 2006, 2011). We selected five frequently used source domains (BUILDING, COMPETITION, JOURNEY, PLANT, and WAR) that had a large number of conventionally used instances of metaphorical expressions. Ten keywords from these conventionally used instances were selected for each source domain (Table 1) after ascertaining they belonged to their respective source domain category using the source domain verification methodology in Ahrens and Jiang (2020). Five additional keywords (*boy*, *girl*, *man*, *woman*, and *lady*) were also included in the experiments as attention checks (see Appendix A in the Online Supplementary Materials).

### 2.2 Procedure and instrumentation

The data were collected through the online survey platform SurveyMonkey ([www.surveymonkey.com](http://www.surveymonkey.com)).<sup>1</sup> The order of the 55 keywords were randomized by participant and presented along with a 7-point Likert scale in the three surveys. Participants were recruited through Amazon Mechanical Turk (<https://www.mturk.com>), an online questionnaire program that allows researchers to collect ratings data from participants who meet certain criteria (Buhrmester et al. 2011; Paolacci et al. 2010). The workers on Mechanical Turk were directed to the surveys we designed on SurveyMonkey. All the participants needed to have an approval rating of 95 % or better on Mechanical Turk. Attention checks were also given after the main task to ensure the participants had carefully read the stimuli (Keith et al. 2017).

**Table 1:** Stimuli items across the five source domains.

| BUILDING            | COMPETITION         | JOURNEY            | PLANT            | WAR             |
|---------------------|---------------------|--------------------|------------------|-----------------|
| <i>building</i>     | <i>competition</i>  | <i>journey</i>     | <i>plant</i>     | <i>war</i>      |
| <i>architecture</i> | <i>championship</i> | <i>destination</i> | <i>blossom</i>   | <i>army</i>     |
| <i>buttress</i>     | <i>contest</i>      | <i>direction</i>   | <i>cultivate</i> | <i>assault</i>  |
| <i>construction</i> | <i>game</i>         | <i>path</i>        | <i>fertilize</i> | <i>battle</i>   |
| <i>foundation</i>   | <i>team</i>         | <i>progress</i>    | <i>fruit</i>     | <i>combat</i>   |
| <i>framework</i>    | <i>race</i>         | <i>road</i>        | <i>harvest</i>   | <i>enemy</i>    |
| <i>structure</i>    | <i>rivalry</i>      | <i>route</i>       | <i>reap</i>      | <i>military</i> |
| <i>support</i>      | <i>sport</i>        | <i>travel</i>      | <i>root</i>      | <i>skirmish</i> |
| <i>roof</i>         | <i>tournament</i>   | <i>trip</i>        | <i>seed</i>      | <i>weapon</i>   |
| <i>wall</i>         | <i>medalist</i>     | <i>way</i>         | <i>sprout</i>    | <i>warrior</i>  |

<sup>1</sup> The studies were conducted in accordance with the Human Subjects Ethics Sub-committee of the Hong Kong Polytechnic University (HSESC Reference Number: HSEARS20200827002).

The procedure for the three studies was as follows:

- Study 1: Participants in Study 1 were instructed to rate the extent to which the words listed in the survey are associated with masculinity, which was defined as “quality or attributes regarded as characteristics of men”. If they thought the keyword was weakly associated with masculinity, they were instructed to select 1. If they believed the keyword was strongly associated with masculinity, they were asked to select 7. If they thought the keyword was somewhere in between “weakly associated” and “strongly associated” with masculinity, they were asked to select the number on the scale that, in their opinion, best represented the strength of association. If they did not know the meaning of the word, they were asked to select 8.
- Study 2: Participants in Study 2 were asked to rate the extent to which the words in the survey are associated with femininity, which was defined as “qualities or attributes regarded as characteristics of women”, using a similar scale to that in Study 1.
- Study 3: Participants in Study 3 had to rate the extent to which the words in the survey are associated with either masculinity or with femininity as defined in Studies 1 and 2. If they thought that the keyword was strongly associated with masculinity, they were asked to select 1. If they felt that the keyword was strongly associated with femininity, they were asked to select 7. If, in their opinion, the keyword was somewhere in between “strongly associated with masculinity” and “strongly associated with femininity”, they were asked to select the number on the scale that best represented the strength of association. If they did not know the meaning of the word, they were asked to select 8. Screenshots of the sample survey scales of the three studies can be seen in Appendix B in the Online Supplementary Materials.

After completing the ratings task, participants were asked several demographic questions regarding their gender, age, language background, and nationality. In exchange for participation, each participant was paid US\$1.00.

## 2.3 Participants

We set specific inclusion criteria of participants as follows: we only included participants who completed the entire questionnaire, reported a US nationality, grew up only speaking English, and identified as either male or female. Furthermore, we checked the ratings on the five attention check items and calculated the means and standard deviations, dropping cases where:

- participants had missing values on at least one of the attention check items (*man, boy, girl, lady, woman*), indicating that they did not know these (standard) English words
- participants had a standard deviation of less than 1 on the five attention check items combined (*man, boy, girl, lady, woman*), indicating that they did not differentiate between explicitly male and female labels
- participants scored the mean of the masculine words (*man, boy*) lower on masculinity than the mean of the feminine words (*lady, girl, woman*) in Study 1; participants scored the mean of the masculine words (*man, boy*) higher on femininity than the mean of the feminine words (*lady, girl, woman*) in Studies 2 and 3.

In Study 1, 538 participants were recruited, of whom 228 participants met the above inclusion criteria (42.4 %). Their average age was 39.04 years old (SD = 11.73). A total of 87 participants self-identified as female (38.2 %) and 141 participants self-identified as male (61.8 %).

In Study 2, 537 participants completed the study, of whom 354 participants met all inclusion criteria (65.9 %). Their average age was 40.53 years old (SD = 12.35). A total of 184 participants self-identified as female (52.0 %) and 170 participants self-identified as male (48.0 %).

Out of the 593 participants who completed the survey in Study 3, 478 met all inclusion criteria (80.6 %). Their average age was 40.66 years old (SD = 13.24). A total of 281 participants self-identified as female (58.8 %) and 197 participants self-identified as male (41.2 %).

The numbers of participants (de)selected split by the different (de)selection criteria are provided in Appendix C in the Online Supplementary Materials. Given that more participants were excluded in Studies 1–2 than in Study

3, we ran post hoc power analyses for all three studies that demonstrated that power was highly sufficient (100 % for Study 1; 91.7 % for Study 2; and 99.5 % for Study 3) for all three studies (see Appendix D in the Online Supplementary Materials for details).

## 2.4 Data analysis

All data cleaning and analyses were conducted using R (R Core Team 2021) with the package lme4 (Bates et al. 2015). For all three studies, the likelihood ratio test demonstrated that the best fit was obtained for a model with fixed effects for source domain, participant gender, and their interaction; random intercepts for participants and words; and random slopes for participants. We obtained the main effects of source domain and participants' gender and examined potential interactions between source domain and gender. Data and analysis code are available on the Open Science Framework at <https://osf.io/x2zu3/>.

## 3 Results and discussion

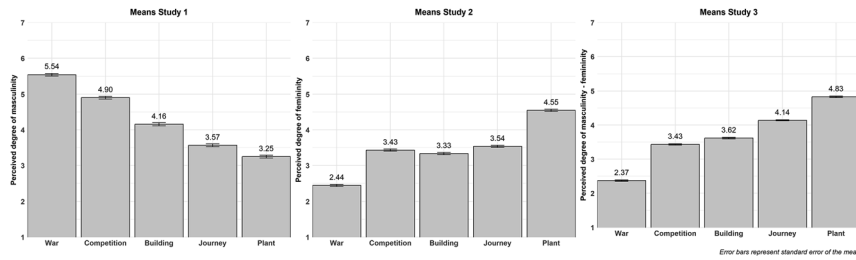
In this section, we report the findings of the main effects between source domains, as well as the interaction of source domain and participant gender.

### 3.1 Control analyses

First, we conducted several control analyses to see whether the domains that we chose were indeed associated with masculinity or femininity. For this to be the case, each domain needed to have a score that was significantly different from the scale's midpoint of 4. To test for this, we conducted one-sample *t*-tests for each source domain. These analyses reveal that the keywords in the domains were significantly different from the scales' midpoint. Specifically, in Study 1, the keywords in the domains of WAR ( $t(2,243) = 42.60, p < 0.001$ ), PLANT ( $t(2,252) = -18.65, p < 0.001$ ), JOURNEY ( $t(2,261) = -10.73, p < 0.001$ ), COMPETITION ( $t(2,266) = 24.23, p < 0.001$ ), and BUILDING ( $t(2,236) = 3.68, p < 0.001$ ) were rated significantly different from 4. In Study 2, the keywords in the domains of WAR ( $t(3,501) = -52.95, p < 0.001$ ), PLANT ( $t(3,509) = 17.24, p < 0.001$ ), JOURNEY ( $t(3,531) = -15.35, p < 0.001$ ), COMPETITION ( $t(3,526) = -18.47, p < 0.001$ ), and BUILDING ( $t(3,451) = -19.78, p < 0.001$ ) were rated significantly different from 4. In Study 3, the keywords in the domains of WAR ( $t(4,698) = -79.96, p < 0.001$ ), PLANT ( $t(4,467) = 38.39, p < 0.001$ ), JOURNEY ( $t(4,686) = 8.80, p < 0.001$ ), COMPETITION ( $t(4,249) = -28.94, p < 0.001$ ), and BUILDING ( $t(4,576) = -17.01, p < 0.001$ ) were rated significantly different from 4. More detailed *t*-test results can be found in Appendix E in the Online Supplementary Materials.

### 3.2 Main effects of source domains

Figure 1 shows the means for the keywords associated with the different source domains across the different studies, and Table 2 shows results from the linear mixed-effects analyses (see also Appendices F and G in the Online Supplementary Materials). Table 2 demonstrates that, regardless of source domain, men and women did not differ in their perceptions of perceived masculinity (Study 1:  $b = 0.19, p = 0.238$ ) and femininity (Study 2:  $b = -0.17, p = 0.179$ ). However, in Study 3, men rated the keywords regardless of source domain as significantly more feminine compared to women ( $b = -0.19, p = 0.001$ ). The overview with mean values of masculinity (Study 1), femininity (Study 2), and masculinity-femininity (Study 3) for each keyword can be found in Appendix H in the Online Supplementary Materials.



**Figure 1:** Means of perceived masculinity (Study 1), perceived femininity (Study 2), and perceived masculinity-femininity (Study 3) across the different source domains.

**Table 2:** Fixed effects estimates and variance-covariance estimates for the models predicting perceived masculinity and perceived femininity (SD = source domain).

| Predictors                        | Perceived masculinity (Study 1) |                |        | Perceived femininity (Study 2) |               |        | Perceived masculinity-femininity (Study 3) |               |        |
|-----------------------------------|---------------------------------|----------------|--------|--------------------------------|---------------|--------|--------------------------------------------|---------------|--------|
|                                   | Estimates                       | 99 % CI        | p      | Estimates                      | 99 % CI       | p      | Estimates                                  | 99 % CI       | p      |
| (Intercept)                       | 5.41                            | 4.86 to 5.96   | <0.001 | 2.53                           | 1.99 to 3.07  | <0.001 | 2.46                                       | 2.06 to 2.85  | <0.001 |
| SD [competition]                  | -0.71                           | -1.36 to -0.07 | 0.004  | 1.02                           | 0.31 to 1.73  | <0.001 | 1.04                                       | 0.47 to 1.60  | <0.001 |
| SD [building]                     | -1.57                           | -2.22 to -0.93 | <0.001 | 0.87                           | 0.16 to 1.57  | 0.002  | 1.19                                       | 0.63 to 1.74  | <0.001 |
| SD [journey]                      | -2.24                           | -2.89 to -1.60 | <0.001 | 1.17                           | 0.46 to 1.88  | <0.001 | 1.79                                       | 1.24 to 2.34  | <0.001 |
| SD [plant]                        | -2.52                           | -3.16 to -1.88 | <0.001 | 2.18                           | 1.47 to 2.89  | <0.001 | 2.45                                       | 1.90 to 3.00  | <0.001 |
| Gender [male]                     | 0.19                            | -0.23 to 0.61  | 0.238  | -0.17                          | -0.49 to 0.15 | 0.179  | -0.19                                      | -0.34 to 0.04 | 0.001  |
| SD [competition] * gender [male]  | 0.14                            | -0.08 to 0.36  | 0.092  | -0.07                          | -0.24 to 0.10 | 0.297  | 0.05                                       | -0.08 to 0.18 | 0.330  |
| SD [building] * gender [male]     | 0.32                            | 0.10 to 0.54   | <0.001 | 0.04                           | -0.13 to 0.21 | 0.566  | 0.19                                       | 0.06 to 0.32  | <0.001 |
| SD [journey] * gender [male]      | 0.46                            | 0.24 to 0.68   | <0.001 | -0.16                          | -0.33 to 0.00 | 0.012  | -0.07                                      | -0.19 to 0.06 | 0.181  |
| SD [plant] * gender [male]        | 0.40                            | 0.18 to 0.62   | <0.001 | -0.16                          | -0.33 to 0.21 | 0.013  | 0.07                                       | -0.06 to 0.19 | 0.187  |
| <b>Random effects</b>             |                                 |                |        |                                |               |        |                                            |               |        |
| $\sigma^2$                        | 1.91                            |                |        | 1.86                           |               |        | 1.35                                       |               |        |
| $\tau_{00}$                       | 1.25 <sub>Participant</sub>     |                |        | 1.17 <sub>Participant</sub>    |               |        | 0.26 <sub>Participant</sub>                |               |        |
|                                   | 0.29 <sub>Target Word</sub>     |                |        | 0.37 <sub>Target Word</sub>    |               |        | 0.22 <sub>Target Word</sub>                |               |        |
| ICC                               | 0.45                            |                |        | 0.45                           |               |        | 0.26                                       |               |        |
| N                                 | 228 <sub>Participant</sub>      |                |        | 354 <sub>Participant</sub>     |               |        | 478 <sub>Participant</sub>                 |               |        |
|                                   | 50 <sub>Target Word</sub>       |                |        | 50 <sub>Target Word</sub>      |               |        | 49 <sub>Target Word</sub>                  |               |        |
| Observations                      | 11,263                          |                |        | 17,523                         |               |        | 22,681                                     |               |        |
| Marginal $R^2$ /conditional $R^2$ | 0.180/0.545                     |                |        | 0.120/0.519                    |               |        | 0.271/0.464                                |               |        |

Note: The domain of WAR was the reference category for source domain; female was the reference category for gender.

RQ1 asked about the extent to which the keywords associated with the source domains of (a) WAR, (b) COMPETITION, (c) JOURNEY, (d) BUILDING, and (e) PLANT were perceived as masculine or feminine. Table 2 demonstrates that, in Study 1, the keywords in the domains of COMPETITION ( $b = -0.71$ , 99 % CI  $[-1.36, -0.07]$ ,  $p = 0.004$ ), BUILDING ( $b = -1.57$ , 99 % CI  $[-2.22, -0.93]$ ,  $p < 0.001$ ), JOURNEY ( $b = -2.24$ , 99 % CI  $[-2.89, -1.60]$ ,  $p < 0.001$ ), and PLANT ( $b = -2.52$ , 99 % CI  $[-3.16, -1.88]$ ,  $p < 0.001$ ) were perceived as less masculine compared to the keywords associated with the domain of WAR. Post hoc analyses with Tukey corrections further demonstrate that the keywords associated with the domain of COMPETITION were seen as more masculine compared to the keywords in the domains of BUILDING ( $p = 0.01$ ), JOURNEY ( $p < 0.001$ ) and PLANT ( $p < 0.001$ ). In addition, the keywords in the BUILDING domain were seen as more masculine compared to the keywords associated with the domain of PLANT ( $p = 0.002$ ), but not of JOURNEY ( $p = 0.1021$ ). The keywords associated with the domains of PLANT and JOURNEY did not differ in perceived masculinity ( $p = 0.7209$ ).



These findings indicate that the keywords associated with the WAR and COMPETITION source domains on the one hand and JOURNEY and PLANT source domains on the other hand are highly contrastive with respect to their perceived masculinity.

In Study 2, the keywords associated with the domains of COMPETITION ( $b = 1.02$ , 99 % CI [0.31, 1.73],  $p < 0.001$ ), BUILDING ( $b = 0.87$ , 99 % CI [0.16, 1.57],  $p = 0.002$ ), JOURNEY ( $b = 1.17$ , 99 % CI [0.46, 1.88],  $p < 0.001$ ), and PLANT ( $b = 2.18$ , 99 % CI [1.47, 2.89],  $p < 0.001$ ) were perceived as more feminine compared to the keywords associated with the domain of WAR. Post hoc analyses with Tukey corrections further demonstrate that the keywords in the domain of COMPETITION were seen as less feminine compared to those in the domain of PLANT ( $p < 0.001$ ), but not as compared to those in the domain of BUILDING ( $p = 0.9965$ ) or JOURNEY ( $p = 0.9955$ ). In addition, the keywords associated with the PLANT domain were seen as more feminine compared to those in the domains of BUILDING ( $p < 0.001$ ) and JOURNEY ( $p = 0.0019$ ). The keywords associated with the domains of BUILDING and JOURNEY did not differ in perceived femininity ( $p = 0.9472$ ).

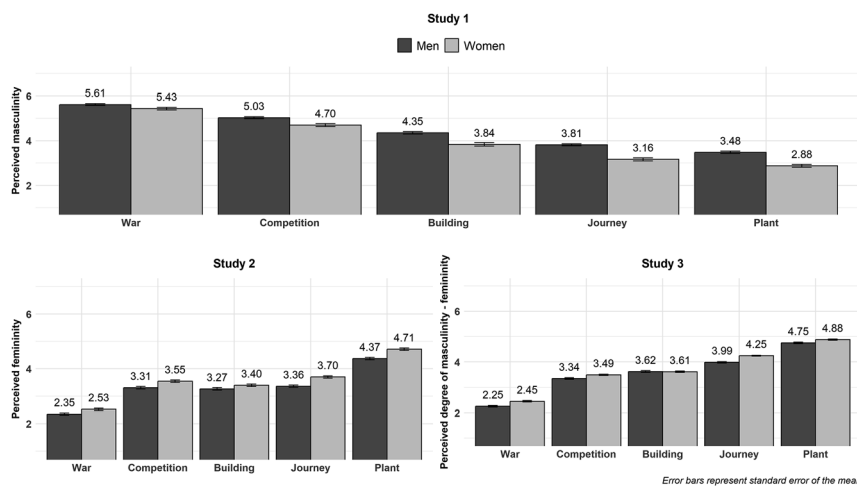
In Study 3, the keywords in the domains of COMPETITION ( $b = 1.04$ , 99 % CI [0.47, 1.60],  $p < 0.001$ ), BUILDING ( $b = 1.19$ , 99 % CI [0.63, 1.74],  $p < 0.001$ ), JOURNEY ( $b = 1.79$ , 99 % CI [1.24, 2.34],  $p < 0.001$ ), and PLANT ( $b = 2.45$ , 99 % CI [1.90, 3.00],  $p < 0.001$ ) were perceived as more feminine compared to those in the domain of WAR. Post hoc analyses with Tukey corrections further demonstrate that the keywords in the domain of COMPETITION were seen as less feminine compared to those in the domains of JOURNEY ( $p = 0.0131$ ) and PLANT ( $p < 0.001$ ), but not compared to BUILDING ( $p = 0.8521$ ). In addition, the keywords in the PLANT domain were seen as more feminine compared to those in the domains of BUILDING ( $p < 0.001$ ) and JOURNEY ( $p = 0.006$ ). The keywords in the domains of BUILDING and JOURNEY did not differ in perceived femininity ( $p = 0.1704$ ).

These analyses reveal how, overall, the keywords associated with different source domains vary with whether they are perceived as masculine or feminine, with the keywords in the WAR domain being perceived as most masculine, followed by those in the COMPETITION source domain. In contrast, the keywords in the PLANT domain were seen as most feminine. This demonstrates that the keywords in the domains related to competition and conflict were perceived as more masculine, while the keywords in the domains related to nurture and growth were seen as more feminine.

### 3.3 Interaction between source domain and gender

RQ2 asked about the extent to which (a) perceived masculinity and (b) perceived femininity of the keywords in the source domains differed between male and female participants (see Figure 2).

Post hoc analyses with Bonferroni corrections showed that, for perceived masculinity (Study 1), men perceived the keywords associated with the source domains of BUILDING ( $b = -0.51$ ,  $SE = 0.16$ ,  $p = 0.0017$ ), COMPETITION



**Figure 2:** Means of perceived masculinity (Study 1), perceived femininity (Study 2), and perceived masculinity-femininity (Study 3) by source domain and participant gender.

( $b = -0.34$ ,  $SE = 0.16$ ,  $p = 0.0399$ ), JOURNEY ( $b = -0.65$ ,  $SE = 0.16$ ,  $p < 0.001$ ), and PLANT ( $b = -0.60$ ,  $SE = 0.16$ ,  $p < 0.001$ ) as more masculine compared to women. Men and women did not differ in their perceptions of masculinity of the keywords in the source domain of WAR ( $b = -0.19$ ,  $SE = 0.16$ ,  $p = 0.2382$ ).

To specify, results of the post hoc analyses with Bonferroni corrections from the linear mixed modeling analysis reveal significant differences in the masculinity ratings for the keywords associated with the source domains of BUILDING ( $p = 0.0017$ ), COMPETITION ( $p = 0.0399$ ), JOURNEY ( $p < 0.001$ ), and PLANT ( $p < 0.001$ ) between men and women, with men considering keywords in these source domains as more masculine. There is no difference in how masculine the keywords in the source domain of WAR are viewed by men and women, which is considered highly masculine by men and women. The clines for both genders are similar, with keywords in the WAR and COMPETITION source domains viewed as being highly masculine while those in the JOURNEY and PLANT metaphors are much less so.

In Study 2, post hoc analysis with Tukey corrections showed that women perceived keywords in the source domains of PLANT ( $b = 0.33$ ,  $SE = 0.12$ ,  $p = 0.0081$ ) and JOURNEY ( $b = 0.33$ ,  $SE = 0.12$ ,  $p = 0.0079$ ) as more feminine than men. Men and women did not differ in their perceived femininity of the keywords in the source domains of WAR ( $b = 0.17$ ,  $SE = 0.12$ ,  $p = 0.1785$ ), COMPETITION ( $b = 0.24$ ,  $SE = 0.12$ ,  $p = 0.0583$ ), and BUILDING ( $b = 0.13$ ,  $SE = 0.12$ ,  $p = 0.2977$ ). In Study 3, we found that women perceived the keywords in the source domains of WAR ( $b = 0.19$ ,  $SE = 0.06$ ,  $p = 0.0013$ ), COMPETITION ( $b = 0.14$ ,  $SE = 0.06$ ,  $p = 0.0191$ ), JOURNEY ( $b = 0.25$ ,  $SE = 0.06$ ,  $p < 0.001$ ), and PLANT ( $b = 0.12$ ,  $SE = 0.06$ ,  $p = 0.0368$ ) as more feminine than men. Men and women did not differ in their perceptions of the keywords in the domain of BUILDING ( $b = -0.001$ ,  $SE = 0.06$ ,  $p = 0.9831$ ).

## 4 General discussion and conclusion

The aim of this study was to investigate whether keywords associated with five source domains (following Ahrens and Jiang 2020) that are often discussed in metaphor studies differed in perceived masculinity and perceived femininity. Across three rating studies, we consistently found that keywords in the source domains of WAR and COMPETITION were viewed as more masculine, while keywords in the PLANT metaphors were considered more feminine. Our rating studies also showed that keywords in the BUILDING and JOURNEY source domains are, as predicted, neither strongly masculine nor feminine, but we did see that there is a preference for keywords in the BUILDING domain to be associated with masculine attributes in contrast to those in the PLANT domain, while keywords in the JOURNEY domain are more associated with femininity, suggesting a subtle difference between these two conceptual domains.

These findings support previous discussions in the metaphor literature, which have proposed WAR, SPORTS, and GAME as masculinized domains that are predominantly used in stereotypically male arena such as business discourse (Koller 2004b: 77). Additionally, Elmore and Luna-Lucero (2017) argued that the SEEDLING metaphor (part of the PLANT source domain) is more likely associated with increasing the perception of female characteristics. Our findings also align with a previous study categorizing the source domains of WAR and GAME as “masculine-oriented” and LIVING ORGANISM as “feminine-oriented”, but differ for the source domains of BUILDING and JOURNEY, which were viewed as “neutral” (Zeng et al. 2020: 147).

The second research question asked the extent to which the gender stereotypicality of the keywords associated with common source domains differed between male and female participants. We found that men and women differ in how masculine they view the keywords in the BUILDING, JOURNEY, COMPETITION, and PLANT source domains to be, with men viewing them as more masculine. In contrast, Study 2 demonstrated that women perceived the keywords in the source domains of JOURNEY and PLANT as even more feminine than men. This is interesting in light of Su et al.’s (2021) corpus-based study on Mandarin data which found that the dominant usage pattern of a gendered linguistic device, such as gender modification, can either confirm or contradict a stereotype. As their study was corpus-based, Su et al. (2021) were unable to ascertain whether men or women were more likely to pattern one way or the other. Future research will look at how gendered metaphors are used to potentially reinforce or confront stereotypes in various professional settings by both men and women.



In Study 3, we also found that there was a significant difference in the masculinity-femininity ratings on WAR, JOURNEY, COMPETITION, and PLANT, with women considering them more feminine than men did. These differences between how men and women view keywords associated with source domains have implications for how they may use and interpret them, which will be the focus of future studies.

Gender stereotypes are pervasive across different social situations (Ellemers 2018). One caveat of our study is that we only looked at US speakers of English. Given that gender stereotypes may differ across cultures (Obioma et al. 2022) and over time (Eagly et al. 2020), it may be worthwhile to replicate our study in other cultures and contexts to demonstrate similarities and differences in gender perceptions around the world.

In this paper, we conducted online surveys with US English speakers to collect data on the gender stereotypicality of specific keywords determined to be associated with five common metaphorical source domains: BUILDING, COMPETITION, JOURNEY, PLANT, and WAR. All metaphorical keywords were presented and rated with reference to their gender stereotypicality by 1,060 participants. We ran statistical analyses to evaluate if these keywords were rated as more associated with femininity or masculinity in order to reflect the extent to which specific source domains are gender stereotyped. We demonstrated that the keywords associated with the WAR and COMPETITION source domains were viewed as most masculine in comparison to the other source domains, and that the keywords associated with the PLANT and JOURNEY domains were viewed as being associated with femininity. In addition, the results provided herein establish a basis for future work on understanding the effect of gender stereotyping in real-world situations, including how gendered metaphors are processed (e.g., Messer and Kennison 2018) as well as the influences of gendered metaphors in judgments (e.g., Elmore and Luna-Lucero 2017).

**Research funding:** This research was supported by the Hong Kong Research Grants Council (RGC) General Research Fund Scheme (Project Number: 15602420) to the first author. The second and fourth authors would like to acknowledge the Hong Kong RGC Postdoctoral Fellowship Scheme (Reference Number: PDFS2122-5H01).

## References

- Ahrens, Kathleen & Menghan Jiang. 2020. Source domain verification using corpus-based tools. *Metaphor and Symbol* 35(1). 43–55.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Buhrmester, Michael, Tracy Kwang & Samuel D. Gosling. 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science* 6. 3–5.
- Campbell, Anne, Steven Muncer & Edward Coyle. 1992. Social representation of aggression as an explanation of gender differences: A preliminary study. *Aggressive Behavior* 18(2). 95–108.
- Charteris-Black, Jonathan. 2004. *Corpus approaches to critical metaphor analysis*. Basingstoke: Palgrave Macmillan.
- Charteris-Black, Jonathan. 2006. Britain as a container: Immigration metaphors in the 2005 election campaign. *Discourse & Society* 17(5). 563–581.
- Charteris-Black, Jonathan. 2011. *Politicians and rhetoric: The persuasive power of metaphor*. Basingstoke: Palgrave Macmillan.
- Crawford, Jarret T., P. Leynes Andrew, Christopher B. Mayhorn & Martin L. Bink. 2004. Champagne, beer, or coffee? A corpus of gender-related and neutral words. *Behavior Research Methods, Instruments, & Computers* 36. 444–458.
- Deignan, Alice. 1995. *COBUILD English guides 7: Metaphor dictionary*. London: Harper Collins.
- Eagly, Alice H., Christa Nater, David I. Miller, Michèle Kaufmann & Sabine Sczesny. 2020. Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist* 75(3). 301–315.
- Ellemers, Naomi. 2018. Gender stereotypes. *Annual Review of Psychology* 69. 275–298.
- Elmore, Kristen C. & Myra Luna-Lucero. 2017. Light bulbs or seeds? How metaphors for ideas influence judgments about genius. *Social Psychological and Personality Science* 8(2). 200–208.
- Friedman, Susan Stanford. 1987. Creativity and the childbirth metaphor: Gender difference in literary discourse. *Feminist Studies* 13(1). 49–82.
- Gidengil, Elisabeth & Joanna Everitt. 2003. Talking tough: Gender and reported speech in campaign news coverage. *Political Communication* 20(3). 209–232.
- Hines, Caitlin. 1999. Foxy chicks and Playboy bunnies: A case study in metaphorical lexicalization. In Masako K. Hiraga, Chris Sinha & Sherman Wilcox (eds.), *Cultural, typological and psychological perspectives on cognitive linguistics*, 9–23. Amsterdam: Benjamins.
- Keith, Melissa G., Louis Tay & Peter D. Harms. 2017. Systems perspective of Amazon Mechanical Turk for organizational research: Review and recommendations. *Frontiers in Psychology* 8. 1359.
- Koller, Veronika. 2004a. Businesswomen and war metaphors: "Possessive, jealous and pugnacious"? *Journal of Sociolinguistics* 8(1). 3–22.

- Koller, Veronika. 2004b. *Metaphor and gender in business media discourse: A critical cognitive study*. Basingstoke: Palgrave Macmillan.
- Kövecses, Zoltan. 2010. *Metaphor: A practical introduction*. Oxford: Oxford University Press.
- Lakoff, George & Mark Johnson. 2003. *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Messer, Rachel H. & Shelia M. Kennison. 2018. Individual differences in the processing of novel, gender-stereotyped metaphors. *North American Journal of Psychology* 20(1). 37–54.
- Misersky, Julia, Pascal M. Gygax, Paolo Canal, Ute Gabriel, Alan Garnham, Friederike Braun, Tania Chiarini, Kjellrun Englund, Adriana Hanulíková, Anton Öttl, Jana Valdrova, Lisa Von Stockhausen & Sabine Sczesny. 2014. Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods* 46(3). 841–871.
- Murphy, Peter F. 2001. *Studs, tools, and the family jewels: Metaphors men live by*. Madison, WI: University of Wisconsin Press.
- Obioma, Ihuoma F., Tanja Hentschel & Alina S. Hernandez Bark. 2022. Gender stereotypes and self-characterizations in Germany and Nigeria: A cross-cultural comparison. *Journal of Applied Social Psychology* 52(8). 764–780.
- Paolacci, Gabriele, Jesse Chandler & Panagiotis G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5(5). 411–419.
- Philip, Gill. 2009. *Non una donna in politica, ma una donna politica*: Women's political language in an Italian context. In Kathleen Ahrens (ed.), *Politics, gender and conceptual metaphors*, 83–111. Basingstoke: Palgrave Macmillan.
- R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: <https://www.r-project.org/>.
- Schubert, Thomas W. 2004. The power in your hand: Gender differences in bodily feedback from making a fist. *Personality and Social Psychology Bulletin* 30(6). 757–769.
- Scott, Graham G., Anne Keitel, Marc Becirspahic, Bo Yao & Sara C. Sereno. 2019. The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods* 51. 1258–1270.
- Su, Qi, Pengyuan Liu, Wei Wei, Shucheng Zhu & Chu-Ren Huang. 2021. Occupational gender segregation and gendered language in a language without gender: Trends, variations, implications for social development in China. *Humanities and Social Sciences Communications* 8(1). 1–10.
- Velasco Sacristán, María Sol. 2005. A critical cognitive-pragmatic approach to advertising gender metaphors. *Intercultural Pragmatics* 2(3). 219–252.
- Winter, Bodo, Sarah E. Duffy & Jeannette Littlemore. 2020. Power, gender, and individual differences in spatial metaphor: The role of perceptual stereotypes and language statistics. *Metaphor and Symbol* 35(3). 188–205.
- Zeng, Winnie Huiheng & Kathleen Ahrens. 2023. Corpus-based metaphorical framing analysis: WAR metaphors in Hong Kong public discourse. *Metaphor and Symbol* 38(3). 254–274.
- Zeng, Winnie Huiheng, Dennis Tay & Kathleen Ahrens. 2020. A multifactorial analysis of metaphors in political discourse: Gendered influence in Hong Kong political speeches. *Metaphor and the Social World* 10(1). 141–168.

---

**Supplementary Material:** This article contains supplementary materials (<https://doi.org/10.1515/lingvan-2024-0021>).